ABSTRACT

Title of Dissertation: Antecedents and Effects of Retail Shelf Stockouts

Hadiye Pinar Celebi, Doctor of Philosophy, 2019

Dissertation directed by: Associate Professor Philip T. Evers

Department of Supply Chain Management

Retail shelf availability research has been limited by the inability to measure stockouts. Not being able to fully capture stockout occurrences has led to studying either the effects of stockouts or their antecedents. It has also led to using various fundamentally different stockout attributes as measures across studies. The relationship between stockout attributes is not clear, making it difficult to have a consensus on either the drivers or the impact of stockouts.

This thesis considers both antecedents and effects of stockouts by incorporating actual stockout events under two different risk pooling methods. The first set of models simulate stockout-based customer switching (the inventory effect) to study pooling by substitution for a retailer setting service level goals for two products. The second set of models study pooling by postponement, termed “instore logistics postponement,” using archival data from a new shelf sensor technology that captures actual stockout events. An extension to the second part of this study examines the nonlinear relationship between stockout attributes.

Both parts of the dissertation contribute to the stockout literature in different ways. The simulation work contributes towards reconciling opposing views on the performance effect of risk pooling through substitution, also showing how different performance measures may accentuate or mask the impact of stockouts. The shelf technology work contributes to logistics postponement by studying how a two-tier inventory within the store may affect stockouts along more than one stockout attribute, and whether less frequent but longer stockouts are linked to better performance than shorter but more frequent stockouts.

ANTECEDENTS AND EFFECTS OF RETAIL SHELF STOCKOUTS

By

Hadiye Pinar Celebi

Thesis submitted to the Faculty of the Graduate School of the

University of Maryland, College Park in partial fulfillment

of the requirements for the degree of

Doctorate of Philosophy

2019

Advisory Committee:

Professor Philip T. Evers, Chair
Professor Francis Alt
Professor Thomas Corsi
Professor Niratcha Tungtisanont
Professor Wayne H. Slater

Table of Contents

List of Figures

List of Tables

Acronyms and Notation

ABM	Agent-based model
ABMS	Agent-based model simulation
ASL	Aggregate service level
AVE	Average variance extracted
B	Reorder point including safety stock, in units
CAP	Shelf capacity, total units allocated for a product (volume)
CPS	Case pack size
CPG	Consumer packaged goods
CU	Consumer unit
cv	Coefficient of variation
D	Daily demand
DC	Distribution center
DES	Discrete-event simulation
DSI	Demand and supply integration
E (M>B)	Expected lead time stockout in units
FRH	Fill rate heuristic
ILP	Instore logistics postponement
IRO	Instore replenishment (or indirect or secondary) stockout
KS	Kolmogorov-Smirnov
L	Lead time
LR	Likelihood ratio
M	Actual demand during lead time
MSE	Mean squared error
OLS	Ordinary least squares
OOS	Out-of-stock
OPQ	Order packing quantity
OSA	On-shelf availability
POS	Point-of-sale
P/S	Postponement and speculation
Q	Fixed order quantity
R	Average annual demand
RCM	Retail category management
RMSE	Root mean square error
S	Safety stock
SC	Supply chain
SCM	Supply chain management
SHO	Shelf stockout
SKU	Stock-keeping unit
SL	Service level
SO	Stockout
STO	Store stockout
TSL	Target fill rate
WS	Willingness to substitute

Chapter 1 Introduction

1.1 The retailer's unique position in the supply chain

Retailers uniquely add value to the supply chain (SC) with customer service tasks not present with other SC members. Retailers provide an assortment of consumer-packaged goods[1] (CPG), break bulk, hold inventory, and provide services for customers (Levy & Weitz, 2004, p. 7). The assortment a retailer provides brings together items, usually from different manufacturers, leading to a "depth of assortment" within a category of products. Different product categories are the "assortment breadth" or "product variety" of the retailer's offerings[2]. Since customers may buy on impulse (Xiao & Nicholson, 2013) or may buy just one unit of a product, the store functions as the point of breaking bulk and attracts customers through assortment depth and breadth. Offering an array of products available in single or a small number of multiple units makes the retailer the primary SC member whose inventory reduces the end user's cost of holding at home. Finally, store services vary from those in other SC members. Buyers perform a type of self-service,

---

[1] A *product* is a tangible good that is for sale. A *consumer* is an end user who can buy a product in various forms or sizes. For example, a consumer may buy mint toothpaste as a large or small tube, stand-up bottle pump, or in multiple tube packages. For multiple tube packs, a single *consumer unit* with two large tubes of mint toothpaste is the same product volume as two single large tube consumer units of the same product. While they are the same product (mint toothpaste) the size or packaging differences result in offering the consumer different *items*. Items are tracked by retailers as SKUs or *stock-keeping units* so that all sellable items, whether they are the same product or not, are distinguishable from one another for inventory management purposes. Such packaged products are commercial items known as *consumer packaged goods (CPG)*. CPG are repeatedly purchased at a lower price and consumed in a shorter period of time than durable goods such as cars or home furniture and appliances.

[2] "Product variety" at higher levels of the supply chain usually refer to the number of unique items produced, stored or tracked by that firm; the equivalent of "product variety" in a retail store within a product category is called "assortment depth." A retailer's product variety refers to the number of different product categories it offers and is also called "assortment breadth." The total unique number of items a store tracks would be the sum of all depths across its breadth.

where they take on the tasks of walking around the inventory (the store) and pulling items from the shelves themselves. This self-service provides customers the opportunity to browse, touch and feel products. Furthermore, the store environment allows for personalized feedback and information while browsing, cash payments for purchases, immediate gratification from getting the product as soon as it is purchased, and even entertainment and social experiences beyond the purchase transaction (Levy & Weitz, 2004, pp. 81-82).

Retailers also face unique risks related to their tasks that are largely overlooked in terms of strategic importance to the SC (Hochrein, Glock, Bogaschewsky, & Heider, 2015, p. 268). Since a retailer breaks bulk, it has a greater number of customers who have lower order quantities than the customers of higher supply chain members. Frequent smaller orders at the manufacturer level means stockout frequency will increase because of transportation variability (Bowersox & Closs, 1996). For a retailer, transportation variability at the manufacturer level is akin to instore logistics activities: receiving orders from distribution centers (DCs) (Cardos & Garcia-Sabater, 2006), instore product handling (Curseu, Woensel, Fransoo, Donselaar, & Broekmeulen, 2009), and shelf replenishment from the backroom (Eroglu, Williams, & Waller, 2013). Especially since product handling costs within the store are generally higher than inventory costs (van Donselaar, van Woensel, Broekmeulen, & Fransoo, 2005), current research focuses on how to improve operational efficiency *within* the store, instead of interpreting these tasks as strategically important for the supply chain (Samli et. al 2005) beyond concepts of supply chain flexibility (Jafari, Nyberg, & Hilletofth, 2016). While supply chain management's (SCM) traditional view of retailers as "simply intermediaries" now

recognizes large retailers in a position of greater power, the focus on such research remains on how upstream member activities respond to retailer pressure (Zentes, Morschett, & Schramm-Klein, 2007). SCM research lags behind practitioner efforts to establish instore activities as having strategic importance in the SC (Randall, Gibson, Defee, & Williams, 2011).

Retailing Strategy	Administrative Tasks
Determine retail strategy: target market, format, how to achieve sustainable competitive advantage	*Promote firm, items, services*: plan communication programs and budget, select media outlets, plan special promotions and specific displays, manage public relations
Determine financial strategy: evaluate investment opportunities, allocate resources, assign accountability, set profit and turnover performance goals, adjust strategy	*Manage human resources:* develop personnel policies, hire and train managers, plan careers, keep employee records
Determine location, organization, and flow systems for information and goods: economies of scale versus cannibalization, accessibility, trade areas, estimating regional demand	*Distribute merchandise:* find warehouses, receive items, store, and ship them to stores, return to vendors
Determine customer relationship approaches: loyalty programs, market basket analysis, lifetime value, customer retention, personalization, community development	*Establish financial control:* provide financial performance reports, forecast sales, raise capital from investors, bill customers, provide credit, monitor cash flow

Table 1-1 Strategic retail decision making areas (Levy & Weitz, 2004)

Similar to other SC members, retailing theory has traditionally categorized tasks into strategic and tactical areas. Strategic decisions are typically considered to be those that commit significant resources to develop a long-term competitive advantage, while tactical decisions are said to impact operational efficiency for a shorter term. Strategic decisions are comprised of two components: retailing strategy and administrative tasks as listed in Table 1-1. Attaining sustainable competitive advantage with these strategic decisions requires any combination of the following: customer loyalty, location, human resource management, distribution and information systems, unique merchandise, vendor relations, and customer service (Levy & Weitz, 2004). Customer service in terms of basket analysis involves deciding on which categories to include in assortment breadth (product variety). Assortment breadth decisions are similar to whether or not the store implements information tracking technology or where it locates stores in that they all involve longer planning horizons. Similarly, merchandise distribution in administrative strategic decisions are involved with product flow outside of the store, with planning horizons that are long-term.

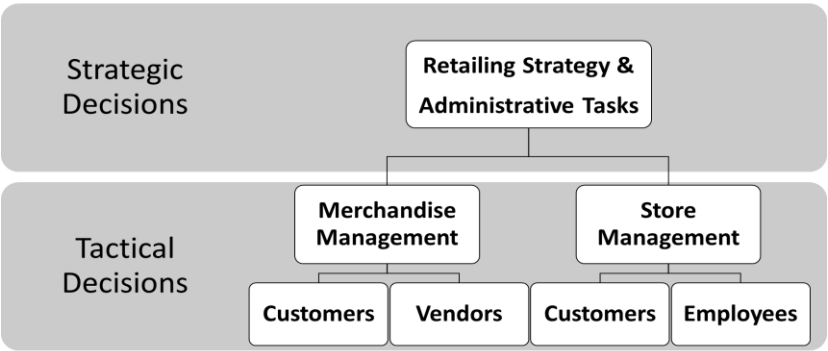


Figure 1-1 Strategic retail decisions determine store tactical goals

Traditionally, strategic decisions determine tactical tasks which fall under two areas: merchandise management and store management, as in Figure 1-1, (Levy & Weitz, 2004). Merchandise management can be viewed as stores focusing on customers through

vendor-related tasks, while store management focuses on serving customers with store employee-related tasks. In terms of merchandise management, stores decide on sourcing options while considering how customers are influenced by the store's assortment depth. Known as demand reshape (Eynan & Fouque, 2003) the supply tasks that have an effect on customer demand in relation to vendor issues include: choosing a profitable assortment (Chernev, 2011), forecasting sales (Ali & Pinar, 2016) to determine order size and timing information, implementing specific buying systems (Williams & Waller, 2011), setting and adjusting prices (Grewal, et al., 2010), and offering promotions (Laroche, Pons, Zgolli, Cervellon, & Kim, 2003). On the other hand, store management is "the primary interface between planning, coordination and operations" (Bowersox & Closs, 1996, p. 201). Tactical decisions of store management include managing employees (Plouffe, Bolander, Cote, & Hochstein, 2016), controlling store costs (Betancourt & Gautschi, 1988), maintaining store facilities, preventing shrinkage (Mishra & Prasad, 2006), presenting merchandise in terms of how they are displayed or where in the store they are located (Lam, 2001), and providing service such as responding to customer complaints and answering product questions (Snyder, Witell, Gustafsson, Fombelle, & Kristensson, 2016).

More recently, Hubner, Kuhn, and Sternbeck (2013) have proposed a retail operations framework (Table 1-2) which moves away from the traditional strategic versus tactical decision areas, choosing to categorize by planning horizons and explicit mapping of information flows. While the physical flow of products is said to go from procurement (P) to warehousing (W), to distribution (D) and then to sales (S), the information flow is said to go back and forth between each of these subcategories as well as up the planning

horizon, from short- to mid-term and mid- to long-term planning (Hubner, Kuhn, & Sternbeck, 2013, p. 517). Examples of long-term planning decisions (Table 1-2) are the number and location of DCs and stores as well as which vendors the retailer works with, or decisions which have to do with the retail network structure. Decisions at this level of planning horizon affect a retailer's tasks indefinitely, or until additional long-term decisions once again alter the network structure. Mid-term planning guides a retailer's activities for 6 to 24 months (Hubner, Kuhn, & Sternbeck, 2013), with decisions on how the physical flow will take place within an established network structure, which products to carry for which target market, and the overall estimated labor needs for each store. Short-term planning involves decisions about what to do that very business (store) day, such as personnel shifts and promotional activities. Planning horizons directly involving tasks within the retail store are all of the sales (S) subcategory in Table 1-2: outlet planning in the long-term horizon, master category planning and instore planning in the mid-term horizon, and instore fulfillment in the short-term horizon.

The first noticeable difference between traditional and recent retail operations framework stems from separating activities into 3 categories (by horizon) instead of 2 (strategic versus tactical). Long-term planning horizon activities in Table 2 still fit into the strategic activities listed in Table 1-1: determining location and organization of stores

Planning Horizon	Long-term	Procurement logistics (P)	Sourcing; supplier selection and contracting; inbound logistics
		Warehouse design (W)	Number, locations, function and types of warehouses
		Distribution planning (D)	Physical distribution structure
		Outlet planning (S)	Store type and location; strategic layout planning
	Mid-term	Product segmentation and allocation (PWD)	Planning of product delivery modes ; assignment product-warehouse-outlet; selection of dispatch; building segments with order patterns; selection of transportation means (modes)
		Inbound planning (P)	Supplier order management; Master inbound route planning
		Production planning (W)	Capacity and personnel; warehouse management
		Distribution planning (D)	Outlet order management; master outbound route planning
		Master category planning (S)	Category sales planning; assortment architecture; promotion planning
		Instore planning (S)	Store personnel planning; store logistics planning
	Short-term	Order fulfillment (P)	Supplier order dispatching; operative inbound route planning and transportation scheduling; inbound ramp management
		Production scheduling (W)	Personnel schedules; sequencing and job release
		Transport planning (D)	Outlet order dispatching; operative outbound route and transportation scheduling; outbound ramp management
		Instore fulfillment (S)	Short-term sales planning; store personnel scheduling; short-term instore logistics management

Table 1-2 Compilation of the operational framework proposed by Hubner, Kuhn and Sternbeck (2013)

and the physical distribution structure, involving warehouse number and location decisions. Similarly, short-term planning horizon activities in Table 1-2 still fit into the tactical activities described as a part of merchandise and store management illustrated in Figure 1-1: managing employees, adjusting prices, maintaining store facilities. However, in the mid-term planning horizon the activities do not all belong to the traditional strategic or tactical categories. For example, in Table 1-2, "product segmentation and allocation" fits into the strategic task in Table 1-1 of distributing merchandise, while "master category planning" in Table 1-2 fits into the tactical task of merchandise management as illustrated in Figure 1-1. Similarly, store personnel planning is traditionally seen as a tactical activity of store management whereas it is a component of "instore planning" in the mid-term horizon. This illustrates how the strategic value of mid-term planning decisions which take place in purchasing (P), warehousing (W), and distribution (D) is traditionally accepted, whereas instore planning and master category planning activities at the store level (S) have been considered tactical activities having to do with operational efficiency instead of competitive advantage.

A second difference when comparing the traditional versus recent framework in retail operations is explicitly accounting for information flow. Information flows from short- to long-term planning in the proposed framework while the traditional view simply posits that strategic decisions affect tactical tasks. The new framework implies that instore fulfillment information (item sales) must be used for master category and instore planning, which are then both needed for outlet planning. Similarly, information flow must occur within planning horizons between the retailer's business units. For example, retail personnel doing instore fulfillment planning should exchange information with

personnel who plan transportation from the retail DC to the store, with personnel scheduling DC operations, and with personnel who order from vendors and plan receipt of goods to DC. In the traditional view, there is no explicit accounting for information flow or sharing information between business units and across firms. This lack of a "formal feedback loop" leads to poor "demand responsiveness" and misalignment between different planning horizons (Dreyer et al. 2017). In this way the traditional view is one of a supply chain, whereas the newer framework is of a "demand and supply chain" (Dreyer et al. 2017).

Demand and supply integration (DSI) relies on integrated information and the understanding that bringing together demand and supply processes is strategically important in terms of customer value (Esper, Ellinger, Stank, Flint, & Moon, 2010). DSI defines demand management as a firm's processes involving the firm's customers with supply management processes involving the firm's suppliers (Esper, Ellinger, Stank, Flint, & Moon, 2010, p. 7). DSI has been studied for four decades and is also referred to as demand-chain management and demand-supply chain management (Bonomi Santos & D'Antone, 2014). In their systematic literature review, Bonomi Santos and D'Antone find that integrated activities have so far been empirically shown to improve performance outcomes for firms, between departments and within departments (Bonomi Santos & D'Antone, 2014, p. 1020). Given its link to performance improvement, it makes sense that practitioners are in a search for the best approaches to accomplish DSI for different types of products (Boone & Ganeshan, 2015) and scholars continue to push for more collaborative decision-making processes (Fawcett, Waller, & Fawcett, 2010).

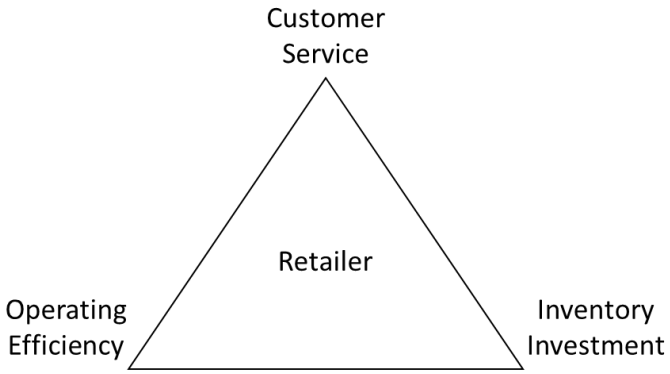


Figure 1-2 A retailer's opposing objectives. Adapted from (Tabucanon & Farahani, 1985)

Even in a fully integrated retail supply chain, where all processes are planned in an integrated system, aligning a store's goals still needs to consider how these goals may oppose each other, as illustrated in Figure 1-2 (Tabucanon & Farahani, 1985). Operational efficiency refers to how well the firm's constraints or inputs of labor, space, stock, time and money are utilized in terms of output and include measures such as unit or dollar sales per personnel, per square foot, per selling time or per costs incurred. The firm's objective is to maximize operational efficiency and this measure is often used to gauge firm profitability (Kumar, Anand, & Song, 2017). Customer service refers to the quality or degree to which customer demand is met and is generally measured in terms of firm or product revenue or market share, units sold or proportion of demand fulfilled (fill rate), which is the most common measure in the retail industry (Teunter, Syntetos, & Babai, 2017). The firm's objective or goal is to maximize such customer service measures. Finally, inventory investment is the actual number of units held or dollars invested in keeping items (SKUs) present at a store. There are various measures of inventory level or value (Stangl & Thonemann, 2017): the amount of safety stock, the

shelf capacity, the target service level or fill rate, the reorder point, and inventory cycle order size. The firm's objective is to minimize inventory investment, so it is one of the opposing objectives in Figure 1-2.

Inventory management and deployment goals oppose each other because tasks which work towards one goal often work against another goal. High customer service goals push a store to higher inventory investment and lower operating efficiency by stocking more items in terms of variety and quantity (Ton & Raman, 2010), hiring more personnel (Hise, Kelly, Gable, & McDonald, 1983) for peak hours, or expanding space allotted for items in the store (Arndt & Olsen, 1975). High efficiency (productivity or profitability) goals push the store to larger less frequent orders and fewer personnel, likewise requiring higher inventory investment and possibly less flexibility in responding to shifts in customer demand (Balakrishnan, Pangburn, & Stavrulaki, 2004). Similarly, inventory minimization goals push the store towards greater inefficiency because of frequent or rushed shipments, increased labor for more frequent shelf replenishment, as well as lower customer service performance (Eroglu & Hofer, 2011). As these examples illustrate, each one of the three inventory management and deployment goals can be considered antecedents (drivers or retailer inputs) or effects (performance outcomes) of retailer decisions or of phenomena. Regardless of how they are considered for a specific research question, each goal involves tasks that work against the other two goals.

Business logistics theory presents methods to better balance these opposing objectives or goals through five areas of risk pooling: transportation, procurement, production, inventory storage, and sales and distribution (Oeser, 2010). The overarching

idea behind all of them is that when a firm considers different streams of the same type of risk as one combined source of uncertainty, then the overall variability is lower than the variability observed within each stream. The source of variability may stem from uncertainty in demand, supply (lead time), or both. Transportation pooling refers to transshipments and cross-filling, while procurement refers to order splitting and centralized ordering. Pooling in the production area includes form postponement, component commonality and capacity considerations, while inventory storage is concerned with centralizing warehousing tasks.

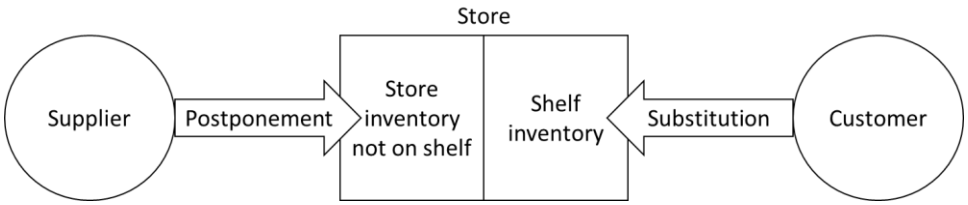


Figure 1-3 Sales and distribution pooling mechanisms considered in this dissertation

The last area of risk pooling, sales and distribution, include two mechanisms which provide the theoretical lens of this dissertation: product substitution and postponement (Figure 1-3). Product substitution occurs when a customer initially prefers one item but ultimately buys another. Time postponement occurs in tandem with speculation, where items are held close but purposefully delayed from moving forward to customers (Oeser, 2010). In the retail context, this refers to purposefully holding more inventory within the store than the shelf capacity will hold, to be able to quickly replenish the shelf with the additional inventory already within the store. This dissertation defines this strategy of merchandise management and shelf replenishment as *instore logistics postponement*. Shelf replenishment is also at the heart of product substitution in two

ways. First, a customer's initially preferred product may not be available on the shelf because it has not been replenished before customer arrival to the shelf. Secondly, if the customer decides to substitute a product for the initially preferred one, then the alternate product must be available (must have been replenished) on the shelf to be a substitute. While pooling through postponement deals with store actions leading to product availability, pooling through substitution takes place with customer actions after facing product unavailability (Figure 1-3).

Product availability for a retailer differs from inventory concerns of other SC members because of the store's unique role with its customers. Customers have been allowed to walk around and select items from store shelves themselves without relying on a store clerk to bring a product upon request since 1916 (Regan, 1960). This shift to self-service made it possible for a product to be stocked out or demanded by customers without the knowledge of store personnel. The practitioner's lack of knowledge of when an item stocks out has also made it difficult to gauge an item's original (initial or direct) demand levels and pattern. Stores have focused on increasing on-shelf availability while having a constrained overall shelf capacity and wanting to keep inventory investment levels low. Research streams in these topics are reviewed in the remainder of this chapter starting with retail-based contexts and followed by pooling methods of substitution and postponement. The chapter ends with research deficits, corresponding questions and the dissertation framework.

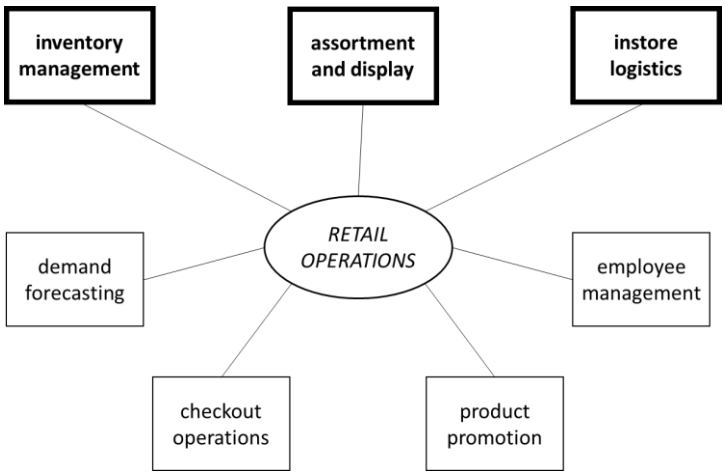


Figure 1-4 Dissertation areas in bold, graphical representation of Mou, et al. (2018)

1.2 Retail-specific research

This section introduces retail-based research streams. These research streams exist only within retail operations instead of being a retail context or setting for the study of a general business theory or phenomenon. The main purpose of this section is to present terminology used in subsequent sections of product availability, substitution and postponement. A more encompassing higher-level literature review of retail operations is provided by (Mou, Robb, & DeHoratius, 2018). Generated to summarize their overview of retail operations, Figure 1-4 illustrates the three areas of this dissertation. Under inventory management, Mou et al. (2018) include inventory-dependent demand studies. This dissertation looks at substitution pooling in a continuous review inventory system, when demand switches from one item to another due to stockout of the initially preferred item. Under assortment and display in Figure 1-4 is shelf allocation work. Under instore logistics is research on shelf replenishment as well as backroom use. This dissertation

combines shelf allocation, backroom use and shelf replenishment to introduce the concept of instore logistics postponement as a type of pooling by postponement and speculation.

1.2.1 Retail category management

Retail category management (RCM), an approach which started in the 1990s (ACNielsen, 2006) , is a subset of retail operations, though a consensus hasn't been reached on which retail decisions RCM encompasses. Generally speaking RCM focuses on how much space to allot, if any, for each item in a product category. A product category is a group of products shelved together at a store, which "consumers perceive to be interrelated and/or substitutable" (Kurtulus & Toktay, 2011, p. 47). Including an item within a product category is an assortment question, while choosing how many units of limited shelf space to dedicate to each item in a fixed assortment is a shelf allocation question. These two questions have generally been handled as separate research streams (Hubner & Kuhn, 2012) though recent work directs scholars to classify them together as the "assortment and display" subcategory of retail operations (Mou, Robb, & DeHoratius, 2018). Planogram allocation[3] is a third area of assortment and display research and also assumes a fixed assortment but simultaneously varies shelf allocation for each item by considering how much each item's sale will vary depending on the space allotted for it (elasticity) and how much each item's sale will vary depending on the space allotted for another item (cross-elasticity) (Frontoni, Marinelli, & Rosetti, 2017). While shelf space

[3] Planograms are diagrams with drawings or pictures of every item in a product category as it should be arranged on store shelves. How the items "should" be arranged is planogram allocation and reflects the marketing plan for that product category.

literature streams have substitution criteria partially integrated in their work, assortment planning partially integrates shelf space constraints (Hubner & Kuhn, 2012, p. 202).

A recent study (Karampatsa, Grigoroudis, & Matsatsinis, 2017) on RCM categorizes retail decisions by customer switching behavior instead of item characteristics. Customers can switch from an initially preferred product to an available substitute because: the store does not carry the preferred item (assortment-based switching), the price of the preferred item is higher than another product (price-based switching), the amount of space dedicated to the preferred item differs from that of another item (space-based switching), or the preferred item is currently not available on the store shelf (stockout-based switching). Karampatsa et al. (2017) classify prior RCM work by which factors the retailer must consider for each customer behavior, as in Figure 1-5. For example, for inventory management research on stockout-based switching, scholars so far have considered retailer decisions on order size and timing of placing and receiving orders, the upper limit to total shelf space, and the store budget in terms of ordering or holding power. For shelf-space planning, in addition to all of these factors they have also considered the ability of store personnel to restock store shelves. Besides inventory management and shelf-space planning, the Karampatsa et al. (2017) RCM framework also includes assortment planning and price planning.

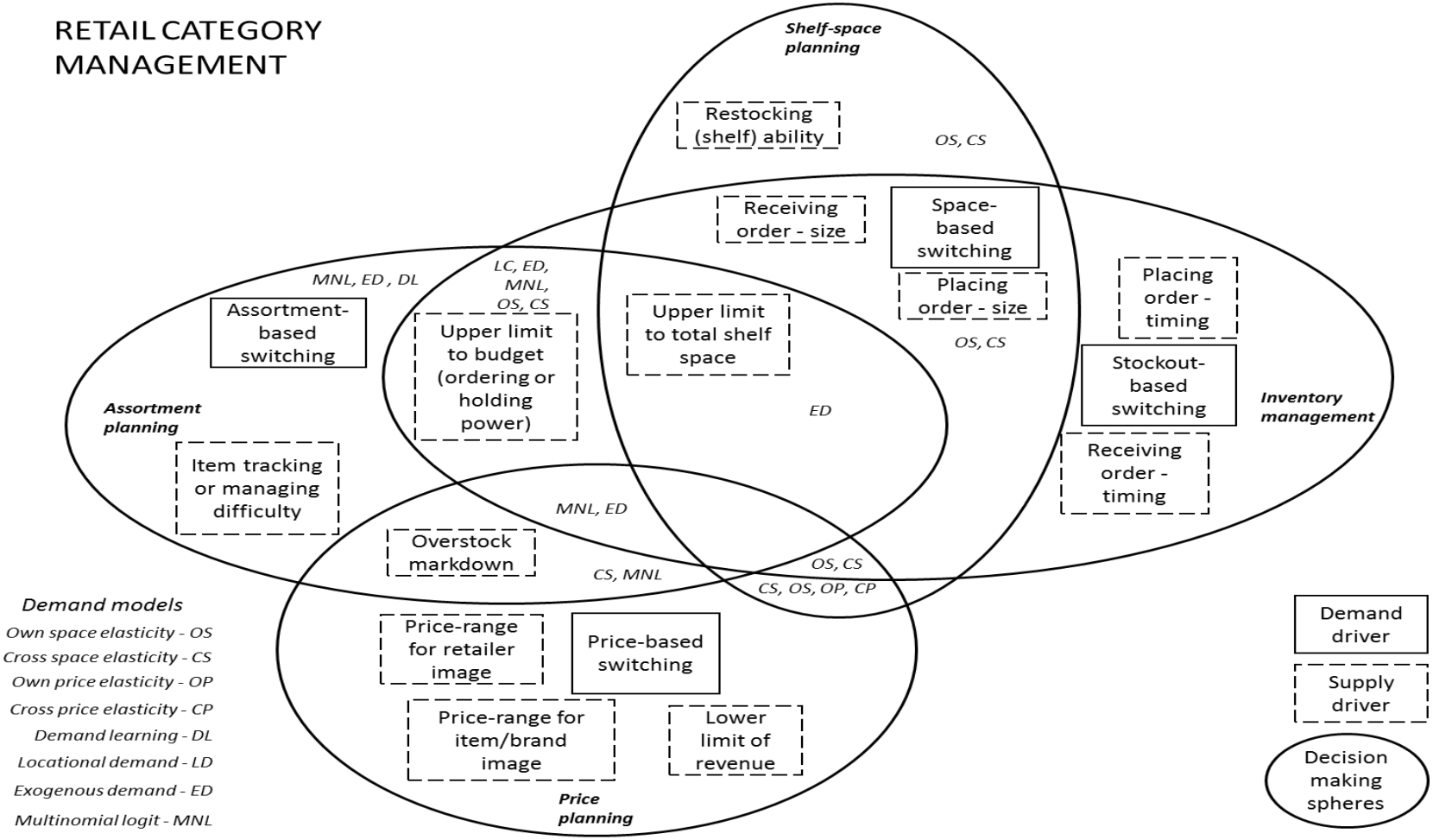


Figure 1-5 Created using works reviewed in Karampatsa et al. (2017)

Both analytic and empirical work reviewed by Karampatsa et al. (2017) is included in Figure 1-5, to illustrate the lack of agreement between RCM scholars on what constitutes RCM research. The demand models used in analytic work all are based on either assortment or shelf-space planning. Models based on assortment-based switching use multinomial logit, exogenous demand and demand learning models while shelf-space planning work uses own and cross space elasticity since demand is a function of shelf space. The shelf to sales ratio has been studied empirically too, using experiments that test how consumer attention to the product category increases with increased space, or how the vertical positioning of the item is linked to different sales performance. According to Karampatsa et al. (2017) empirical work in assortment is usually based on field studies and finds that increasing the assortment does not necessarily improve sales, and the consumer perception of the assortment may be skewed by allotted space and planogram-based arrangement on store shelves. Finally, product availability research overlaps with RCM when the study focuses on stockout-based switching behavior to "verify that substitution behavior is inherent in the retail environment and needs to be reflected in category planning optimization" (Karampatsa, Grigoroudis, & Matsatsinis, 2017, p. 41).

While there is a lack of agreement on the scope and categorization of RCM research, it encompasses a broader level of analysis than most retail operations research. "[U]nlike the traditional brand-by-brand or SKU-by-SKU focus…[RCM]…emphasizes the management of a product category as a whole, allowing the decision maker to take into account the customer response to decisions made about substitutable or interrelated products" (Kurtulus & Toktay, 2011). In contrast, shelf management focuses on brand

sales or item sales, and while it may consider the effects of substitution (Borin, Farris, & Freeland, 1994), its scope of research is on space allocation, shelving decisions, and sales of a single focal brand or item. Similarly, instore logistics research studies the shelving and replenishment of multiple items but does not consider the interrelatedness between the items in terms of customer purchasing behavior. Therefore, while RCM is at the product category level of analysis, shelf management is at the brand or SKU level and instore logistics is at the labor unit of analysis, such as total stacking time (van Zelst et al. 2009) (Curseu et al. 2009).

Retailers manage the overall product category in an eight-step process: category definition, role, assessment, scorecard, strategies, tactics, implementation and review (ACNielsen, 2006, p. 64). Category definition is determining which items to carry in the product category, or deciding on the assortment depth. The category role (destination, routine, seasonal or convenience category) is the purpose of the category compared to all the other categories in the store, so it is determining the category's place in the assortment breadth (ACNielsen, 2006, p. 78). An assessment is then made of each item, by SKU, by brand, segments and subcategories. The assessment is made from four perspectives: the market (market share and benchmarks), the consumer (purchase behavior and profile), the supplier (share and efficiency), and the distributor (contribution and productivity) (ACNielsen, 2006, p. 98). The scorecard is the measurable sales or availability rates for the entire category. After the goals have been determined, category strategies is the step of choosing what sort of supply, marketing efforts, and in-store service is needed for the category overall. Category tactics are the actual things that need

to be done for that category's role, scorecard and strategies. The planned tactics are implemented and a review of the category measures and monitors category progress.

The eight-step process reflects a "paradigm shift" in retail that alters typical retail strategies. Strategies are normally considered to be the traditional strategy of pushing goods to customers, or focusing on one of the following instead to pull customers: product variety (Ton & Raman, 2010), high tech solutions (Dubelaar et al. 2001) , customer loyalty (Terblanche, 2017), and customer service (ACNielsen, 2006). The shift instead alters the focus from these retail strategies to the following six approaches (ACNielsen, 2006, pp. 313-314):

- Assortment/shelf plans → Category marketing plans
- Average customer → Segmented customers
- Average market/store → Competitive clusters/formats
- Independent category plans → Integrated marketing plans
- Fair share gaps by retailer geography → Share-of-wallet by competitive clusters
- Single category captain → Lead and validate category captains

The fair share gap refers to "the difference between what sells and what you could sell" (p. 80). It considers what proportion of the customers in a trading area the retailer could be capturing. The shift to share-of-wallet reflects a consumer-centric view. It considers what proportion of a customer's purchases of the product category are being made through the retailer. To shift to any of the above integrated and consumer-centric approaches requires increased data collection and sharing between business units (Eroglu, Williams, & Waller, 2013) and SC members (Fawcett & Magnan, 2002).

1.2.2 Fast moving consumer goods

Separating goods by how quickly they sell over a given period of time caught the attention of store managers and accounting scholars almost a century ago (Pogson, 1923). Accountants recommended auditing, which is checking inventory records against actual physical inventory, more often for fast moving goods since their quicker movement could mean a greater inaccuracy in records than slow-moving items over the same time period. Keeping inventory records accurate through auditing was preferred over "waste and unnecessary expense" from obsolete stock written off at a loss, stockouts from pilferage and product misplacement, interest charges for capital tied up in inventory, excess inventory costs, handling and re-handling of excess inventory, and cost of space used up by excess inventory (Pogson, 1923, p. 437). Overage, underage and holding costs were perhaps not explicitly accounted for by store managers a century ago, but there was a debate about where in the store to place fast-moving items since these goods were purchased either frequently or by many customers. If placed in "fast-moving space" (David, 1922) in the store, customers could quickly be served and leave, whereas other spaces in the store could pull customers into the store, to spend more time and perhaps buy other items along with the initially preferred fast-moving item.

Today, despite at least 25,000 peer-reviewed articles involving the fast moving consumer goods (FMCG) industry, there is no standardized classification method of items into either fast- or slow-moving consumer goods. FMCG is a subset of CPG and is distinguished by higher turnover. Inventory turnover is a sales and distribution performance measure and refers to the cost of goods sold of an item over its average

inventory during that time period (Gaur, Fisher, & Raman, 2005). Instead of a ratio, fast-moving goods can also be identified by the number of consumer units sold within a unit of time. Stores renew their FMCG inventory more often than other products for two main reasons. The first is that the FMCG items are more popular in terms of customer demand and theft. Secondly, some FMCG items are perishable and their unsold consumer units must be removed from the shelf instead of remaining like items with longer shelf life. While perishable goods are beyond the scope of this dissertation, there is a literature review of perishable goods research and a call for research (Mou et al., 2018) into how the short shelf life of an item influences instore operations. Overall, it is not clear how CPG should be classified as either slow- or fast-moving, or if there is also a third group in between them (normal-moving) or at zero movement (non-moving). It is not clear whether an entire product category is slow- or fast-moving or if item-level categorization is possible. Classifications are largely retailer- and study-based (Wintle & Patch, 2003). For example, one study identifies a slow-moving item as "those SKUs with 12 or fewer movements a year" (Johnston, Boylan, & Shale, 2003, p. 834), where a movement is neither turnover nor units sold, but transaction frequency, essentially the number of customer visits which result in a sale. ABC analysis can also be used to classify items by their sales volume, and while this process is a standardized one, it is extremely specific to the store (Gelders & van Looy, 1978). Any implications of such research for upper level SC members may not be generalizable.

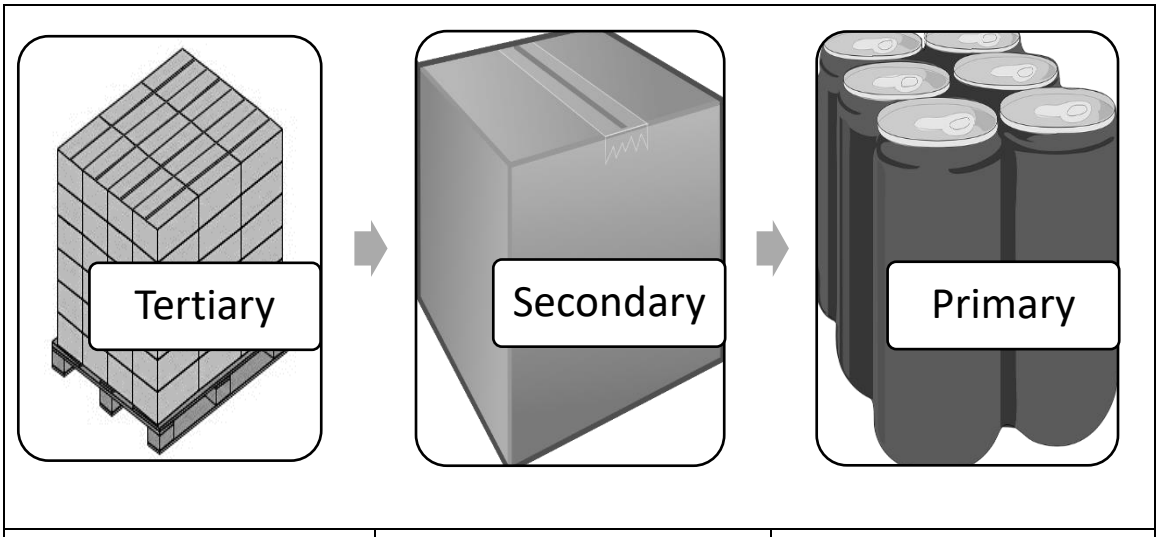
Although there are no standard definitions for slow- and fast-moving items, there are theoretical and practical reasons for making the distinction. A retailer buys inventory in bulk and has store personnel arrange it on a limited capacity of customer-accessible

store shelves. For slow-moving items the cost advantage of buying in bulk may be lost if the items stay on the shelf for too long (Whitin & Youngs, 1955). How much space set aside on the store shelf and how often to have instore personnel replenish the shelf may matter more for fast moving items as the empty shelf will likely get more visits than an empty shelf for a slow-moving item over the same length of time (Gruen & Corsten, 2008). Practically speaking, slow items may have insufficient sales or stockout occurrence for researchers to study. Retailers may also prefer to focus on FMCG because of their "revenue-generating ability" (Wintle & Patch, 2003) compared to slower items. Even for items with low profit margins FMCG have a higher volume of sales than the slow-moving items.

While there is empirical research studying differences in distribution decisions and performance outcomes between fast- and slow-moving goods, most of the time FMCG is simply the domain or context of other research. Sales increase at a higher rate for FMCG than slow-moving items, given the same increase in shelf space (Curhan, 1973), or promotional effort (Cooper, Baron, Levy, Swisher, & Gogos, 1999), while stockouts increase at a higher rate for FMCG than slow-moving items, given the same increase in inventory review periods (Sezen, 2006). In general, however, fast- and slow-moving goods are separate streams of research where they simply serve as context. For example, Grubor and Milicevic find that items with higher turnover (measured in average monthly sales) have more stockouts than items with lower turnover (2015). However, this analysis is done within the FMCG context, so it does not encompass slow-moving items at all. Rather, within quickly moving items, those that sell more on average are linked to more stockouts. A question that arises from this example is whether FMCG is

being used interchangeably with CPG. Overall, a taxonomy on distribution and sales management of FMCG would help to more clearly outline the arc of this retail research. The FMCG role in instore operations is still unclear.

1.2.3 Packaging logistics



Handled only with machinery to load and unload containers or trucks or to move from one spot to another within a bulk storage facility.	Manually handled by store personnel as individual boxes or with handcarts and dollies (trolleys) for multiple boxes.	Manually handled one consumer unit at a time onto and off of store shelves by customers and store personnel.

Figure 1-6 Product packaging types

There are 3 levels of product packaging, as in Figure 1-6. In this example, the primary packaging, made available on store shelves, is a six-pack of cans. Cans may be sold individually, or the same product (the same brand and type of drink) may be offered as different items or SKUs because of single, 6-, 8-, or 12-pack primary packaging options, not to mention bottled packaging of the same product instead of cans. Depending

on the primary packaging, therefore, a consumer unit could be a single can or multiple cans sold together as a single sellable unit. Several consumer units of at least one SKU (item) are placed into secondary packaging often referred to as a case pack. These case packs can then be loaded onto pallets, or tertiary packaging, to be handled more easily during loading and unloading activities in transport between firms. Each packaging type is "successively broken down into smaller units as it moves downstream" (Bartholdi & Hackman, 2016, p. 12) so that retail shelves consist generally of primary packaging. However, "shelf-ready packaging" (SRP) is used directly on store shelves even though it is secondary packaging. SRP (Dujak, Ferencic, & Franjkovic, 2014) is designed so that a portion of the secondary packaging or case pack is easily removed, leaving a tray of primary packaged consumer units with which to quickly replenish store shelves. Beyond the scope of this dissertation, this "hybrid" packaging type may be empirically studied to see if certain products benefit more than others in terms of increased efficiency in replenishing the shelves or even linked to increased sales.

Packaging literature identifies three functions of packaging: ethics, marketing and logistics, with no consensus on which of the latter two is its main function (Regattieri & Santarelli, 2013). The ethical dimensions of packaging include how easy it is to be recycled, reused, user-friendly, and socially responsible, produced efficiently, honest and complete in providing information, and reducing risk of human injury, (Vernuccio, Cozzolino, & Michelini, 2010). In terms of marketing, packaging serves a differentiation function, promoting product or brand awareness and informing the customer about product specifications. This marketing purpose generally refers to the product's primary packaging, and a recent review highlights the effects of packaging on consumer attitudes

towards a product (Lo, Tung, & Huang, 2017). Identifying the product is also a useful aspect in logistics functions. Packaging allows for tagging the product so that it can be tracked while in transit or in storage. Tagging technology, like radio frequency identification (RFID) is cost-prohibitive at the primary packaging level and is generally implemented at the secondary packaging (case pack) level (Condea, Thiesse, & Fleisch, 2012) for retailers, or with tertiary packaging higher in the supply chain where items still move in bulk.

Besides allowing for identification and tracking of products, packaging also serves protection and handling functions in logistics. Packaging functions as protection at every step of the supply chain (Bowersox & Closs, 1996) from tertiary packaging when products leave the manufacturer, to primary packaging which protects against theft as well as mechanical damage (Hellstrom & Saghir, 2006) at the store shelf. Handling functions consist of four areas: volume efficiency, weight efficiency, consumption adaptation, and handleability (Regattieri & Santarelli, 2013, p. 193). Handleability refers to how ergonomic the packaging design is in terms of ease of holding the packaging or opening it. Consumption adaptation considers how much packaging is stored and used, and how the same packaging can be adjusted (flexibility of use) for different levels of consumer needs. Weight efficiency is of special importance for any packaging that is manually handled; packaging must be as light as possible to minimally increase product weight. Volume efficiency involves the degree to which a package is filled with a product. Standard packaging that isn’t customized to the dimensions of a product may contain unfilled space (volume inefficiency). This inefficiency is a logistics concern for every level of packaging and also with respect to marketing functionality (Wever, 2011).

Store personnel deal with different levels of packaging depending on their store type. Three store types have been identified by scholars (van Zelst, et al, 2009, p. 624)

> 1- dense retail outlets: receive tertiary packaging (crates, pallets, etc.) composed of multiple items, where each shipping box has fewer consumer units of an item than its normal case pack size
> 2- stores which receive secondary packaging (totes, case packs, ship packs, etc.) and stack consumer units onto store shelves
> 3- discounters: receive and stack secondary packaging

van Zelst et al. (2009) claim that the second store type is the only one which has a "nonlinear shelf stacking cost" because handling activities are not a "constant and linear rate in the number of CUs." In other words, store type (1) gets crates packed at a DC specifically for a set of shelves at the store where the CUs will fit, while store type (3) gets case packs and puts them directly on store shelves. It is only store type (2), claim the scholars, where personnel regularly handle items at both bulk and consumer unit quantities. This difference in handling activities is a function of how the goods are packaged, arrive to the store, and are put on store shelves.

The different levels of packaging and their handling affects logistics processes although this interaction is often ignored in the literature (Hellstrom & Saghir, 2006) (Waller, Williams, Heintz Tangari, & Burton, 2010). Furthermore, since packaging functions involve different business units and packaging levels interact with various SC member operations (see Appendix I), a roadmap for integrated packaging design has been suggested (Azzi, Battini, Persona, & Sgarbossa, 2012). Indeed, there are recommendations for supply chain members to collaborate even at the research and development stage of the item's life cycle, and proposals that packaging design should drive product development (Olander-Rose & Nilsson, 2009). As a relatively new

discipline packaging logistics still needs to develop a framework for comparing packaging levels, performance effects of any designs and of design changes (Saghir & Jonson, 2001). Because packaging affects handleability and shelf adaptation, the same call for integrated packaging processes is made in shelf management and instore logistics literature.

1.2.4 Shelf management and instore logistics

Instore logistics is often grouped with general logistics research because the retailer's unique position in the SC is not recognized. Instore logistics differs from out-store logistics by concentrating on different activities. Instore logistics consists of handling items, arranging, ordering and processing items within the store, while out-store logistics is the movement of large volumes of items, storing and delivering them to the store (Samli, Pohlen, & Jacobs, 2005). Shelf management is a part of the instore logistics process. Although a subset of instore logistics, it is a separate stream of research because of the self-service nature of CPG retailers and the importance of the various forces customers are exposed to at the store shelf, as illustrated by the "shelf management model" by Borin et al. (1994, p. 364) in Figure 1-7.

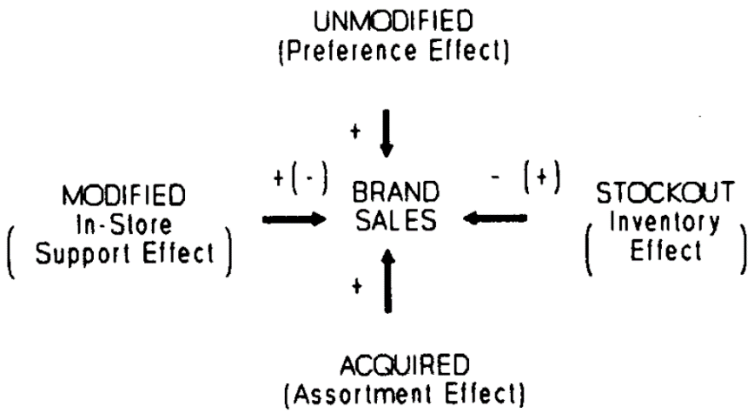


Figure 1-7 The shelf management model (Borin, Farris, & Freeland, 1994)

The dedicated space (retail shelf space for an item) is where several forces can come together, as in Figure 1-7, to affect the sales of the focal SKU (Borin, Farris, & Freeland, 1994). Borin et al.'s model shows that the SKU's sales ("brand sales") of this focal item always increase with the preference effect. The preference effect is when an unmodified demand, that is, a customer who initially prefers this SKU (primary customers), arrives at the shelf space for that item and purchases it. SKU sales can also increase when customers who originally prefer a different item that the store does not offer in its assortment (assortment effect) switch to this focal item as a substitute (secondary customers). Primary customers not willing to substitute may switch to another store so that there is no loss in sales of the focal item. Since store sales would decrease in this scenario, assortment planning literature studies how merchandise can best be managed in terms of which SKUs to offer. Assortment planning decisions, their antecedents and effects are beyond the scope of this dissertation. Shin et al. (2015) provide a thorough overview and classification of assortment planning research.

The remaining two effects in Figure 1-7, which may either increase or decrease the focal item's sales according to Borin et al. (1994), both involve product availability.

A temporary shelf stockout (the item is in the store's assortment but is not on the shelf) spurs the customer to switch from the stocked out item to the focal item, increasing its sales with the inventory effect. The inventory effect can also decrease sales as this additional stream of demand is not due to the assortment effect, which can be somewhat planned for as the initially preferred product is never in stock. This unexpected additional demand can cause the item to stockout earlier than planned, thereby creating dissatisfied primary customers. Customers may be dissatisfied enough to switch away from the focal item because of the instore support effect as well. Instore support activities can be merchandise or store management tasks. Beyond the scope of this dissertation, though controlled for in models, merchandise management activities include promoting the product (Taylor & Fawcett, 2001), reducing its price (Shin, Park, Lee, & W.C., 2015), or decisions about how to display the product to attract customers (Pizzi & Scarpi, 2016). The amount of space allotted to an item on the shelf (its shelf capacity), as determined by the number of facings it has on a shelf "are a result of marketing planning and therefore given exogenously" (Sternbeck, 2015). Shelf space elasticity (Eisand, 2014) stems from this stream of instore support activities and is also beyond the scope of this dissertation.

1.2.4.1 Instore logistics processes

The overarching goal of instore logistics activities is to manage the shelf in such a way that customers arriving at the shelf are not faced with a stockout of their preferred product. The stockout of the product in Figure 1-7 results in a negative (-) effect on brand sales. The stockout of a different product results in a positive (+) effect on brand sales of this product. This positive or negative effect is the inventory effect and occurs only with a

shelf stockout (SHO). Stockout research covers what drives the stockout phenomenon and what effects it has on customer or firm outcomes. While stockout occurrence is a separate stream of research, the phenomenon of shelf stockouts may be an undesirable result of inefficient or ineffective shelf management processes.

Shelf management and instore logistics research focuses on the processes themselves. Antecedents to these processes, like the number of units the personnel needs to shelve and the number of units to be unpacked from case packs (Curseu, Woensel, Fransoo, Donselaar, & Broekmeulen, 2009) are tested to see how these processes vary by these factors. The processes are studied in terms of level of efficiency using constrained resources such as time (van Zelst, van Donselaar, van Woensel, Broekmeulen, & Fransoo, 2009), space and labor (Reiner, Teller, & Kotzab, 2013). For the effects of instore logistics processes, these inputs of time, space, and labor are studied to see what sort of outcomes the retailer faces in terms of item turnover, service level, lost items (Reiner, Teller, & Kotzab, 2013), customer loyalty to the store (Bouzaabia, van Riel, & Semeijn, 2013) and even inventory record inaccuracy and misplaced items (Raman, DeHoratius, & Ton, 2001).

Research has begun to define instore logistics and shelf management processes as a series of tasks. In their empirical analysis Curseu et al., (2009) model how items are shelved onto store shelves by dividing the process into 7 tasks:

(1) grabbing and opening case packs[4],

[4] A case pack is a manufacturer-owned shipping box. Case packs come in different case pack sizes (CPS) which refers to the number of consumer units of the item in the shipping box from the manufacturer. If the shipment to the store is from a retail distribution center (DC), OPQ or order packing quantity is used in the literature instead of CPS. Generally speaking, OPQ has fewer consumer units of an item than CPS because

(2) searching for where the product is located on the shelves,

(3) walking to that location,

(4) preparing the shelf for the new items by removing any existing items,

(5) putting the new inventory on the shelves (first replenishment),

(6) putting the old inventory back on the shelves (second replenishment), and

(7) disposing of packaging waste.

They use hierarchical regression because these activities take place in sequential order. They find that case pack size and the order line size (number of ship packs per order) affect different tasks. Case pack size, for example, is directly related to task (1) and (6), and insignificant elsewhere while the order line size affects the time for task (5) the most. They also find that there is a minimum amount of time to setup which is a constant. Because of this, they recommend retailers either have larger order line size or have case pack size increased to save product handling time. Within product handling tasks, inventory overflow--where a portion of the products in the case pack does not fit the designated space capacity (Eroglu, Williams, & Waller, 2013)--is not considered by Cursue et al. (2009). In practice though, items that do not fit onto dedicated space are sent to a *shared space*, which may be the backroom of the store or at the very top of store shelves where customers are unable to reach the products without assistance.

Another way that the instore logistics processes have been defined is into core elements (5S): stock, systems, standards, space and staff (Pal & Byrom, 2003). "Stock" refers to inventory investment, or the actual number of units, holding cost, or monetary

---

case packs are generally opened at retail distribution centers to be repacked with other items to various stores. The general term for the box size used in shipping to a store is SPS, or ship pack size, and can refer to the number of consumer units per box from either the DC or manufacturer. Shipping boxes, whether from a vendor or retail DC are all secondary packaging.

value of the goods held in stock, and is an element the store wishes to minimize. Systems, standards and management of space and staff are all related to operational efficiency. Systems refer to the technology and decision-making framework for store operations, and standards are how well staff are trained and adhere to systems set in place. Last of all, space and staff both serve as the "input" values in productivity or operational efficiency measures. All of these are in the tactical area of store management activities.

Though these 5S core elements have yet to be empirically studied in this framework, their conceptual development appear to have grown out of prior research. For example, van Donselaar et al. state that the "space in the retail store should be considered as a constraint rather than as a cost factor," (2005, p. 14) because handling costs (staff) are greater than (stock) holding costs (Curseu, Woensel, Fransoo, Donselaar, & Broekmeulen, 2009). Holding costs generally come into play when considering the amount of space to allot to a product and balancing the costs of holding the product with the customer service level to be achieved. Since space for items affects demand for those items (Eisand, 2014) and space for overflow inventory in terms of backroom space affect shelf availability (Milicevic & Grubor, 2015), it makes sense to group one of the logistical activities as space. Similarly the 5S (Pal & Byrom, 2003) overall parallel the category drivers developed for stockout research (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016). This grouping could be helpful in ensuring that shelf management models account for all 5 core elements of operations, and even to see which of them affect performance more than others.

However, both higher-level groupings such as this and explicit task breakdown studies often do not consider testing tasks related to substitution between products or storage of items in both dedicated and shared space. Indeed, even retail category management software used by retailers on average disregard (Hubner & Kuhn, 2012) limited shelf space, substitution, or any demand linked to the amount of inventory visible, known as the inventory billboard effect (Wu, Zhai, & Liu, 2015). A closer look at the shelf management process followed by a thematic look at the factors in play may provide a fuller picture of instore logistical activities.

1.2.4.2 A closer look at the shelf replenishment process

This section presents 4 graphs of an item in a store over a period of 8 days. The dedicated space inventory shows how many units of the item is on the store shelf space that is assigned (dedicated) to that item. It is a measure of customer service because stores are self-service environments and customers will serve themselves to items on the shelf only if those items are actually in their dedicated space. The shared space inventory shows the total units in the store that are not in dedicated space. While the shared space may be any location within the store, in this work only the backroom is considered and "shared space" is used interchangeably with "backroom." The inventory investment graph shows the total units on-hand (in the store) regardless of where they are located. Last of all, the operational efficiency graph shows where the shelf replenishment process is inefficient. The section ends with a short discussion on the research issues the figures illustrate and what they imply for multiple items in a store.

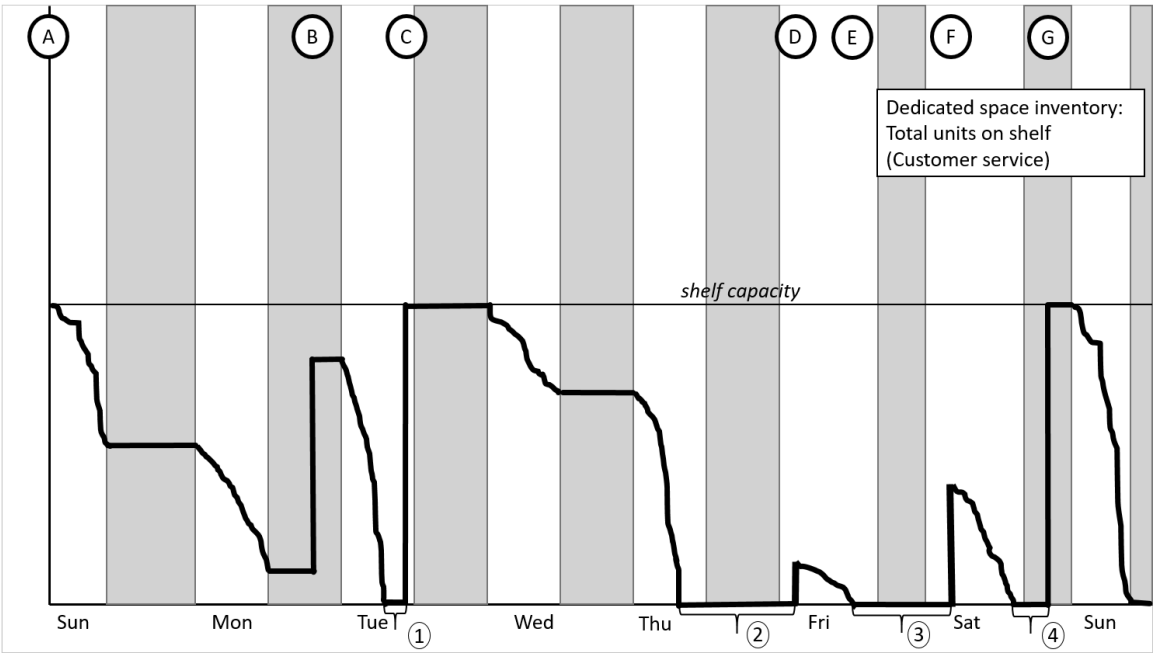


Figure 1-8 Graph of total units of an item on a store shelf

The illustration in Figure 1-8 is a typical sawtooth model of shelf inventory levels (units, vertical axis) in a store over a time span of 8 days (horizontal axis). The line shows the number of units within the dedicated space (shelf capacity) assigned to the item. Store hours vary on weekend nights and Sundays, so the hours open (not shaded) and closed (shaded) vary in width (length of selling time) depending on daily operating schedules. There is an increase of the inventory level whenever personnel place units onto the shelf (points A, B, C, D, F, and G) whereas the general trend of the inventory level is decreasing. Shelf inventory decreases whenever customers remove units of the item from the shelf. The inventory level is flat (neither increases nor decreases) whenever the store is closed, is open but has no customer demand, or the shelf has stocked out (SHO) and is awaiting replenishment. Zero-level inventory intervals are marked as intervals 1-4 in Figure 1-8. During these time intervals, there is poor customer service because customers who arrive to the dedicated shelf space do not find the item since the shelf has not been replenished.

Shelf stockouts (SHO) and replenishments can each occur under two different conditions, and while they appear as a flat zero or an increase (spike) in the customer service graph, one condition for each situation is indistinguishable from another in Figure 1-8. In Figure 1-8, replenishment of the dedicated shelf space occurs at points A, B, C, D, F, and G. However, deliveries of the product to the store take place at points A, C, E, and G. Therefore, although indistinguishable from the other points, the increase in units on the shelf at points B, D, and F occur from instore replenishment. Instore replenishment is when store personnel move consumer units of an item from shared space inventory (backroom) to fill empty dedicated (shelf) space. If there is inventory in the backroom and none on the shelf, then the SHO is an "instore replenishment stockout" (IRO), whereas if both spaces are stocked out the SHO is a store stockout (STO). In Figure 1-8, interval 1 is a STO while interval 2 is an IRO, but there is no distinct difference in the customer service line. Without the outside knowledge of when deliveries take place or whether there is any shared space inventory, zero points or increases in inventory levels in Figure 1-8 all appear to be indistinguishable stockout (SHO) and replenishment events. Additionally, outside knowledge of a delivery (point E) is not reflected in Figure 1-8 since those delivered units remain in shared space.

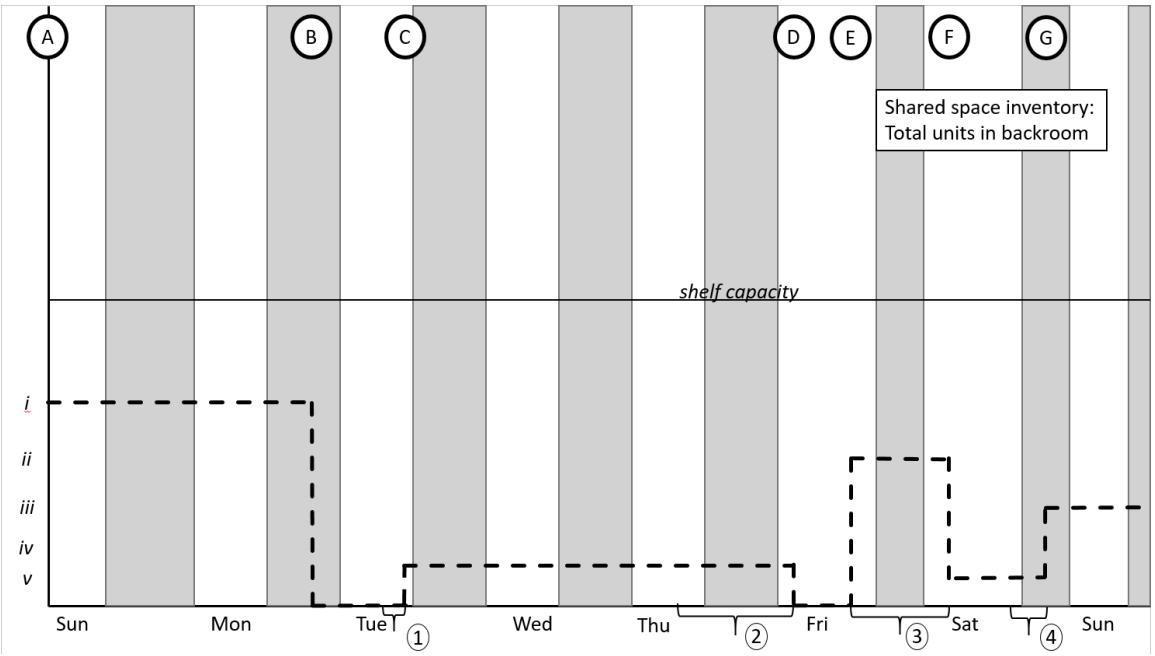


Figure 1-9 Shared space (backroom) inventory level over time

Units of the item located in shared space is illustrated in Figure 1-9. The dashed line shows the number of units of the item that is in the store but not on the dedicated shelf space. These units may most often be stored in the backroom, but may also be at alternate locations than the dedicated space which is known to customers. Customers may go to the dedicated shelf space for the item and not find it, not knowing to look at a different space the item shares with other products, or not having access to that shared space. At interval 1, where the customer service line (Figure 1-8) is at zero units, the shared space (Figure 1-9) also has no inventory. However, at interval 2 the customer service line is also zero but the shared space inventory is positive (iv units). During this time, all of the units of this item that are in the store are in the shared space and are not on the dedicated shelf space. Unless customers have knowledge as well as access to the shared space, the inventory there is not available for purchase. The (ii) units during interval 3 and (v) units during interval 4 are also not in the item's dedicated space but elsewhere in the store.

The shared space inventory graph does not look like the dedicated space inventory graph. Similar to Figure 1-8, the inventory levels in the shared space peak or suddenly increase after deliveries to the store (points A, C, and G). However, since no units of the store delivery at point E are moved to the dedicated space, there is no increase in Figure 1-8 while Figure 1-9 captures this delivery. Additionally, the shared space inventory level does not directly decrease due to a customer purchase since customers do not have access to inventory in this space. Instead the inventory level drops to a lower level whenever *personnel* remove the item from the shared space to replenish the dedicated space. Personnel can remove everything from the shared space, as at points B and D, or only a portion of the shared space inventory as at point F. Partial removal may be from replenishing a partially empty dedicated shelf space or from inefficiently managing shared space inventory. Similarly, the amount of increase in shared space inventory can vary. At point C there are fewer units entering shared space than in point A (iv < i) either because there were no items left on the shelf at delivery time or because the delivery size was smaller than at point A. While the dedicated (shelf) space is replenished in bulk quantities and decreases in unit quantities, the shared (backroom) space is both replenished and emptied in bulk quantities. In this way the shared space inventory line appears to be more like a piecewise or step function rather than a sawtooth model.

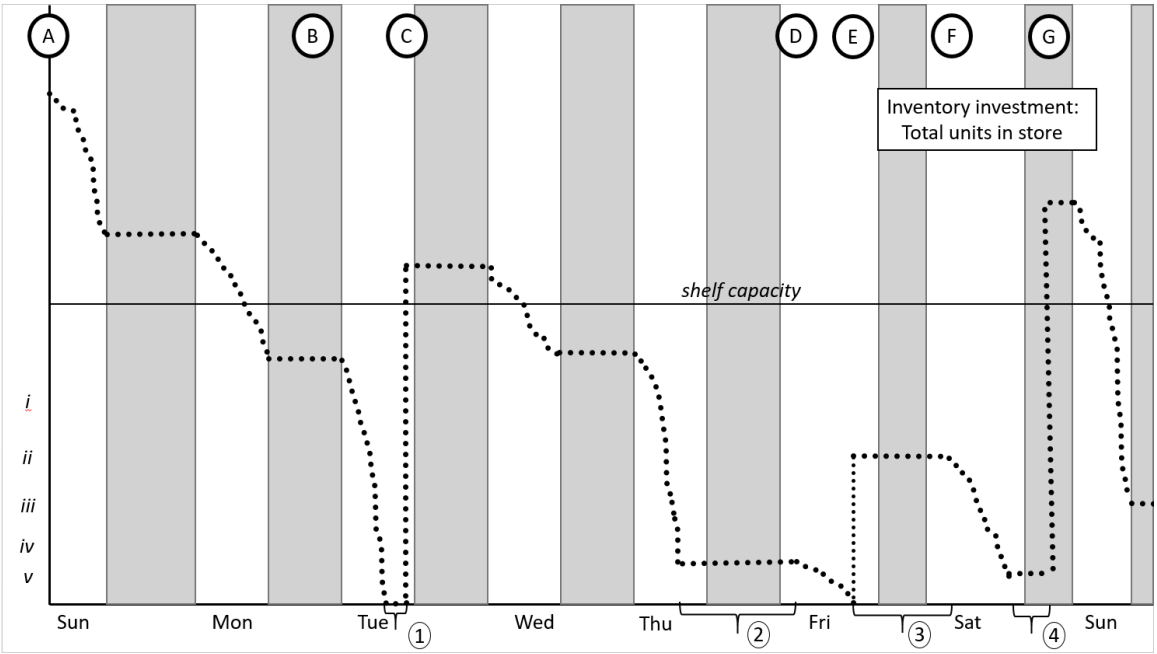


Figure 1-10 Inventory investment - Units on hand over store week

The illustration in Figure 1-10 is a graph of on-hand inventory levels, which is a combination of the shared and dedicated space inventories, and is typically the inventory records retailers keep for their stores. Deliveries to the store (points A, C, E, G) account for increases in inventory levels while customer purchases decrease on-hand inventory levels. In Figure 1-10, whenever the on-hand inventory level is above the horizontal shelf capacity line, the store has more units of the item than can fit on the dedicated shelf space. Such "excess" units, at point C for example, are stored in shared space at a minimum, up to the entirety of what is in the store at that time. On the other hand, when the inventory investment is at a number of units at or below the dedicated shelf capacity, it does not necessarily follow that all of the inventory in the store is actually in its dedicated shelf space. Since inventory investment is defined as the number of units the store physically keeps at a store, it is used interchangeably with on-hand inventory and ignores any purchases that have yet to be delivered to the store.

There are 3 different situations where the inventory investment line is flat (horizontal) over time instead of generally decreasing or with a sudden store delivery peak. Obviously, in all 3 situations neither supply nor demand is driving the line up or down, but the reason for the lack of demand varies in each situation. First of all, the inventory investment line is flat whenever the store is closed since customers do not have access to store shelves during these periods (shaded blocks) of time and are unable to buy the item. Secondly, the inventory investment line is flat and at 0 units, as at interval 1 in the store shelf inventory graph (Figure 1-8), whenever the item is stocked out of the store (STO), so that even if the store is open customer demand cannot further lower inventory investment levels (no backorders). Last of all, the inventory investment line is flat but not at zero inventory and not during closed hours, at intervals 2, 3, and 4 in Figure 1-8, when there are no customer purchases.

There are two possible reasons for no customer purchases during these intervals: (1) no customers are coming to the store demanding the item throughout this time, (2) customers are coming to the store demanding the item but the item is not available on store shelves. In the latter case (2) it is clear that the SHO situation is an IRO and not a STO, because the on-hand inventory level is nonzero as at intervals 2, 3, and 4 in Figure 1-10. Periodic monitoring of store shelves by personnel can confirm an IRO. However, it is more difficult to continuously monitor store shelves for customer arrival to confirm the former reason (1) of no demand. While the inventory investment line is an example of the information that stores generally keep of on-hand inventory, there is no knowledge of what proportion of that on-hand inventory is available to customers on store shelves.

While STO conditions are clearly visible in the inventory investment graph, IRO conditions are masked by the lack of knowledge of what is in shared and dedicated spaces. The interval marked as 1 in Figure 1-10 is a STO, while the intervals marked as 2, 3, and 4 are IRO. Though they are all SHO situations, the inventory investment line shows no stockout situation during IRO intervals. At interval 2, there is some (iv units) inventory overflow from a store delivery at point C. The overflow was placed into the backroom and remained there until point D. At interval 3, a store delivery of inventory investment units takes place at point E but the entire delivery remains in the backroom. Finally, at interval 4, because only part of the shared space inventory (ii-v units) at point F is moved to dedicated space, there remains v units in shared space that is inaccessible to customers.

Similarly, instore replenishment activities are also masked by the lack of knowledge of what is in shared and dedicated spaces. In this example, instore replenishment occurs at points B, D, and F. Since the shared space inventory from point A to B is so high, the inventory investment line does not reflect the fact that the shelf started out half empty on Monday and was almost SHO by the end of that store day. An instore replenishment occurring at point B shifts the shared space inventory to dedicated space, but is indistinguishable in the inventory investment line. Indeed, there is no indication that the iv units in interval 2 or ii units in interval 3 were moved to the shelf at points D and F. In this way, the records that stores keep of their on-hand inventory poorly reflects product availability at the shelf, simply because they are a combination of both dedicated and shared space inventories.

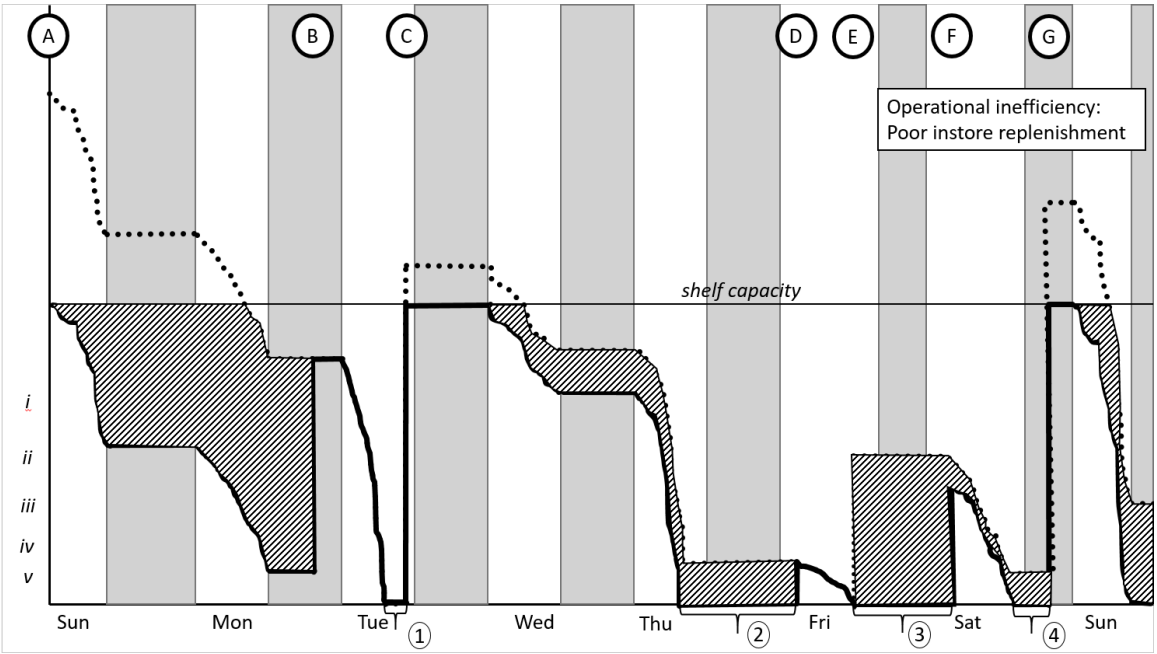


Figure 1-11 Operational efficiency graph

Figure 1-11 maps the areas of operational efficiency in shelf replenishment. Shelf replenishment is inefficient between these two lines of total on-hand (dotted line) and dedicated space (solid line) inventories and within the shelf capacity the store has dedicated to this item (below the thin horizontal line). The area between E and F, for example, has no shelf inventory; it is all in the backroom. The shelf need not be empty, but it is. The inefficiency continues as only part of the inventory in shared space is moved to dedicated space, so that the operational inefficiency (in replenishing the shelf) continues until point G. In terms of in-store logistics, however, this area may be necessary for the store to efficiently use store personnel. If personnel during this period need to focus on checkout operations, then the store is efficiently managing its labor while inefficiently managing its dedicated space. In such a situation, shared space is being used to temporarily hold items while personnel are completing checkout tasks.

A few characteristics about these areas of inefficiency emerge from the figure. First of all, the start and end time boundaries coincide with store deliveries and instore replenishments, respectively. At point C a store delivery is made and the shelf is filled to capacity with some inventory overflow remaining. Starting at that point in time, as soon as customers have begun taking consumer units off of the shelf, the inefficiency begins by there being empty space on the shelf while there are units to fill that empty space in the backroom. This inefficiency continues until point D, where an instore replenishment is made. Secondly, the height of the inefficient regions are also within a range of values

in terms of units. The difference can never be greater than the number of units in the backroom, but is also constrained by the shelf capacity. This is illustrated most clearly between points A and B. Throughout the day on Sunday there is a region of inefficiency from empty space on the shelf that could be filled by backroom inventory, but the level of inefficiency doesn't reach the level of inventory in the backroom until midday Monday. In other words, any inventory in shared space beyond what fills the shelf capacity is not considered inefficient in terms of instore replenishment.

In this way, Figure 1-11, clearly depicts only one type of operational efficiency, related to instore replenishment, and raises questions about inventory investment and customer service. The earlier discussion of 5S core elements (Pal & Byrom, 2003) of instore logistics processes concluded that operational efficiency can be in terms of systems, standards, management of space, of stock and of staff. The inefficient areas are poor management of dedicated (shelf) space. If retailers choose to dedicate a certain capacity of shelf space to an item, they do so with a managerially chosen target fill rate (TSL). That TSL can determine not only the shelf capacity but how often to replenish that dedicated space. Smaller amounts of space would need more frequent replenishment than larger dedicated space, to maintain the same TSL. To efficiently manage TSL (a standard), stores may be pushed to increase inventory investment (stock) beyond dedicated space (shelf space) and have personnel (staff) perform instore replenishment. In other words, efficiently managing an operational standard requires minimizing operational inefficiency in terms of space and increased inefficiencies in stock and staff management.

Overall, figures 1-8 through 1-11 raise several research-related issues for shelf management. Store deliveries in this example occurred 3 times (at points A, C, and E) while shelf replenishment occurred 6 times (A, B, C, D, F, G) and stockout events occurred 4 times (Tuesday, Thursday, Friday, Saturday). These 3 events (store delivery, shelf replenishment, shelf stockout) are distinct though related events. Since personnel can replenish from shared space inventory, there are more frequent shelf replenishments than there are store deliveries. Studies on how prevalent it is to have a store delivery not fully fit onto designated shelf space would be useful in understanding how much backroom replenishment has an effect on task breakdowns. Although Gruen and Corsten (2008) do state that 25% of stockouts in their study had backroom inventory available, such as that illustrated from point E to F in Figure 1-11, it is not clear how often fully moving units from shared to dedicated space (as in points B, D, and G) would occur relative to partially moving shared space inventory (as in point F).

The task breakdown for shelf management research needs to better include instore replenishment activities. Taking into consideration the various ways SHO and shelf replenishments can occur extends Cursue et al.'s (2009) list of 7 instore logistics activities by at least 4 more steps. First of all, an 8th step of putting inventory into shared space would reflect both planned and unplanned actual use of this space. Instore replenishment activities would add 3 more steps: (9) noticing there is some amount of empty dedicated space, (10) searching and finding items in shared space, (11) bringing items from shared to dedicated space. There is no way to tell whether shelf inventory graphs spike with a replenishment from step (5) or step (11) or what type of SHO (STO or IRO) preceded these replenishments. Not only does this example illustrate current

tasks need to be extended from 7 to 11 different activities, it also shows that the process is extended by an additional 3 steps (steps (4), (5), and (6)) for actual restocking when the shelf was partially full during instore replenishment. Thus the number of tasks possible for one shipment delivered to the store, until all of its consumer units are sold, is doubled. The many steps involved may better explain why replenishing shelves has been found to take more than half of personnel work time (Kotzab & El-Jafli, 2015) and cost 38% of logistical costs as compared to in-store inventory's 7% cost on average (van Zelst, van Donselaar, van Woensel, Broekmeulen, & Fransoo, 2009).

Another research-related issue is modelling such a complicated system since there are different parties involved for different extended periods of time. For store personnel, it may be ideal to replenish shelves when the store is closed, so that during open hours they can attend to customers. For the retailer's instore logistics costs, it does not matter whether the store is closed or not because staff are still carrying out operations at some cost to the retailer. However, while the store is open it is possible that replenishing the shelves come at a higher cost to the retailer because the process may take longer as staff attend to or wait for customers within the same aisle. For customers, as long as items are available (on the shelf) when customers arrive to the shelf, it does not matter when the items were replenished prior to customer arrival. However, there may be some cost in terms of decreased service quality when staff are replenishing during store hours and are

in the way of customers, especially when orders to the store are received and larger boxes take up aisle space from customers.

While demand (customers) can only arrive to the store during the store's open hours, the time period for activities related to supply depend entirely on the retailer and can include closed store hours. If logistical activities were modelled with time series considerations, a unit of time could be a 24-hour day. Looking at the effects of the replenishment activities on stockout duration for example, would include time spent replenishing the shelves during closed store hours. However if the duration for the stockouts were counted during closed hours, it would inflate its mean values as stockouts in places the customer cannot access (store shelves during closed hours) are not valid stockouts. Similarly, for models accounting for demand autocorrelation, the time intervals involved are not evenly spaced. Transformation into evenly spaced time intervals (Nieto-Barajas & Sinha, 2015) or some sort of uneven time series model or advanced modelling would be required.

Although the inefficiencies that may arise out of unintentional shared space use are being explored, the implications of the purposeful use of shared space has received little attention. The "backroom effect" (Eroglu, Williams, & Waller, 2013) has been introduced for store deliveries where some units of an item do not fit on dedicated shelf space. This inventory overflow is said to "induce more shelf [stockouts]" (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016) because the shared space inventory is poorly managed and replenishment from shared space does not properly take place. A stark example of poor shared space management in the illustration is at point F, where only

part of the shared space inventory is used for replenishment. The stockout that occurs after point F may have been prevented if all shared space inventory had been used. Inventory overflow studies assume that shared space units are not used for shelf replenishment because they are simply overflow, or units of the item that unintentionally did not fit on the shelf. Studies on how often and under what conditions stores *purposefully* have an order size exceeding dedicated shelf capacity would allow better understanding of how well shared space is being managed by personnel without explicitly keeping shared space inventory records.

The "opposing goals" (Tabucanon & Farahani, 1985) of customer service, inventory minimization, and operational efficiency can be better visualized through Figures 8-11. The customer service line of on-shelf inventory levels can have shorter intervals of zero units if the inventory investment exceeds the shelf capacity at the time of store delivery, allowing for instore replenishment. This almost instantaneous replenishment using store inventory may have improved product availability as compared to scheduled or fixed deliveries from DCs, which take transportation time. However, store inventory investment should ideally be minimized instead of increased, not only because the store can avoid holding costs but because a centralized inventory at a DC can better distribute the item to multiple store locations as needed. Lowering inventory investment would also both increase and decrease OE. Space and labor would be more efficiently managed since there would be less shared space use and no instore replenishment, while stock would be less efficiently managed since SHO would be more likely to occur and for longer intervals.

Finally, Figures 8-11 have further implications beyond their illustration of a single item at a store. Stores have multiple items within a product category (assortment depth) as well as multiple product categories (assortment breadth, or product variety). This example with illustrations of on-hand, dedicated and shared space inventories, as well as operational inefficiencies considers only a single item. Increases (both store and instore replenishments) and decreases (customer purchasing and personnel relocating items) in the graphs appear as if neither the customers nor store personnel interact with any other item in the store. In reality, a customer may decide whether to buy this item depending on the shelf availability of other items in the product category or store. Or if this item is SHO then they may buy another item as a substitute. Similarly, personnel may replenish this item before a SHO occurs while replenishing another item which has SHO. Or they may not be able to carry out an instore replenishment of this focal item but may have brought out other items from the backroom in this product category which they then replenish. Last of all, personnel may not be able to replenish the dedicated space of this item simply because they are attending to the replenishment of other items. In this way the shared and dedicated inventory levels of a single item can be correlated with the shared and dedicated inventory levels of one or more other items in the store.

1.2.4.3 Overview of current theory

Store handling operations have not been studied much (van Zelst, van Donselaar, van Woensel, Broekmeulen, & Fransoo, 2009). In fact, “compared to classical logistics functions of transportation and warehousing, instore logistics plays a minor role in publications about retail logistics” (Kuhn, 2013, p. 9). In certain studies, in-store

logistical activities are not the focus of the hypothesis testing (Mckinnon, Mendes, & Nababteh, 2007) (Waller, Williams, Heintz Tangari, & Burton, 2010), since there are no measures for the logistical activities of store personnel or of other resources directly measuring instore logistical activity.

Instead, scholars use the characteristics of the store or of items, as a measure of logistical activities. For example, in testing the effects of case pack size on item sales relative to other items, Waller et al. (2010) assume that the size of the case pack can be used to measure the store's reliance on overflow areas, using the theory that larger case pack sizes raise the likelihood of overflow inventory, which is then brought back to shelves less efficiently than shelves which are replenished with items newly delivered to the store (Gruen & Corsten, 2008). Curseu et al. (2009) do find a positive relationship between case pack size and the time it takes to complete certain logistical activities, but do not distinguish between replenishment from deliveries to the store and those from within the store.

Backroom space is one of the store characteristics that scholars suggest have an impact on the SC (Pires, Pratas, Liz, & Amorim, 2017). First, backroom areas could help DCs with capacity issues, by being able to push goods downstream instead of keeping them at the DC location until store customers demand them. Secondly, the backroom could enable the DC to plan routing to stores over a greater time period, as the sales area does not need to be open or be affected by the traffic from a delivery to receive the order. Transportation costs may be reduced, say the scholars, because multiple trips to the store throughout the day are avoided as consumer units that do not immediately fit on the

dedicated shelf space can remain in the backroom until needed later in the day. The fourth way backrooms may impact the SC is by enabling online purchases to be picked up or shipped out from store locations, so the backroom is an extension and expansion of the distribution network. The last two ways the scholars suggest backrooms can impact the SC is related to the design of the backroom. It is not only an area for product storage but one that holds employees (#5) who can sort delivery paperwork and enter inventory information into store records (#6).

While warehouse planning has long been considered a strategic component of SC networks (Ballou, 1976), store backrooms have not been included in such studies and they do differ from warehouses. The Pires et al. (2017) discussion on "backroom particularities," are summarized (Table 1-3) here according to the 5S core elements of retail (Pal & Byrom, 2003). Pires et al. (2017) emphasize that backrooms are designed in an ad hoc manner, often by building new stores similar to existing store formats. The backroom is generally split into a social area for personnel and a product area, with products split into food, non-food, and chilled areas. The total backroom area competes with the sales area, as there is an overall store size that both types of space must share.

Area	How a backroom space differs from a warehouse (W)
Systems	• Backroom design depends on the type of retail store • Does not use aisles which take up too much space • Has fewer and less powerful equipment
Space	• Square footage more expensive since stores are located in residential or city/downtown spaces

	• Must “coexist with the selling area, which competes for the same space” (p. 236)
Stock	• Demand for backroom stock is less predictable than warehouse demand and is only triggered when there is a SHO • Deliveries to store are more frequent and smaller compared to W
Staff	• Has fewer personnel • No dedicated pickers; personnel must split time between backroom and sales area
Standards	• Not as organized as warehouse space • Not considered a priority within the store since personnel focus on attending to customers

Table 1-3 Warehouses versus backrooms (Pires, Pratas, Liz, & Amorim, 2017) in 5S areas

Furthermore, according to Pires et al. (2017), this space is more expensive per square foot than the space in a DC or warehouse. Backroom space is located either in a residential area or other non-industrial and high-customer traffic area. This more expensive and constrained space has systems in place that differ from typical warehouse space. It has “a low level of automation” compared to warehouses since store personnel “rely on manual operations” (Pires, Pratas, Liz, & Amorim, 2017, p. 236). There are also fewer number of store employees who work solely in the backroom space as compared to the number of employees who work in warehouses This is because the priority in stores is the sales area and serving clients, whereas in warehouses the main priority is managing the storage space. Since the CUs delivered to the store for each item are generally fewer than quantities arriving to warehouses, backroom space is less of a priority than warehouse space in terms of organizing any CUs of the item that remain in the backroom after shelf replenishment.

Shelf management and instore logistics decisions can affect store profit. Hubner and Schaal (2017) show how a specific retailer can increase profits by 22% for two shampoo items by using both shared and dedicated space. The mixed-integer programming model considers two types of replenishment costs: total direct replenishment (comprised of fixed cost to replenish an item, variable cost for each unit replenished, and dedicated space holding costs per unit) and total backroom replenishment (number of shelf refills from backroom, backroom inventory and backroom inventory holding costs per unit). The objective is maximizing profit by optimal choices of many facings each item has, the orientation of each item, and how often each item is ordered. Since they are given elasticity, cost and demand data for both items, they can calculate overall profit with different scenarios of shared and dedicated space use. From their model and case study they conclude items with higher profit margins should be allotted more dedicated space than items with low profit margins, which could be kept in the backroom and replenished more frequently. They also recommend: "if retailers have the opportunity to use backrooms for intermediate storage, they should leverage them, because backroom space allows for more flexibility in planning showroom shelf-space and in-store replenishment processes" (Hubner & Schaal, 2017, p. 145).

Overall, shelf management and instore logistics are considered tactical tasks that only affect daily operational efficiency of the store. While the below statement refers to how instore handling is linked to operational costs, further research on sales may prompt

strategic decision makers to also incorporate instore logistics into their decision making process.

> "A main deficit in retail practice and research is therefore the fact that instore handling operations are not well embedded in planning procedures on higher planning levels, even though instore handling operations contribute to nearly half of the total operational costs of grocery retailers." (Kuhn, 2013, p. 15)

1.3 Product availability

Product availability, or its counterpart of stockout occurrence, has great impact on firm performance. It is one of the "fundamental dimensions" (Bowersox & Closs, 1996) of customer service in terms of determining sales. Globally, retailers incur a revenue loss of $634.1 billion annually (Buzek, 2015) from stockouts which occur 4.9% to 12.3% of the time (Gruen, Corsten, & Bharadwaj, 2002, p. 13), depending on the product category. A store loses an estimated $200 to $800 per week on "lost time cost" of store personnel tracing stockout events and customers lose 20% of their shopping trip time waiting on more information about a stocked out item (Gruen & Corsten, 2008, pp. 8-9). Approximately 50% (Zinn & Liu, 2008) to 81% (Corsten & Gruen, 2003) of customers who face a stockout switch stores, and repeated stockout experiences affect customer store and brand loyalty (Turk, 2012) (Walter & Grabner, 1975) with retailers losing short- (Wu, et al. 2013) and even greater long-term (Campo, Gijsbrechts, & Nisol, 2003) market share.

1.3.1 Measures

Though important to firm performance, the retailer's unique way of adding value in the supply chain makes it difficult to capture stockout occurrence. As customers browse and pick from the store's accessible inventory, store personnel are not aware of whether the customer has taken the last unit of a product from the shelf to purchase or even to misplace in another part of the store. Indeed, store personnel cannot know exactly when a shelf stocks out unless they either have emptied it themselves during replenishment activity, are constantly monitoring the shelves, have a technology that

monitors shelf stock levels for them, or have a customer inquiry. Additionally, stores usually have some amount of inventory overflow (Eroglu, Williams, & Waller, 2013), where the total number of units of a product in the store's possession does not fit on the shelf space allotted for it. Stores could also be purposefully ordering more consumer units of an item than the shelf space dedicated for it, to be able to provide greater assortment depth with constrained total shelf space. Not knowing when that dedicated space has stocked out and needs to be replenished, or not being aware there is inventory overflow in the store for an item may result in poor instore replenishment of that item. Because of this, it is possible for a product to be stocked out on the store shelf but still be in the store out of the reach of customers and thus unavailable. The distinction between products that are in the store but have not properly been replenished on the shelf (IRO) and products that are neither on the shelf nor in the store (STO) is not made in this dissertation. In general, if a product is not available on the shelf space dedicated for it, it is a shelf-out (SHO) and this dissertation studies SHO without limiting the phenomenon to only STO or only IRO cases.

Since it is so difficult to measure an entire SHO event for a product on the store shelf, there have been multiple approaches to estimating SHO occurrence. Until now, only vending machine studies (Anupindi, Dada, & Gupta, 1998) have directly tracked a product's SHO occurrence in entirety. Stockouts usually appear in the literature measured in one of the methods listed in Table 1-4. Each method has its own strength and limitations and the use of one over the other may depend on the research question of the study. Experiments generally present a SHO event to study its effects in terms of customer

Table 1-4 Current methods of measuring stockout occurrence

Method	Example	Method strength	Method Limitation
Experiment	(Zinn & Liu, 2008)	The exact occurrence (duration, frequency, etc.) is known since it is an experimental factor	More appropriate for testing customer reaction to stockout rather than drivers of stockout occurrence
Surveys	(Dubelaar, Chow, & Larson, 2001)	Respondents familiar with store and may make strategic inventory decisions. Data for variable(s) of interest can be easily obtained	Captures general trends instead of actual stockout events. Not all customers report or ask store personnel about every stockout occurrence.
Simulation	(Wu, Huang, Blackhurst, Zhang, & Wang, 2013)	Accurate tally of stockout occurrence, ability to track stockout duration and frequency	Combinations of experimental factors used may not actually exist in practice
Perpetual inventory aggregation	(Avlijaš, Simićević, Avlijaš, & Prodanović, 2015)	Can gather information on more stores and SKUs, for a longer time than first two methods	Inaccuracies when physical units in store differ from recorded levels.
Manual audit	(Chuang, Oliva, & Liu, 2015)	Accurately flag a product as stocked out	No information of stockout duration; Additional costs incurred if obtained through third party audit
Point-of-sale data (POS)	(Grubor & Milicevic, 2015)	Exact day and time of purchase along with other items (basket shopping) made by a single customer, or aggregated by day	Items often incorrectly scanned at register. Affected by non-demand factors. Can only track individual item movement with RFID. Cannot use to study stockout effects on sales.

reactions or firm performance. With this method, it is not possible to also capture antecedents to SHO as the stockout occurrence itself is a researcher-controlled experimental factor. Similarly, point-of-sale (POS) data can be used to estimate when a product must have stocked out on the shelf. POS-based estimates of when a product must have stocked out varies: from straightforward calculations of assuming a SHO must have occurred during periods of time where sales is 2 standard deviations below the mean sales (Hausruckinger, 2005), to complex algorithms requiring naïve Bayes types of machine learning (Papkiriakopoulos & Doukidis, 2011) to estimate SHO. SHO estimated through POS data are generally used for studies on the antecedents to SHO, since the sales values cannot again be used to measure performance effects of SHO. POS data and perpetual inventory records are often confirmed with periodic manual audits. Without manual audits inventory records and sales records (POS data) may be only 45% and 85% accurate, respectively (Gruen & Corsten, 2008). Similarly, when directly observing retail shelves with manual audits, despite additional costs incurred from having store personnel or a third party physically at the shelves, SHO events are still not observed in their entirety (from the moment the product has stocked out until the moment of replenishment). This is why the two most often used methods to measure SHO are manual audits and POS estimates (Grubor & Milicevic, 2015) and inventory records used along with manual audits (Simicevic, 2015).

The remaining methods of measuring stockout occurrence, through survey and simulation, enable the study of both the antecedents and effects of stockouts but also have limiting considerations with their use as listed in Table 1-4. First of all, using surveys researchers can capture information about drivers, stockouts and outcomes such as sales.

These surveys are questionnaires to store personnel, not to customers (when customers are respondents, the study is to learn about intended customer reaction to stockouts). While store personnel responsible for shelf replenishment may know when stockouts occur, the information obtained through surveys is still general in nature, without providing exact stockout and replenishment times of individual items. Simulation is used to study which conditions drive stockouts (Eroglu, Williams, & Waller, 2011), how different lengths of stockout duration affects retailer market share (Wu, Huang, Blackhurst, Zhang, & Wang, 2013), or a system-wide analysis of how inventory record inaccuracy spurs stockouts which then affect fill rates and average inventory levels (Nachtmann, Waller, & Rieske, 2010). The starting values of parameters in simulations are often taken from a real-life numerical example or dataset.

None of the current stockout measuring methods actually capture entire stockout events, so stockout research has focused on different attributes that can be measured from a stockout event, as in Table 1-5. When the stockout event is an experimental condition, it essentially has a binary value of test (1) or control (0) (Huang & Zhang, 2016) or has a categorical value if the study focuses on how customer response varies depending on the cause of product unavailability (Anderson, Fitzsimons, & Simester, 2006). When stockouts are estimated or measured instead of researcher generated, Gruen and Corsten (2008) recommend that more than one of the following attributes of a SHO event are considered: the SHO frequency, duration (or availability), intensity, and the breadth.

Stockout Measures	Attribute type	Measure	Example research	Limitations
Item OOS event rate	Frequency	$\frac{Item\ SHO\ frequency}{Given\ unit\ of\ time}$	Gruen&Corsten (2008) Eroglu, Williams, Waller (2011)	-Must confirm replenishment between manual audits -Hard to find process issues
Category OOS event rate	Breadth	$\frac{Number\ of\ different\ OOS\ items}{Total\ number\ of\ items\ in\ assortment}$	Simicevic (2015) Gruen&Corsten (2008) Trautrims et al. (2009) Fernie& Grant (2008) Ehrenthal&Stolzle(2013) Papakirakopoulos (2012) Che, Chen, Chen (2012) Taylor & Fawcett (2001) Goyal et al. (2016)	-Hard to find problem items -A single snapshot in time, no duration
OOS duration rate	Duration	$\frac{Total\ time\ an\ item\ is\ OOS}{Given\ a\ period\ of\ selling\ time}$	Wu et al. (2013)	-Audits can't established moment of SHO, only from observation to replenishment -May consist of multiple SHO events which are masked with total SHO time
Lost Sales	Intensity	Unit $\frac{Total\ unit\ sales\ lost\ due\ to\ an\ item\ OOS}{Total\ units\ sold + units\ sale\ lost}$ Monetary $\frac{Total\ monetary\ loss\ due\ to\ OOS}{Total\ monetary\ sales + Loss}$	Grubor&Milicevic(2015) Anderson, Fitzsimons, Simester (2006)	-Must make assumptions about demand distribution (when and how many customers would have arrived during the SHO period) -Same issues as in OOS duration rate valid here

Table 1-5 Stockout measures by attribute type

Table 1-5 includes limitations to each type of stockout rate as well as examples of research using each attribute. The table does not list availability rate or customer impact rate. Availability rate is defined as (1-OOS duration rate) and still carries the same considerations as the duration rate in Table 1-5. Similarly, OOS customer impact rate is an intensity attribute with the same concerns as the other intensity attributes of unit and monetary lost sales rates. As this table shows, of all the different attribute types, breadth is the most commonly used and discussed (such as 8% global stockout rates) measure in empirical work by far (Gruen & Corsten, 2008) while duration is the least common (Wu, Huang, Blackhurst, Zhang, & Wang, 2013).

The least-often used attribute can be a problematic measure for multiple reasons. For SHO duration, the rate is the total time that the product is stocked out divided by the store's total selling time. For example, if a store is open for 10 hours, and within that period, a product is stocked out for 4 hours, then the stockout duration is: 4/10) = 0.40 or 40%. The duration attribute may give retailers a better understanding than the frequency attribute of how likely customers are to face an empty shelf. A 5% SHO duration is probably less of an issue than a 90% SHO duration in terms of the number of customers not being able to find the item in its dedicated space. However, the duration attribute requires knowledge of the exact moment of SHO occurrence and may be underestimated if it is being measured through shelf auditing since there may be a time lag between when the item is SHO to when the store is aware of the SHO.

Additionally, the duration attribute on its own does not provide enough information on where the replenishment process needs improvement. For example, consider two items, each with a weekly SHO duration of 30%. One item has a SHO frequency of 5 during that same time period while the other item has a SHO frequency of 1. The item which stocked out once stayed stocked out for a longer *consecutive* time period than the item which stocked out 5 times during the week. The total amount of time each one stocked out may be exactly the same (equal SHO duration) but the item which stocks out multiple times is being replenished multiple times throughout the week, while the single frequency SHO item is not. If the retailer has only SHO duration information it is possible that the inventory management of both items will be adjusted in the same manner. Instead it may be more appropriate to increase the shelf space allocated to the item frequently stocking out and the order size of the item stocking out only once.

The SHO frequency attribute also faces challenges when considered on its own. It is defined as the number of times the product is stocked out divided by a given unit of time. For example, if a product is stocked out 2 times in one 10-hour selling day, the frequency is 2/1 = 2 because the unit of time is one day. Care must be taken that the unit time is the same actual length of time across all observations for this frequency attribute. Within the unit of time where SHO frequency is counted, it is not clear if the item was not available that entire amount of time or for only a few minutes, so the frequency attribute does not provide the depth (or length) of any SHO event. An item which was stocked out once for a 10-minute portion of the 10-hour day would be viewed as equally stocked out as an item which was stocked out once for a 6-hour portion of the 10-hour day. Because of this, the SHO frequency attribute is more useful for seeing how an

individual item's availability has changed over time when only one aspect of its inventory management has been changed (larger order sizes, etc.). It may also be useful in comparing availabilities of multiple items over a small amount of time (store day or a few hours) to compare the pattern of SHO frequency over time from one item to another.

The intensity of a stockout refers to its effects on the retailer or to customers. It can be in terms of units lost (units lost to stockout/(units lost + units sold)), or sales loss which is similar to units lost in monetary terms. For example, if a store lost 40 of the original 100 units demanded by customers the intensity of the stockout would be 1-[(100-60)/100]= 0.4 or 40%. The customer impact rate considers the number of customers affected by the stockout in terms of baskets (Anderson, Fitzsimons, & Simester, 2006). It assumes if a customer faces a stockout, then all items from that shopping trip will be lost because the customer will purchase the entire basket of goods elsewhere. In this way it attempts to capture the intensity of a SHO event in terms of including lost sales of other items which may be in other product categories and may not even be SHO.

The breadth attribute is the most often used measure for product availability studies but also has its limitations. It is the number of different SKUs that are out of stock divided by the total number of SKUs offered. This is a snapshot in time. For example, at a certain time if 4 out of the 50 items offered in a product category were simultaneously stocked out the SHO breadth for the category would be 4/50 = 0.08 or 8%. This SHO attribute is useful for comparing entire categories, stores, manufacturers, or across stores. It is easily obtained by simply visually looking at a store shelf and counting the number of items that are SHO. Depending on the unit or grouping of analysis (category, store

over time, manufacturer, etc.) it may help retailers see which groups have a greater proportion of items available than other groups of items, but lacks any measure of each item's SHO pattern within each group. Additionally, the groups may vary in the total number of SKUs offered. For example, one product category may have 40 items in its assortment while another product category has 220 items in its assortment. A SHO breadth of 10% means 4 items are simultaneously stocked out in one product category while 22 items are simultaneously stocked out in another. Treating the 10% SHO rate as some sort of ordinal measure comparable across groups may make it difficult for retailers to decide which product category needs more attention.

Even when scholars use the same attribute across studies, the stockout measure itself can vary from study to study, making it difficult to compare stockout research across studies. For example, in one study (Trautrims, Grant, Fernie, & Harrison, 2009) the breadth attribute is a ratio of the count of the stores that have an SKU in stock at the end of a day, over the total number of stores carrying that SKU during a period of time, as a measure of availability. It is the availability rate for each individual item, and the authors call it OSA. Simicevic (2015) measures breadth of stockout in a different manner: "...if we observed three stores, two SKUs, and a period of 2 days, and for one SKU in one store for 1 day the inventory level at the end of the day was zero, the stock-out rate in this case would be 8.3% (1/12)." This measure is called the stockout rate, and is a measure of how many items in how many stores are stocked out for how many days (relative to the total sample). By definition an availability rate is equal to 1 minus (or 100 percent minus) the stockout rate (Gruen & Corsten, 2008), but in this example the two

studies do not allow for a comparison of rates even though they are both measuring the same attribute of SHO breadth.

Additionally, the terminology for stockouts is far from standardized and often does not properly express the attribute of stockout being studied. Stockout-related research uses terms and abbreviations about the topic such as: on-shelf availability, stockout, stock-out, stock out, out of stock, out-of-stock (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016), or OSA, SO, OOS. Since OOS and stockout refer to the same thing, the measure "OOS rate" (Grubor & Milicevic, 2015) and "stockout rate" (Simicevic, 2015) is expected to be a comparable measure of stockout. Instead the Grubor and Milicevic (2015) OOS rate uses a measure for the intensity of the stockout while the Simicevic (2015) stockout rate uses a measure for the breadth of the stockout. Even more confusing is when a study (Trautrims, Grant, Fernie, & Harrison, 2009) is said to be measuring availability (by definition based on SHO duration) but is measuring breadth of availability (based on count of products available among those offered). Trying to reach a consensus on either the antecedents or effects of SHO occurrence is difficult when the terminology is inconsistent or points to a different attribute than what is being measured.

There has been a push to bring a common set of language to stockout research, but even these recommendations have not been able to sufficiently capture stockouts. Gruen and Corsten (2008, p. 14) recommend as best practice to use their terminology to clarify what type of stockout research is being done: item OOS event rate (for frequency attribute), category OOS event rate (for breadth), OOS duration rate (duration) (and its

counterpart of shelf availability rate (availability)), OOS lost unit sales rate (intensity), OOS sales loss rate (intensity), and OOS customer impact rate (intensity). Given this framework, Grubor and Milicevic (2015) would use the term "OOS lost unit sales rate" and Simicevic (2015) would use "overall OOS event rate." Even in this example, the stockout terminology is not one the seven terms that Gruen and Corsten have suggested. Their terms do not sufficiently describe Simicevic's stockout rate as the number of total stockout occurrences for all SKUs in all stores for all days even though it is indeed a type of breadth measurement. Not only would scholars have to adjust the names Gruen and Corsten (2008) recommend to properly reflect their unit of analysis, but even with this terminology "OOS" simply refers to a shelf-out or "SHO" as presented in this dissertation. Researchers who are specifically limiting their research to SHO that are due to poor instore replenishment (IRO) or poor inventory investment decisions (STO) would benefit by making this clear at the very onset of introducing their study (Kotzab & El-Jafli, 2015) instead of the general "OOS" or "OSA" nomenclature inconsistently used today.

1.3.2 Antecedents and effects

Stockout research is known to exist with two separate streams of focus (Aastrup & Kotzab, 2010). Just as any phenomenon has antecedents and effects, stockout research too has streams that are concerned with what characteristics drive stockouts (antecedents) and how stockouts impact performance (effects). However, there is little research that considers both the antecedents to and effects of stockout, and this is most likely due to the difficulties in measuring stockout. For example, consider stockouts estimated using

POS data. Those estimated stockouts cannot be used to study their effects on sales, since the stockouts were estimated using sales. Similarly when stockouts are a factor in an experimental design it makes little sense to study the antecedents to stockout. Additionally, different areas of research (marketing, operations management, etc.) assign themselves to different aspects of retail shelf availability, so scholars have accepted studying stockouts in the two separate streams of antecedents and effects.

Studies on the effects of stockout look at different outcomes depending on the research topic. Focusing on store brands for example, since a stockout of one brand may mean an increase in demand of another brand (Borin, Farris, & Freeland, 1994) it is possible that retailers will choose to purposefully let items from other brands stockout so that their own brands with higher profit margins will be purchased instead (Shah, Kumar, & Zhao, 2015). Usually, however, research on the effects of stockout find poor performance results as listed in the introduction of the "Stockout research" section (Section 1.2). These studies are then usually used as practical motivation for future research on minimizing stockouts. If the research focus is on customer service, studies on stockout effects revolve around customer reactions to stockout. Customer reaction studies are in the dozens (Sampaioa & Sampaio, 2016) but generally cover behavior where the customer will either substitute (to another item whether or not of the same brand, or to another store), delay purchase, or leave (Zinn & Liu, 2001).

Since customer reactions may mean a sales loss for that stocked out item, for the brand, or for the store, studies look to optimize stockout for customer service and profit (Trautrims, Grant, Fernie, & Harrison, 2009). In this case the customer behavior becomes

the phenomenon of interest and research looks at the antecedents and effects of customer behavior, to be able to explain or predict the shifts in demand. The goal in operations management research is to keep the stockout event the focal phenomenon and study what the impact of resulting demand streams will mean for retail operations and inventory management. So the research focus is managing stockouts through better informed forecasting. The notion of stockout costs (Walter & Grabner, 1975) arose from studying the effects of stockout in this manner, and assortment planning research has a stream of analytic models to determine how much and which items retailers should hold, based on customer choice (Honhon, Johnalagedda, & Pan, 2012).Empirical research studying stockouts costs have only begun fairly recently (Anderson, Fitzsimons, & Simester, 2006) with quasi-experiments where customer response is studied to get a better understanding of reducing stockout costs through marketing tactics. In this case, the focus of the research is how to minimize lost sales (customer reactions that result in lost sales) instead of minimize stockout occurrence.

Studies on the antecedents to stockout also vary by research topic. A recent taxonomy classifies stockout drivers into 5 categories: operational (O), managerial (M), behavioral (B), coordination (C) and systemic (S) drivers (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016). A complete list of drivers under each category is provided, with permission from authors in the appendix (Appendix II). Operational drivers have to do with merchandise management and administrative retail tasks that decide how many units of a product are sent to a store, how many units are put on the shelf, how often they are replenished and other shelf management issues. Managerial drivers relate to employee training in terms of emphasis on the need to mitigate stockouts and to managerial

turnover which affects the manager's familiarity with store issues. Behavioral issues refer to various store employee actions ranging from errors to reckless disregard for store policies or processes. The authors refer to coordination drivers as issues of communication or being in synch with business units or firms outside of the store's replenishment personnel. Finally, systemic drivers do not have to do with store planning or execution as do other drivers. They are said to be "de facto constraints" that increase stockout risk and require "a structural alteration of the overall business strategy" (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016).

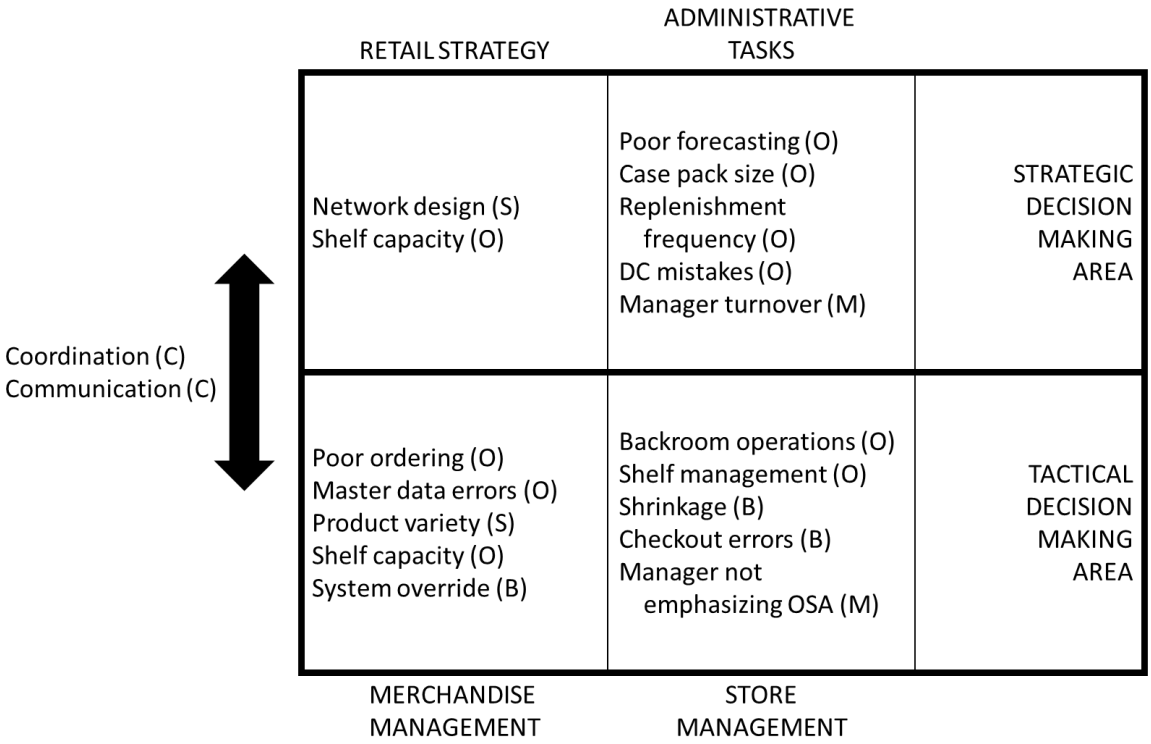


Figure 1-12 Stockout drivers (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016) organized in terms of the retailer's 4 decision making areas

Viewing these stockout drivers in terms of strategic and tactical retail decisions outlined in Section 1.1 of this dissertation sheds new insights. Figure 1-12 lists retailer-

based drivers from Moussaoui et al (2016) with a letter signifying their classification grouping in parenthesis. The systemic drivers having to do with demand (autocorrelation, unpredictability, velocity) are not included since they are not within the retailer's direct control. Each driver is placed in its decision area based on the definitions of strategic and tactical tasks reviewed in Section 1.1. A category of drivers may be split into different decision areas. For example, the driver "manager turnover" is listed as a part of the strategic human resources activities under administrative tasks because that is the decision area which includes hiring and training managers. However the "manager not emphasizing OSA" driver is a tactical store management task because it falls under the topic of managing store employees. Moussaoui et al.(2016) classify both as being in the "managerial drivers" category. Furthermore, it's possible to have one driver in multiple decision areas. For example, shelf capacity appears in two locations because the retailer determines the strategy, location (planograms) and flow of goods in consideration of both economies of scale and cannibalization, while the store itself can adjust facings[5] for store-specific needs. Communication and coordination issues now clearly fall between the different decision-making areas of the retailer, as do shipping size, shelf capacity and backroom replenishment issues. Research on backroom replenishment issues point to a misalignment between the ordering process, shipping size and shelf space (Eroglu, Williams, & Waller, 2013) and the schema in Figure 1-12 illustrates they belong to

[5] The retail shelf space for a product consists of a width, height and depth determined by the shelf dimensions. Items are allotted a certain portion of the total shelf capacity. The number of units, or consumer units, a product is placed horizontally in terms of width of space is a *facing* (e.g. a product with 2 facings has two consumer units side by side on the shelf). A product's *stacking* is when an SKU's consumer units can be placed on top of each other within the shelf's height (e.g. a product with a stacking of 3 has 3 consumer units in a vertical column). Finally, *row depth* refers to the number of consumer units that can be placed deep into the shelf (e.g. a product with a row depth of 4 has four consumer units back to back, from the front of the shelf to its back). Item shelf capacity is facings * stacking * row depth for each product.

different decision-making areas. Overall, all of these drivers are within some decision-making unit of the retailer and as such are “supply drivers” or antecedents to stockout.

The drivers removed from the list of currently known antecedents to stockout are all related to demand and should be considered “demand drivers” to stockout: demand velocity, unpredictability, and shrinkage. Shrinkage (Fernie & Grant, 2008) is a phenomenon that suffers from a lack of standard terminology or operationalization (RILA proposes to redefine 'shrinkage', 2016). However, since it refers to lost items generally stolen by store personnel or others, it can be considered a separate stream of product demand, even if “customers” aren’t paying for those items. Demand velocity (Taylor & Fawcett, 2001) has to do with how many units of a product are demanded within a unit of time. Higher velocity means that the number of units of a product on the shelf will be depleted within a shorter period of time than demand with lower velocity. Since the retailer’s core purpose is to supply incoming demand, higher velocity demand needs to be matched with more frequent shelf replenishment or greater inventory investment in terms of the number of units held on the shelf (capacity). Demand unpredictability (Corsten & Gruen, 2003) refers to a stochastic variation of the velocity of sales, or actualized (realized) demand. Increased variance in demand will lead to poorer forecasting, which may result in stocking too few or too many items. Too few items obviously cannot meet arriving demand, while too many items face the risk of being stored in some overflow or shared area (backroom etc.) where it is lost, forgotten, or otherwise leads to poor backroom (overflow) replenishment.

1.3.3 Overview of current theory

There are currently 3 areas of concern with empirical research of stockouts. The greatest obstacle is the lack of a precise measure of a stockout event. Without fully observing a stockout event, researchers have opted to test certain attributes instead. The second obstacle is the lack of standardized terminology and clarity in how stockout levels are measured for hypothesis testing. This lack of consistency makes it difficult to have a consensus or growing conceptual framework on stockout antecedents and effects, which is the third area of concern in this stream of literature.

Comparing studies with different stockout attributes may result in conflicting findings if the attributes themselves may have different relationships with the antecedent or effect being studied. Constrained product category shelf space may with an increased number of SKUs mean that each SKU has less space than if the variety was lower. With less space, there are fewer units of each SKU of greater assortment depth than there are of lower depth. With fewer units on the shelf, store personnel would likely replenish each SKU more frequently. The breadth of stockout, or simultaneous SKU stockouts, may increase, but the average duration for a stockout may not change or may even decrease if store personnel are in fact replenishing shelves more frequently. In this example, product assortment would have a positive relationship with "stockout rates" (frequency) in one case and a negative relationship with "stockout rates" (duration) in the other. This is why Moussaoui et al. (2016) state the "strong need for research exploring ways in which OOS could be tracked in a more efficient and reliable manner."

1.4 Product substitution

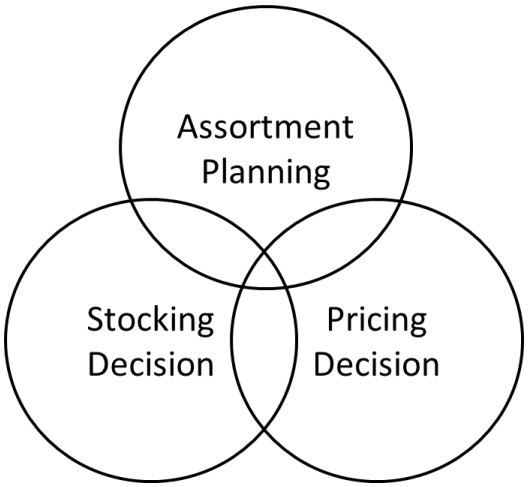


Figure 1-13 Planning for substitutable products, adapted from Shin et al. (2015, p. 687)

Retail management decision areas vary depending on what type of product substitution mechanism is being considered. Figure 1-13 is adapted from Shin et al (2015) to illustrate the different types of customer-driven substitution. Customers may switch from one item to another because of price differences, requiring retailers to carefully consider pricing decisions including discounts and promotions (Borin's in-store support effect in Figure 1-7). Assortment choices, or which products to carry, are a focus for substitution due to product-specific attributes. In this case, there may be a product with different features which drives customers to substitute, or the store may not carry the initially preferred product within its array of offerings (Borin's assortment effect). Finally, retailers focus on stocking decisions for stockout-based substitution. Since customers switch because of the stockout (Borin's inventory effect), retailers must consider decisions in merchandise and store management such as inventory review policy, order quantity, how much (and where) to hold and how often to replenish.

Because of the retailer's unique place in the supply chain, product substitution in retail differs from substitution in upstream members, especially for CPG items. In retail stores, customers take on the active physical role of walking through and looking at inventory and taking items from store shelves where similar items are also stored. The customer decides what to buy after the retailer has decided what to offer in terms of quantity, attributes or pricing. For upstream supply chain members, the customer may not even be aware that a substitution has even taken place if the firm has substituted a product component or sub-assembly during production. Upstream SC members either automatically substitute alternate components or subassemblies, backorder, or explicitly prompt customers to consider a substitute. Substitution actively decided by customers is known in the literature as customer-directed or customer-driven substitution. Supplier-driven substitution assumes the seller is the substitution decision maker and stocks, prices or selects a product assortment accordingly (Shin, Park, Lee, & W.C., 2015). Although some retailers may have personnel who actively prompt customers to consider available alternatives, CPG tend to be self-service items with product attributes and alternatives which are familiar to customers and readily apparent on store shelves.

Customer-driven substitution, though perhaps not as clearly distinct from supplier-based substitution as literature suggests in the retail setting, is the terminology used in this dissertation in accordance with current substitution literature. Since this dissertation only considers the inventory effect on brand sales and store management activities affecting shelf availability, the assortment- and pricing-based substitution is beyond the scope of this work.

1.4.1 General framework of stockout-based product substitution

Stockout-based substitution, or the inventory effect (Borin, Farris, & Freeland, 1994) occurs whenever a customer switches from an unavailable initially preferred product to an available alternate. For example, if mint-flavored toothpaste is not on its designated shelf, the peppermint- or wintergreen-toothpaste or the same flavor in another size or brand may be purchased instead. Indeed, Ge et al. (2009) found consumers observing that a product is sold out experience an "immediacy effect," making it more urgent to buy something immediately instead of delaying the purchase of the stocked out product. Known as customer switching, this type of substitution phenomenon constitutes only a small portion of risk pooling literature (Oeser, 2010). The pooling of demand occurs when spillover demand from a stocked out product is able to be fulfilled by an alternate product.

The stockout-based substitution phenomenon occurs as a combination of 3 conditions relying on both store (supply) and customer (demand) actions:

1. There are no consumer units (CUs) of an item the store usually offers on its designated customer accessible shelf.
2. Store personnel do not stock the empty shelf with any CUs of the item before customer arrival.
3. Customers arrive finding the shelf empty and a portion of them willing to switch to an alternate item do so.

The first condition for the customer-driven substitution to occur is a SHO, which can occur either as an IRO or STO. For an IRO, CUs are available within the store but are not on designated customer-accessible shelves for that item, which points to room for improvement in store replenishment processes. For a STO, the product is neither on its shelf nor at any other location within the store, which suggests either a lead time greater than accounted for with safety stock, or inaccurate forecasting in the administrative task area. Several studies show more than half (51-98%) of stockout incidences are caused by issues within the store (Gruen et al., 2002; Mckinnon et al., 2007; Ehrenthal and Stolzle, 2013) even though only 25% of SHO have been found to be IRO (Gruen and Corsten, 2008).

The second condition for the phenomenon to occur is that store personnel do not replenish the empty shelf space before customers search for that product. Shelf replenishment could be delayed for several reasons. Personnel are often not aware of the instant when the dedicated space stocks out. Some time may pass before knowledge of SHO. Personnel may also be occupied with other tasks already or may be interrupted by other tasks when trying to replenish that shelf space. If the shelf is empty because of STO, there simply are no consumer units of that item anywhere else in the store with which to replenish. If personnel have to search for the item in shared space, the search time may take longer than the time between the item SHO and the next customer arrival to that dedicated space.

The third condition that makes customer-driven stockout-based substitution possible is the customer finding the designated shelf space for the SHO item empty.

Since this substitution decision is based on the inventory effect (Borin, Farris, & Freeland, 1994) of a temporarily unavailable product in the store's assortment, the customer's behavior is said to be *dynamic*. Dynamic choice is when customers "base their choice from the products available in stock at the time of their visit to the store" (Honhon, Gaur, & Seshadri, 2010). Customers have individual preferences of products they would be willing to buy, ranked from the initially preferred product down to any number of items to which they are willing to switch when faced with product unavailability (Honhon, Johnalagedda, & Pan, 2012). Studies show that the majority of grocery store customers "would substitute if their favorite brand-size was not available" (Mahajan & van Ryzin, 2001) although there may also be customers who leave without purchase when faced with a stockout.

Customers may rank preferences according to price, any other product attribute, or just random whim, but the customer does not switch from the more preferred item to an alternate item unless faced with stockout. For those customers willing to substitute, they may prefer any number of alternate items and those items may be ranked in any order. All alternate preferences as well as the initially preferred product are included in the retailer's assortment but may or may not be available for purchase on store shelves at the time of intended purchase. Once faced with the stockout, the customer does not have to make a choice among available items. Instead the customer immediately switches to the next most preferred product, whichever that may be for a specific customer, to see if that alternate item is available. Analytically modelling such switching behavior is said to be NP-hard which is why heuristics are often used for this type of substitution (Honhon, Johnalagedda, & Pan, 2012).

It is necessary to define a specific set of terms to refer to subsitutable items within a store's product category more easily. An item which is the highest ranked item in a customer's preference ranking is referred to as the *preferred item*. If the preferred item is not available, it is referred to as the *stocked out item*, or *SHO item*. The customer's next ranked preference is referred to as the *alternate item*. It is possible that either or both the preferred item and any other alternate items may be unavailable. Any item outside of the preferred item is an unavailable or *stocked out* (or SHO) *substitute*. Since no single item in a product category is the preferred item for every single customer, researchers can select one item as a focal item and distinguish customers by this focal item. Customers who arrive at the store initially looking for the focal item are *initial customers* of their preferred item. Customers who have switched over to this focal item from their own SHO item are *alternate customers* to the focal item.

In stocking a focal item, the retailer accepts that a certain portion of arriving customers will not be able to buy this item because of stockout. How the retailer stocks and replenishes items is determined by the target fill rate, SL_i, it sets for each item. The target fill rate is a "subjective management judgment" (Tersine, 1994, p. 211) which accepts not meeting a certain portion of demand because it would be cost prohibitive to always hold much more inventory than the expected demand for an inventory cycle. If the retailer sets a goal of 90% service level (target fill rate), a series of calculations from inventory management theory--commercial software is often used (Hubner & Kuhn, 2012)--determine how much to stock and when to reorder the item to meet this target fill rate. The target fill rate, in other words, is a measure of the inventory investment the retailer must make to reach some desired level of customer service for an item. Outside of

a small margin of error due to calculations based on a normal distribution (whereas items can only be stocked in integer values), the retailer can expect to meet the SL goal and can serve 90% of customers (measured as fill rates) attempting to purchase the focal item.

What customer service levels (fill rates) are attained with the stockout-based substitution mechanism is involved? There is no straight-forward or simple answer to this question, as the theoretical model with a multitude of different customer paths helps to illustrate in Figure 1-14. The initial node at the very left of the model is the entire demand, *D*, for the product category, and may be averaged over any time interval (day, week, order cycle, quarter, year, etc.). The category demand *D* is split into a total of *n* possible subsets representing the initial customers for each item. For example, $D_X/\Sigma D_{i,}$ is the proportion of total category demand for initial customers of item X. The customers represented by the first set of links are able to either purchase the preferred item or find that it is stocked out. The proportion SL_i of each initial customer grouping is the marginal probability of those who are able to buy the preferred item, while the proportion who faces a stockout is $1\text{-}SL_i$ (or $100\text{-}SL_i$ if it is an integer percentage instead of in decimal format) As such, there are *2n* links of varying marginal probabilities in this second stage of the conceptual design.

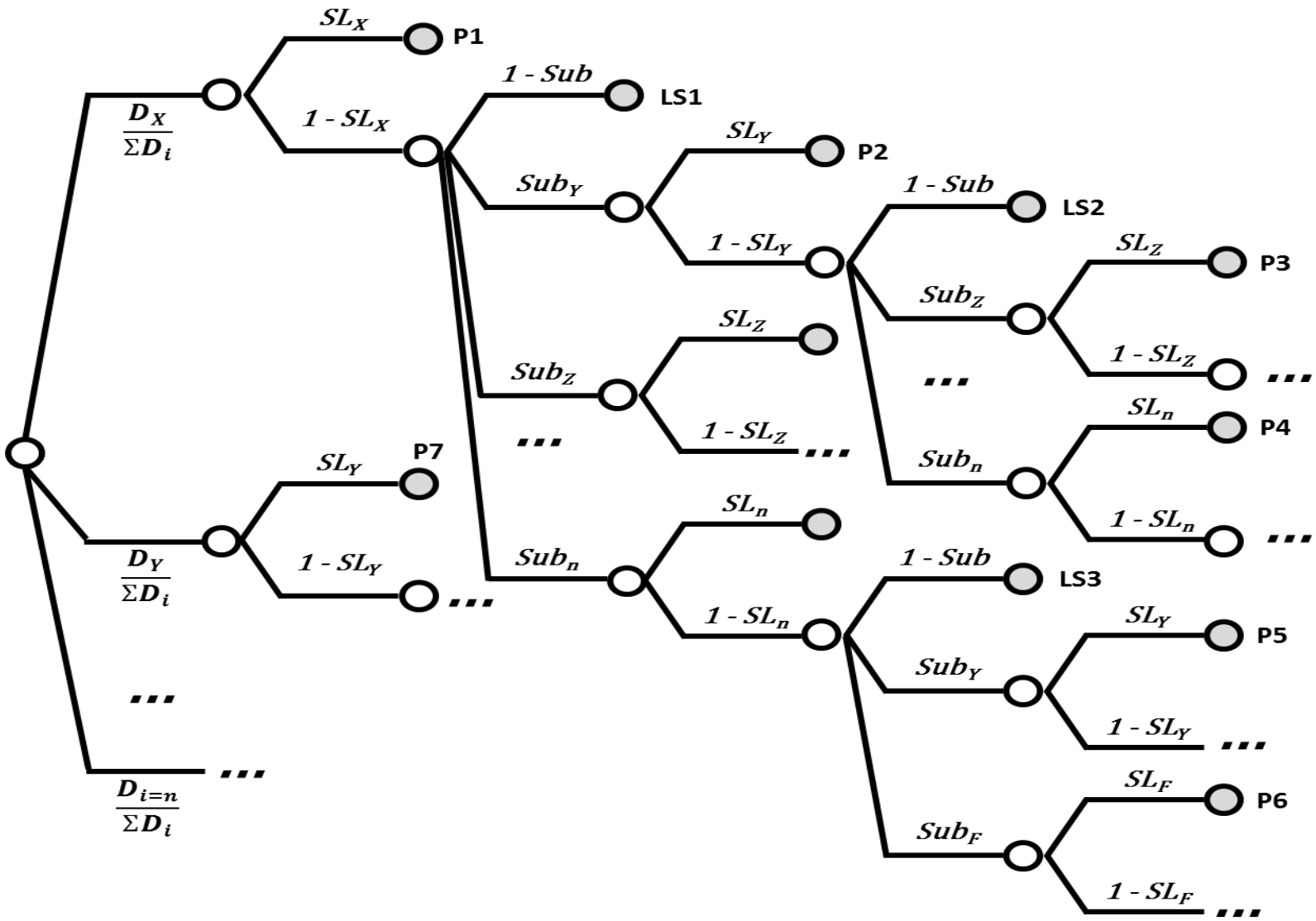


Figure 1-14 Theoretical model of customer switching with n products

Of initial customers facing item SHO (their preferred item is temporarily not on the shelf), a fraction of them are willing to substitute while the remaining customers are not, *1-Sub*. The customers willing to substitute, *Sub*, is a group comprised of *n-1* subsets, each subset representing a unique alternate item preference at some proportion Sub_r. Originally preferring product X, the customer in the first branch of Figure 1-14 now prefers alternate item "*r*," which replaces the stocked out product as the customer's alternate item. Item "r" is an alternate item or SHO substitute for the customer and the customer is an alternate customer for item "r."

For simplicity of presentation of this conceptual model, no further indices are provided for this fraction of remaining customers who are willing to substitute. The fraction of customers whose preferred item is X as opposed to Z may not have the same marginal probability of switching to item Y, Sub_Y. Likewise, the marginal probabilities of a customer preferring to switch to item Y after facing three stockouts as opposed to two stockouts may not be equal, so the various occurrences of Sub_Y in Figure 1-14 are not necessarily of equal value. To properly differentiate between all the possible permutations of customer preference the Sub_r notation would need *n* subscripts. Instead of mapping out every path a customer may take, this figure aims to illustrate the complexity of the customer switching phenomenon.

Customer switching allows for varied outcomes, with each customer either purchasing a product from the product category or leaving the retailer without a purchase if all alternate preferences are exhausted. These two possible general outcome types are represented in Figure 1-14 as shaded nodes, labeled for

convenience as *P* for units purchased and *LS* for units of lost sales. For example, P1 is the count of initial customers of item X who are able to find it in stock and leave the retailer with a purchase with conditional probability $D_X/\Sigma D_i * SL_X$. While P3 is also an outcome where a product is purchased, its conditional probability is the product of the marginal probabilities in its own tree path: $D_X/D_i * (1-SL_X) * Sub_Y * (1-SL_Y) * Sub_Z * SL_Z$. If the conceptual model in Figure 14 had a total of only two items in the category, there would be 8 possible outcomes with different conditional probabilities calculated in this manner.

The number of outcomes possible increases exponentially with increasing product category size, making planning for substitution difficult for both researchers and practitioners to manage. While there are 8 possible outcomes for a category with only two items, there are 3,912 outcomes possible for a category with six items. Alongside scholars currently unable to analytically model dynamic switching behavior, retailers too face increased complexity in their inventory planning and management activities. Wan et al., (2012) find that, for one distribution center, 84 items in their product category best provide increased sales by capturing demand without being offset by the operational costs of managing a variety of products. A high variety of products in retail stores is also linked to greater SHO occurrence (Ton & Raman, 2010) because of phantom products—items within the store but not accessible to customers. Stores may want to capture different customer segments by offering a variety of items and thus reaching a larger market size. But it is more complicated to use product variety to increase fill rates by offering alternate items when initial customers face a SHO item. This is because each item has one initial

customer grouping and up to *n-1* alternate customer groups demanding the item. Rajaram and Tang (2001) call this aggregated demand for an item its "effective demand," and say it is the sum of primary demand (initial customers purchasing the item) and derived demand (alternate customers purchasing the item). As the number of items increases, the derived demand may take up a greater portion of the total demand for the focal item.

However, since the dynamic switching requires 3 conditions to take place, modeling this type of substitution needs simplifying assumptions. While there are many different types of simplifying assumptions to model the stockout-based substitution mechanism, the following are used more often:

- that product demand is known or deterministic (Shin, Park, Lee, & W.C., 2015)
- single-period newsvendor models (Khouja, Mehrez, & Rabinowitz, 1996) where a product's inventory position is not carried through from one period (order cycle) to another and items simply disappear (no inventory value) after the period ends,
- equal costs and prices between items, that assortments are fixed with only two items (Karakul & Chan, 2008) or that substitution can only occur in one direction (Karakul & Chan, 2008).

According to the Shin et al. (2015) taxonomy on substitution literature, which classifies substitution models based on assumptions, this dissertation's substitution mechanism is an inventory-based (stockout-based), customer-driven (customer switches to an alternate item when the preferred item is stocked out), two-way

substitution studying retail stocking decisions. The two-way substitution can occur at varying degrees without being symmetric between two items, and the inventory position is carried over to the next cycle to be able to capture the effects of customer switching on performance fill rates.

1.4.2 Antecedents and effects of the substitution phenomenon

For the product substitution risk pooling mechanism to be applicable, customers must be willing to switch from one item to another. If customers are willing to buy an alternate item when the initially preferred item is stocked out, then the retailer is able to manage store shelves for the varying individual streams of demand for each product as an aggregate demand. Likewise, if the lead time of each substitutable item varies, customer willingness to switch from the SHO item to the alternate item allows the retailer to stock for a pooled lead time variability. The retailer may lower inventory investment overall while meeting customer service goals, or reach even higher customer service levels than the same inventory investment for non-substitutable items (McGillivray & Silver, 1978). In this way the substitution pooling approach enables retailers to better meet the opposing goals for customer service and inventory investment.

It is not just the retailer who is affected by stockout-based substitution. The supply chain consists of a minimum of 3 roles: the consumer, the retailer and the manufacturer or brand owner. Each role as well as the supply chain overall may be affected by customer switching in terms of a change in performance. The concept

behind the classic performance measure of fill rate is illustrated in Figure 1-15. The fill rate for item X is *P1/D1*, and is *P2/D2* for item Y. This fundamental measure is also known as item fill rate and corresponds to the target service level the retailer expects to meet in stocking each product.

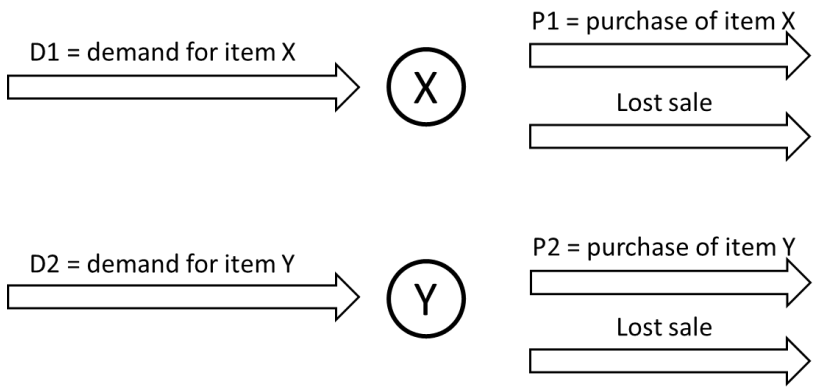


Figure 1-15 The classic view of item demand and sales in determining fill rates

Fill rate measurements, however, are not nearly as straightforward when customer switching is involved. With customer switching, a portion of D1 in Figure 14 may purchase item Y, and similarly demand from D2 may shift to item X when item Y is stocked out. In the generalized case with more than two items, switching behavior makes an even more significant difference in fill rate calculations. Indeed, if the conditional probabilities explained in the general model are fill rates, a basic question is where in the general model (Figure 1-14) the classic definition of fill rate is accounted for.

For example, shaded node P1 is the purchase of item X by a customer who initially prefers this item. This is not the product fill rate for X, because the retailer is also able to sell X to alternate customers. For this subset of category customers who initially prefer X, P1 is the only outcome where their initial demand is met. Customers, who are aware of the high availability of their initially preferred product X, could prefer to go to this retailer or store first (since retailers and stores are also substitutable as are the products they carry). Retailers, likewise, may be guided by knowing how much of the sales of a product goes to those customers initially preferring it. They may adjust their stocking and replenishment of other items accordingly. Finally, the fill rate at P1 may be a useful measure for the brand-owner, who may opt to start or stop promotional or packaging efforts attempting to capture new customers.

However, this P1 rate is not a measure that answers two vital questions. First of all, what is the customer fill rate, or what percentage of customers initially preferring item X had their demand met by purchasing some item (initial or alternate) from this retailer? Secondly, what is the item fill rate, or of people looking to buy item X (as an initial or alternate preference) what percentage of them were able to buy it? While the first question follows a group of customers to find what proportion of them leave the retailer with a purchase from their intended category, the second question is more along the lines of the classic definition of (item) fill rate. This measure does not exist anywhere in the theoretical model of customer switching. In practice, however, items are stocked using this fill rate concept and traditional (static) choice models.

Stocking items according to this classic fill rate measure can count the same individual customer multiple times, for different items. For example, in Figure 1-14, P7 is when customers who initially prefer item Y are able to purchase it. P2 and P5 are the outcomes for customers who initially prefer item X, but purchase item Y after facing one or two stockouts, respectively. When considering what proportion of "item Y demand" (customers attempting to purchase Y as their initial or alternate preference) the retailer has been able to meet, the customers from P2 and P5 are included to "item Y demand." However, they are also included in "item X demand" since this was their initial preference and faced item SHO. How about for a customer who initially preferred product X but was also not able to buy item Y? P6 is an outcome for customers initially preferring item X and facing two stockouts, but in this outcome item F is purchased instead of Y. In all of these examples, the customers moving from one item to another in their preference rankings are considered unique individual demand for every alternate item as well as the preferred product.

Since the classic definition of fill rate is nowhere to be found in the conceptual model of customer switching, for whom is the retailer setting the target service level of a product? When the goal is met, what does this mean? Where customer switching is involved, the retailer's fill rates need to properly account for those who initially prefer it, as well as those who face up to *n-1*stockouts. Customers could face multiple stockouts, for example, if the customer's initial preference and higher-ranked set of alternate preferences are all from the same brand, and that brand has yet to be replenished at that retail site. In this sense, it is useful to also be able to discriminate between what proportions of demand for a supplier's brand is being filled, along with

what proportion of that brand is being purchased by those who initially prefer it. The classic notion of fill rate does not sufficiently describe this phenomenon nor help clarify the phenomenon's impact on different chain members.

Table 1-6 lists the various performance fill rate measures which may be considered in customer switching for a more expanded analysis of the impact of the phenomenon on the supply chain. Category fill rates are included as a measure for the retailer to gauge how much of the total demand for a category it is able to meet, regardless of item, brand, or customer subsets. Based on the theoretical model, if there was absolutely no customer switching, these different measures would be without distinction. The product fill rates would be no different than the customer or brand fill rates, so that in aggregate, they would also be no different than category fill rates.

Fill rate measure	Fill rate numerator: Units of sales	Fill rate denominator: Number of arriving customers	Conceptual model example	Prior work
Item	All sales of item, regardless of customer's initial preference	Those who initially prefer item + Those alternately prefer this item after stockout of higher ranked preference	*Item X fill rate* $\frac{\Sigma P7 + P2 + P5 \ldots}{\Sigma\, D_Y + D_X(\ldots Sub_Y + \cdots) + \cdots + D_n\,(\ldots Sub_Y + \cdots)}$	Military logistics (Hauk, 1964)
Customer (overall)	All purchases by customers with same initial item preference	Those who initially prefer item	*Customer X fill rate* $\frac{\Sigma P1 + P2 + P3 + P4 + P5 + P6 \ldots}{D_X}$	(Zinn, Menzter, & Croxton, 2002)
Customer (Initial pref.)	Purchases by customer of same initially pref. item	Those who initially prefer item	*Customer X direct fill rate* $\frac{P1}{D_X}$	Retail inventory management (Tan & Karabati, 2013)
Brand (overall)	All same-brand item sales regardless of customer's initial pref.	Those initially prefer a product within same brand + those alternately prefer a product in same brand after stockout of higher ranked preference in same or other brands	*Brand A (ex. Items X&Y) fill rate* $\frac{\Sigma P1 + P2 + P5 + P7 \ldots}{D_Y + D_X}$	Brand sales (Borin, Farris, & Freeland, 1994)
Brand (Initial pref.)	All same-brand item purchases by customers of same initial brand pref.	Those who initially prefer a product within same brand	*Brand A (ex. Items X&Y) direct fill rate* $\frac{P1 + P7}{D_Y + D_X}$	Perishable products (Tang, Rajaram, & Ou, 2002)
Category	All purchases made of every item in category	The entire category demand (all customers regardless of preferences)	*Category fill rate* $\frac{\Sigma P1 + P2 + P3 + P4 + P5 + P6 + P7 \ldots}{D}$	Product commonality (Baker, Magazine, & Nuttle, 1986)

Table 1-6 Performance fill rate measures for theoretical model

1.4.3 Overview of current theory

The stream of research on stockout-based substitution face empirical and analytical hurdles. Empirically, stockout-based customer switching decisions are not observable, and analytically modelling customer reactions and store performance effects requires simplifying assumptions for tractability (Honhon, Johnalagedda, & Pan, 2012). The simplifying assumptions include one-way substitution (Honhon, Johnalagedda, & Pan, 2012), uneven costs or profits for the store depending on whether the preferred or alternate item was purchased (Gilland & Heese, 2013), or single-period news vendor models, which assume either product perishability or fixed cycle replenishment and no lead time (Smith & Agrawal, 2000). Even without such simplifications, classic inventory models assume financial measures, such as costs for under- or over-stocking items in the newsvendor model, are known. Even if holding costs were not dwarfed by instore handling costs (Curseu, Woensel, Fransoo, Donselaar, & Broekmeulen, 2009), this newsvendor model assumption limits the retailer's application of such research findings since they are typically unaware of such costs and must rely on benchmarks or approximations.

Approximations or specific data gathering are necessary for empirical research as well, since a substitute purchase is not observable even if stockouts could be observed. A measure of elasticity (Eisand, 2014) is used for substitution between two goods that can be studied empirically by analyzing price, space, or other attributes and estimating how any one of these parameters is linked to differences in sales. However, for the stockout-based substitution mechanism, empirical work has

shown that the number of facings, or the amount of space for each product does not have significant impact on sales "as long as a minimum threshold (to avoid out-of-stocks) was maintained" (Dreze, Hoch, & Purk, 1994).

Observing the stockout based substitution mechanism requires observation of the shelf availability as well as customer arrival and preference rankings mentioned in the general framework earlier. To capture stockout-based substitution, the point at which the customer discovered the stockout must be known and must occur before a shelf replenishment takes place, which must also be known. Such information can be known either in a field experiment or through surveys. By far the most often used method of observing when a customer switches because of stockouts is to survey customers about a prior stockout experience (Zinn & Liu, 2008).

The two closest empirical studies observing availability as well as customer arrival and buying preferences still vary greatly from the context of a CPG store. A quasi-experiment by Anderson et al. (2006) gather data from real customers calling a catalog sales company which experiences actual stockouts of items from time to time. The focus of the study is how a firm can mitigate loss of overall sales (basket orders involving more than one item) during a stockout of a focal item. During an actual stockout, the authors have the catalog personnel offer to backorder the focal item in any one of five different ways: inform the customer that the preferred item is stocked out (control), explain that it is a stockout out due to a supplier issue, stocked out because it is an extremely popular product, to offer a shipping discount, or to offer a dollar discount on the backordered item. In this case the stockout event and the

customer's demand for a stocked out item are observable. However, the focus is on what happens to the overall customer order because of the stockout of a focal item. Customers in the study are said to prefer "one-stop shopping" so when the preferred item is stocked out, then other items that are in their order are cancelled at a higher rate (10%) than when the focal item is in stock. The focal item is said to not be substitutable so the customer switching that occurs is presumed to be to another seller of this product category. This study is not like self-service retail settings, but points to the importance of the customer's willingness to switch since there is a higher rate of losing the sale of an entire basket of goods due to the stockout of one item.

The second study actually observes both stockout and customer switching in a vending machine setting. Anupindi et al. (1998) gather data from 3 vending machines in different locations which carry the same 6 beverages in their assortment. They are able to observe a product's stockout in its entirety along with the sales of alternate items during the stockout. They assign substitution rates between products themselves a priori and based on intuition and knowledge about the brands and product attributes. Customer arrivals are assumed to be Poisson with mean hourly observed values which are then used to estimate how much demand may have switched over to alternate items. Their research question is whether sales can be used as a guide for demand levels when products are substitutable and they compare sales to maximum likelihood estimates, finding that sales estimates are biased with too high values for products that are highly available and to which customers would switch if faced with stockout of another product.

The issues facing substitution research build on those facing stockout research. Stockout research faces fundamental issues of measurement and standardization, and substitution research must account for those issues as well. On top of this, substitution research faces lack of observability of customer preference rankings (Honhon, Johnalagedda, & Pan, 2012), the arrival sequence of different customers to a product on the shelf (Gilland & Heese, 2013), and whether an item's purchase was made by an initial or alternate customer (Tan & Karabati, 2013). Current substitution theory faces other hurdles beyond the scope of this dissertation. For example, customers in this dissertation already have a preference ranking which is a set list of items the customer will buy, from the initially preferred product to any and all alternate items. This preference ranking is a static list and does not account for modified demand through in-store support effects (Borin, Farris, & Freeland, 1994) like discounts or promotions. In such a situation of demand reshape (Eynan & Fouque, 2003) an item ranked as the 4th item in a customer's preference rankings may become the second or even most-preferred item. Just as customers make a dynamic choice depending on the item availability, their choice list may also be dynamic depending on changing item attributes (store efforts).

1.5 Logistics postponement

While recent attention has been brought to the "need to further develop the postponement concept to match the extending scope of its applications" (Yang, Burns, & Backhouse, 2004, p. 483) the literature is still lacking on the application of this concept in the context of product availability on brick-and-mortar shelves (Jafari, Nyberg, & Hilletofth, 2016). Despite the fact that 71% of surveyed US shoppers said they prefer brick-and-mortar stores over buying the same items online (Stanley, 2016), the current research of postponement carried out by retailers is based on online shopping. Investigating how using logistics postponement within a brick-and-mortar store affects product availability and store performance adds to the knowledge on postponement.

1.5.1 Theory development and contexts

Although implemented by firms in the 1920s (Yang, Burns, & Backhouse, 2004) marketing literature first described postponement three decades later as a way of dealing with demand uncertainty. Alderson (1950) stated that delaying activities on the form, identity or inventory location of an item to as late a time as possible would lead to more efficient distribution tasks. Distribution tasks would become more efficient if they focused on areas that most needed them. To learn where marketing tasks were most needed, postponement theory was based on the implicit assumption that during the time period when supply tasks are delayed the firm would receive or

obtain more demand information. New information about current unfulfilled customer demand during this period of delay serves to reduce demand uncertainty. Reduced uncertainty through postponement allows for better decisions on where in the supply chain a product is differentiated or when it is deployed (Yang, Burns, & Backhouse, 2004). These delayed activities are made at a decoupling point, "where in the supply chain the customer order penetrates and that distinguishes forecast and order-driven activities" (Yang, Yang, & Wijngaard, 2007, p. 974). This decoupling point is also known as the "push-pull boundary" so that "pull postponement" (Gattorna, 1998, p. 80) means that the decoupling point is higher up in the supply chain. Pull postponement is synonymous with "time postponement" (Reimann, 2012) which is Alderson's concept of deciding the time at which to differentiate a (unfinished) product or to deploy a differentiated (final) product.

Business logistics theory (Bucklin, 1965) took Anderson's postponement theory and combined it with the concept of speculation, to incorporate hedging or buffering as additional way of dealing with demand uncertainty. This hedging approach is the "converse" of postponement, Bucklin stated, and was developed earlier as a pooling method with "speculators, by shifting uncertainty to themselves, used the principle of grouping, as insurance, to transform [uncertainty] into the more manageable form of a relatively predictable risk" (Bucklin, 1965, p. 27). Speculation leads to "intermediate inventories" which form "whenever its additional costs are more than offset by net savings in postponement to the buyer and the seller" (Bucklin, 1965, p. 31). In other words, stock may be kept in a intermediate space if holding or owning the product costs less than what is gained by physically moving the product at

a different point of time. Bucklin's development of Alderson's postponement theory emphasizes that the title-holding "merchant middleman" keeps inventory in an intermediate place that *carries the risk of not being sold* (Bucklin, 1965, p. 30). This differs from temporarily storing an item while distributing it, as that temporary storage carries no risk of sales loss, and it is also different from the classical view of speculation in economics where items are purchased speculatively with "the expectation of an impending change in the ruling market price as the sole motive of action" (Kaldor, 1939, p. 1).

Hedging against and reducing demand uncertainty through postponement and speculation (P/S) has been studied in many different contexts, all depending on the state of the product (Yang, Yang, & Wijngaard, 2007). If the product is in the stage of research and design, the context is at the manufacturer and supplier levels under product-development postponement or ordering (of raw materials or components) postponement (Saghiri & Barnes, 2016). If the product is semi-finished (work-in-progress) then the focus is production postponement in the manufacturing and assembly context, having to do with when to start production, when to complete it or when to package and distribute the product (Boone, Craighead, & Hanna, 2007). While packaging (Tse, et al., 2012) and distributing (Jishan & Yajun, 2011) can sometimes be in the context of warehouse added value, these postponement decisions have generally been studied in the manufacturing context. All states of the product before it is fully finished are form and identity postponement (Alderson, 1950), also classified as process-level postponement (Zinn & Bowersox, 1988) and are beyond the scope of this dissertation. Indeed, earlier reviews of postponement measures have

emphasized that these types of postponement are "not appropriate for firms at the end of a supply chain (distributors and retailers)" (Li, Rao, Ragu-Nathan, & Ragu-Nathan, 2005, p. 633).

		Logistics	
		Speculation *Decentralized inventories*	**Postponement** *Centralized inventories and direct distribution*
Manufac-turing	**Specul-ation** *Make to inventory*	**The full speculation strategy**	**The logistics postponement strategy**
	Postpone-ment *Make to order*	**The manufacturing postponement strategy**	**The full postponement strategy**

Figure 1-16 Postponement and Speculation matrix from (Pagh and Cooper, 1998)

Postponement and speculation can take place in terms of manufacturing or logistics, as in the P/S matrix by Pagh and Cooper (1998, p. 20), in Figure 1-16. In terms of logistics P/S, items may be quickly moved downstream to decentralized locations at forecasted amounts (speculation) or there may be a delay of forward movement until after customer orders have been received (postponement). The full speculation strategy (top-left cell) refers to speculative behavior in both manufacturing and logistics, such as stocking CPG on store shelves. Manufacturing postponement strategy is when the items are at decentralized locations (stores) but are not in their final form. Examples are souvenir t-shirt printing stores or custom-decorated cakes in bakeries. The full postponement strategy is when goods are not in

their final form until customers have placed the orders, and are stored at either regional warehouses or at the manufacturer's site where they can be turned to their final form after receiving customer orders. An example is personal computer companies where customers purchase the products online and the manufacturer or regional warehouse brings together subassemblies to meet the product specifications of the purchased item. Last of all, the logistics postponement strategy refers to bringing items to their final form before customer purchase and storing them at centralized sites or at the manufacturer. Examples are having retail distribution centers or manufacturers drop-ship previously finished items directly to customers and bypassing stores.

Empirically, logistics P/S in retail has been studied mainly in terms of online retailing. According to Bailey and Rabinovich integrating both postponement and speculation in a dynamic manner is neither studied nor implemented by most firms (2005, p. 163). For online retailers, however, they state that there is "dynamic" use of P/S, where retailers will use both the full speculation and the logistics postponement strategies for their items. Taking into consideration holding, ordering and drop-shipping costs, Baily and Rabinovich develop an analytical model of total costs as a function of product popularity and of firm market share. Their model proposes and numerical analysis supports that as the popularity of items and firm market share both increase, firms will be incentivized (by reduced costs) to both drop-ship items and hold them in stock. Increased product popularity is linked to decreased demand uncertainty, they argue, so retailers would lean towards drop-shipping to further reduce uncertainty, instead of hedging against the uncertainty of a popular product

through speculative stock. The authors do point out that "…in traditional retail environments, […] inventory replenishment decisions are centered on either inventory location speculation *or* postponement" (Bailey & Rabinovich, 2005, p. 161).

The current view of logistics postponement activities being implemented as either "speculation or postponement" in brick-or-mortar environments is partially due to the sparse literature in the retail context, which is a "missing link" in postponement theory development (Jafari, Nyberg, & Hilletofth, 2016). Jafari et al. assert through case studies of three retailers that postponement methods in the retail context differ from existing postponement studies because of the unique "nature of [a retailer's] operations" (p. 460). Pagh and Cooper's P/S matrix (1998) consider make-to-stock items as either in a decentralized (logistics speculation) or centralized (logistics postponement) location. The possibility of postponing the forward movement of some consumer units of items onto store shelves while also speculatively stocking them in a store location that faces the risk of not being sold has received little research attention. Retailer operations differ from other supply chain member operations since their customers serve themselves to items on store shelves. Not only do retailer operations differ from other supply chain members, but existing studies of the postponement construct necessary for theory development revolve around antecedents to or enablers of postponement (Li, Rao, Ragu-Nathan, & Ragu-Nathan, 2005), instead of performance effects of this risk pooling method. Measures

Paper	Postponement	Speculation	Logistics Postponement
Rabinovich and Evers (2003) [empirical; archival data]	Constructs from survey responses: Time Postponement (Likert scale of how frequently transshipment occurs at 3 different product stages) Form Postponement (Likert scale of degree of customization of products)	Single measure construct: percent of units manufactured to inventory	
Jafar, Nyberg & Hilletofth (2016) [qualitative; multiple case study]	Time postponement activities were considered only in terms of non-finished goods.		Inventory centralization, reshoring manufacturing
Li et al. (2005) [empirical; factor analysis for various SCM constructs]	Form Postponement: products designed for modular assembly, products assembled only after customer order received	Unfinished goods: modular process can be rearranged so final product assembly is at distribution center (DC), (final assembly delayed until last possible position in SC, closest to customers. Finished goods: goods stored at distribution points close to customer	

Table 1-7 Postponement and speculation studies in the retail context

This overview of postponement measures covers only the time and place postponement of finished goods, or logistics postponement (Yang, Burns, & Backhouse, 2004). Only finished goods-related postponement methods are considered since the retail store by definition functions as a location for finished goods made available to customers (Wagner, Ettenson, & Parrish, 1989, p. 60). A useful overview of postponement for unfinished goods are in the proceedings of Zhang and Tan (2001) with construct overviews in Li et al. (2005).

Logistics postponement measures of finished goods encompass a limited area of postponement literature, as listed in Table 1-7. Rabinovich and Evers (2003) find time postponement positively linked to form postponement, while form postponement is negatively related to speculation. Time and form postponement along with information systems are more negatively linked to speculation. Jafer et al. (2016) find that among SC members logistics postponement is most widely used in retail. Retailers in the case study increasingly link postponement to flexibility. Finally, Li et al. (2005) develop and validate a measurement instrument for various supply chain practices including postponement. They consider 5 different postponement measures and the final instrument for postponement removes process reordering of unfinished goods and final goods stored at distribution points close to the customer. This latter measure has low factor loading and is the only logistics postponement measure in the study. Their final postponement measures in their study involve form postponement or place postponement for unfinished goods.

1.5.2 Overview of current theory

Logistics postponement, comprised of when to move forward and where to store finished goods inventory, does not currently have an accepted measure in the literature. Indeed, even in general "[d]espite numerous studies on different aspects of postponement theories and applications, limited measured variables have been developed to assess postponement" (Saghiri S. , 2011, p. 6428). Not only is there a call for developing measures of postponement, but the need for studies on "new postponement strategies" at other portions of the supply chain (Boone, Craighead, & Hanna, 2007) such as at the downstream or retail level, is emphasized as future work in postponement literature.

In retail, inventory risk pooling is generally concerned with shelf space (Hubner & Kuhn, 2012) and the backroom, which is an alternate storage location that is not accessible to customers. Since each replenishment order represents an equal chance for a delivery delay (Bowersox & Closs, 1996), holding stock in the backroom reduces the uncertainty to within the store, an area within the store manager's control. Having inventory of a product within the store and successfully using backroom inventory to replenish shelf space is called the store-level fill rate effect (Waller, Heintz Tangari, & Williams, 2008).

A store's backroom (shared space) can serve as a tool in both stocking and distribution tactics to store safety stock and possibly even some cycle stock in a speculative manner (Bucklin, 1965). By minimizing the shelf capacity of each

product, the retailer can provide a greater depth of assortment within the constrained product category shelf space. Backroom stock is then "distributed forward" to customer accessible designated shelf space whenever the product needs replenishment. While the store-level fill rate effect is concerned with lead time uncertainty, the use of postponement and speculation deals with demand uncertainty within shelf space constraints. Conflicting results on whether backroom use improves (Milicevic & Grubor, 2015) or hurts (Eroglu, Williams, & Waller, 2013) firm performance may also rest with product attributes driving demand variability in addition to the backroom effect (Waller, Heintz Tangari, & Williams, 2008) or the store's ability to properly replenish the shelves.

1.6 Research deficit and questions

Research deficits concerning retail shelf availability

- Detection of SHO – when does it occur
- Stockout measurement
    - Actual observation of event (duration, frequency)
    - Use of more than one stockout attribute in an estimation
    - Any relationship between SHO attributes
- Bringing supply and demand drivers of SHO together
    - Isolating customer switching's effect on sales
    - Accounting for sales as a demand driver to stockouts
- Stockout terminology
    - Differentiating between shelf, store, and replenishment stockouts
- Role of postponement within instore logistics

Table 1-8 Current gaps in retail shelf availability research

This section summarizes and highlights the gaps in retail shelf availability work (Table 1-8) and the product availability questions by chapter (Table 1-9).

Retail shelf stockouts (SHO) are an integral part of substitution and postponement risk-pooling mechanisms and, as listed in Table 1-8, limit their theory development because of measurement issues. Retailers fearing that SHO occurrence mean lost sales, may carry alternate items (substitution) or more consumer units of an item in their store than their shelf capacity can hold (postponement). However, neither researchers nor retailers have full information on SHO occurrence, so risk pooling methods are usually linked to a single SHO attribute instead of the event itself. Most of the time only one SHO attribute can be used to study either the antecedent or performance effects of SHO. The relationship between SHO attributes is not clear. The lack of clarity complicates efforts to determine if the risk pooling methods mitigate lost sales from SHO or decrease SHO occurrence.

This dissertation attempts to circumvent SHO measurement issues in order to explore the research questions listed in Table 1-9. Chapter 2 finds a way to get around the measurement problem by recreating a store environment where SHO events organically occur within a discrete agent-based simulation program. Chapters 3 and 4 use initial data captured from a shelf liner technology which allows the researcher to have information on exactly when and for how long a shelf space is empty. In this way, SHO events are neither an experimental factor imposed by researchers nor an estimated occurrence using realized sales figures.

Once there is no need to assume or approximate if and when a SHO event occurs, it is possible to evaluate if and when the risk pooling methods benefit the firm, as intended. In Chapter 2, the number of initial and alternate customers buying an item is thus measurable so substitution's role in the relationship between inventory investment and customer service can be further studied. This also allows for a comparison of different fill rate measures, and the size of the effect of substitution on the relationship between inventory investment and customer service. In Chapters 3 and 4, item-level SHO information expands retail shelf availability knowledge by linking SHO antecedents to performance effects in a single model and using multiple SHO attributes. The retailer's unique relationship with customers in terms of having two different inventories (store shelves and backroom) contributes to logistics postponement theory as being a valid tool to apply within a store. It also opens a new avenue of research to further explore antecedents and effects of instore logistics postponement.

Ch.	Research Questions	Attribute	Inventory investment	Operational efficiency	Customer service
2	How do customers switching from one item to another because of stockouts affect the retail store's customer service levels?	Availability (Breadth)	X		X
2	Can a simple heuristic about the inventory effect serve as a reasonable rule of thumb for stocking decisions of substitutable products?	Availability (Breadth)	X		
3	Does logistics postponement exist within a firm?	Duration, Breadth, Frequency	X		
3	Does instore logistics postponement improve product availability?	Duration, Breadth, Frequency	X	X	
4	How does product availability in terms of both stockout duration and frequency affect sales of that item?	Duration Frequency		X	X
4	How are the stockout attributes of duration and frequency related to one another?	Duration Frequency		X	

Table 1-9 Research questions by chapter

1.7 Research framework and design

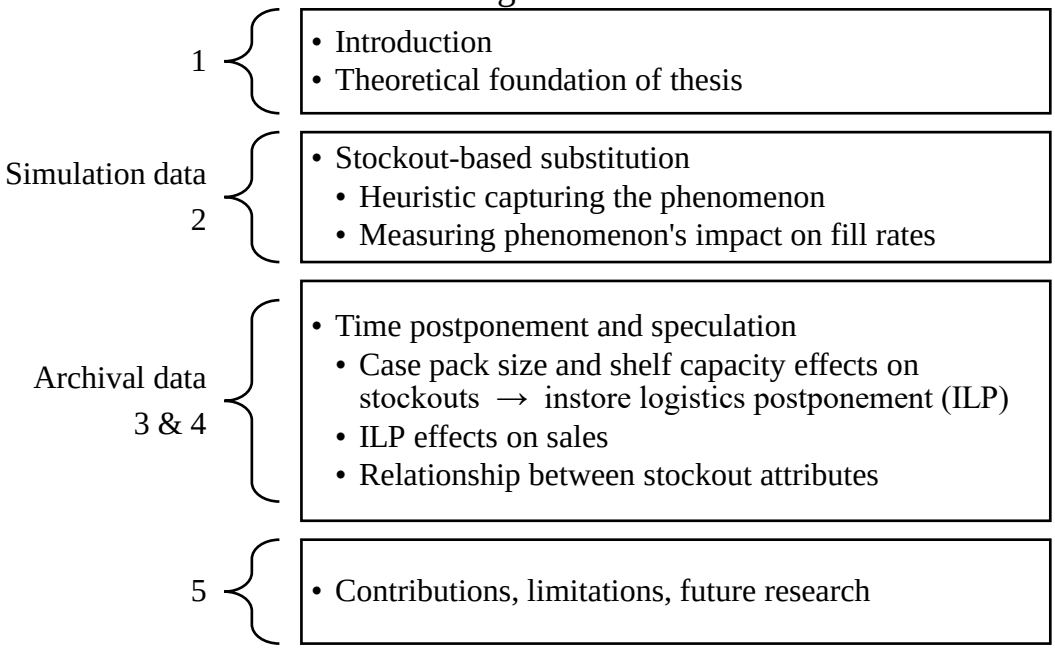


Figure 1-17 Thesis structure

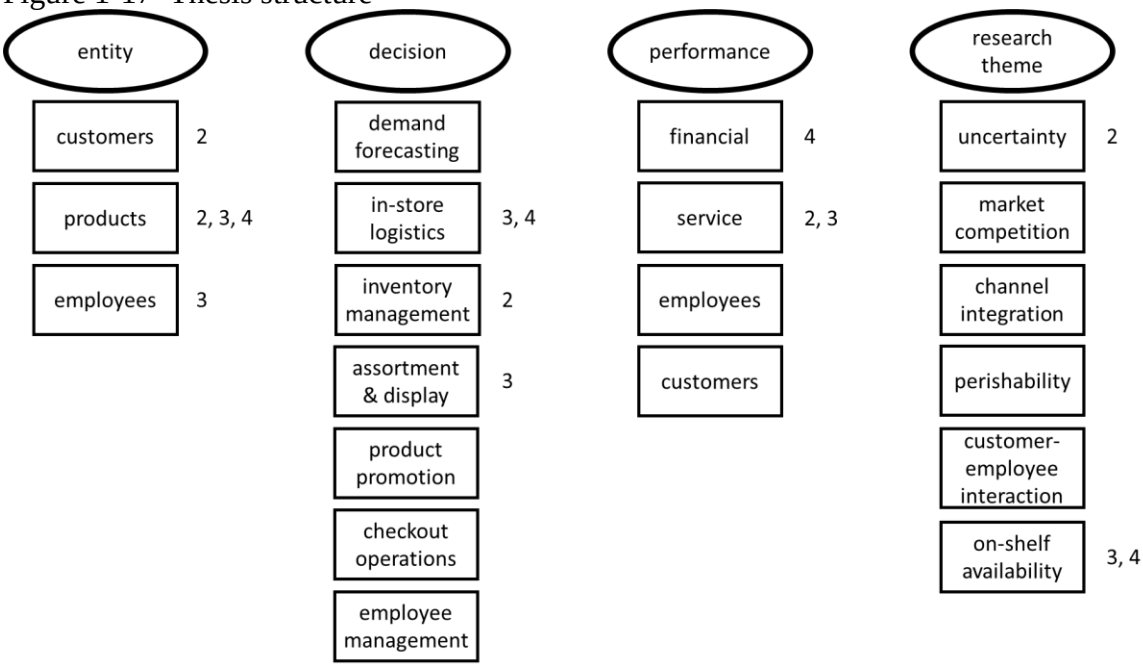


Figure 1-18 Thesis structure under Mou et al. (2018) framework

Figures 1-17 and 1-18 outline the structure of this dissertation on on-shelf availability (OSA). This chapter introduced retail shelf availability and the challenges unique to this supply chain member. The current status of substitution and postponement research as well as streams intersecting this research theme (OSA) were presented. Figure 1-17 shows that secondary (archival) data is used in Chapters 3 and 4, while Chapter 2 is modelled with discrete agent simulation.

According to Mou et al (2018), all research on store operations can be identified within a framework along 4 dimensions: entities, decisions, performance, and research themes. These research themes and components are listed in Figure 1-18. When customers substitute, as in Chapter 2, the entities involved are customers and products. While case pack size is a product attribute, SHO duraton and frequency are attributes involved in shelf replenishment, which means both product and employees (Mou, Robb, DeHoratius, 2018, p. 402) are the entities studied in Chapter 3. While Chapter 4 also uses SHO attributes, it considers them only as antecedents to firm performance instead of as a result of employee efforts.

The retailer decisions studied in this dissertation encompasses 3 of the 7 decisions in the Mou et al. (2018) framework: in-store logistics, inventory management, as well as assortment and display. In-store logistics research itself consists of 3 areas: the use of backroom storage, shelf replenishment, and new in-store logistics operations in omni-channel retailing. Chapters 3 and 4 touch upon both backroom storage and shelf replenishment. Chapter 3 also includes decisions on assortment and display since shelf space allocation "is associated with in-store

logistics" (Mou, Robb, DeHoratius, 2018, p. 407) in terms of the effects of case pack size on replenishment efficiency. Chapter 2 involves the first two of the following five inventory management areas: lost sales, inventory-dependent demand, inventory record inaccuracy, joint inventory and pricing, and ordering behavior.

The retailer's performance is assessed with different measures in this dissertation. In Chapter 2, the focus is on the service perspective in terms of fill rates, or the number of units purchased as a proportion of the number of units demanded. In Chapter 3, the performance measure is with a service perspective as it uses two different ways of measuring OSA, SHO duration and SHO frequency. In Chapter 4, the performance measure, expressed in units sold, can be categorized as a financial measure, since it "primarily relates to bottom-line improvement." Regardless of performance measure, all chapters focus on the immediate or short-term effects of SHO.

The dissertation's theoretical contributions aim to address prior conflicting findings on both supply and demand drivers to SHO and the phenomenon's effects by recognizing risk pooling methods within the retail level as unique to pooling anywhere else in the supply chain. The first portion of the dissertation, where demand that would otherwise be lost is pooled to available items, uses a discrete agent simulation to track dynamic stockout-based demand choices which may be observable to higher level supply chain members facing product unavailability, but not to retailers. It contributes to the understanding of when pooling through substitution improves fill rates and when it does not. The second portion of the

dissertation suggests a different mechanism for pooling through postponement within a retail store as compared to upper stream members. A theoretical framework on this mechanism is presented to explain unexpected empirical outcomes and opens the avenue for further hypothesis development and testing.

The dissertation also contributes to retail practice in terms of inventory deployment and management. When setting service levels of items, to what degree does the substitutability between products matter? What issues or opportunities do the different fill rate measures mask or accentuate? When choosing how much shelf space to dedicate to an item, to what degree does the case pack size of the item matter? What issues or opportunities do the various SHO measures mask or accentuate?

Along with these contributions the final chapter also addresses limitations of each study and of the dissertation overall. From not being able to bring both types of pooling (postponement and substitution) into a single model, to not having item data on entire stores or product categories, a discussion of limitations highlights possible future research and other measures for concepts tested here.

Chapter 2 How does stockout-based switching affect performance fill rates?

This chapter of the dissertation assumes the retailer has full operational efficiency in terms of shelf replenishment including knowledge of the size of initial demand for each product. There is no distinction between shared (backroom) space and dedicated (shelf) space. Everything the store has on-hand is available to the customer, either because there is no backroom or there is instantaneous replenishment of shelves (with no capacity constraint) from the backroom. This may appear unrealistic in terms of store operations—though it is a reasonable portrayal of upper stream member's inventory—but is necessary to remove any confounding effects of inefficient replenishment so that the customer switching phenomenon's effects on the opposing goals of high customer service and low inventory investment is captured. The research contribution is two-fold; (1) reconciling contradictory findings on whether and when substitution improves retailer performance without additional inventory investment, and (2) presenting a more nuanced view of performance fill rate by comparing the impact of customer willingness to switch and target fill rates (inventory investment) on different measures of customer service.

2.1 Introduction

Stores hope that customers who discover their favorite mint toothpaste is stocked out will still make a toothpaste purchase by switching to a comparable available item. An assortment of items within a product category are on store shelves

to attract various customer segments (Rong, et al., 2015). In terms of stockouts, alternate available items may be the same product (mint toothpaste) in a different size or packaging, or an item with a completely different brand or type (different flavored toothpaste, or mint toothpaste with whitening) than the stocked-out item. Customers face stockouts at store shelves (*shelf stockouts* or *SHO*) about 5 to 12 percent of the time (Gruen, Corsten, and Bharadwaj, 2002), and respond to SHO by delaying purchasing, leaving, or switching to another product (Zinn and Liu, 2001). Although customers may substitute one item for another because of price or physical attributes, *customer switching* in this study refers to stockout-based substitution where the customer has a ranked order of preferred products (Honhon & Seshadri, 2010) and automatically switches from one item to another until reaching the highest ranked item that is available on the store shelf. Since customers "base their choice from the products available in stock at the time of their visit to the store" (Honhon, Gaur, & Seshadri, 2010) their purchase depends on instantaneous inventory levels and is called inventory-based dynamic choice.

Some retailers attempt to make use of the customer switching phenomenon as an opportunity to control demand uncertainty of an item (a customer's *initial item*) with already-stocked substitutable items (*alternate items*) without having to increase inventory investment. Risk pooling theory suggests that stores can control each item's risk of lost sales due to stockout by considering all items in the product category as a single source of uncertainty with less variability than observed with each item individually. In this way risk pooling allows stores to capture excess demand for an item (when order cycle demand exceeds that period's on-hand inventory) with excess

inventory of other items (when order cycle inventory exceeds demand during that period). Indeed, McGillivray and Silver (1978) found that customer switching decreases two types of firm costs: stockout costs (because of fewer customer sales lost) and holding costs (because of lower levels of safety stock carried for each item). Risk pooling results in higher fill rates—a greater proportion of arriving customers leaving the store with a purchase—without higher inventory cost or in reaching the same level of performance with lower inventory investment. Making use of the risk-pooling effect of customer switching enable stores to better meet the two opposing goals of high customer service and low inventory investment.

However, instead of improving retail performance, customer switching has sometimes been linked to negative outcomes. Yang and Schrage (2009) built on previous research to coin the term "inventory anomaly," and found both partial substitution, where only a portion of customers are willing to switch to an alternate item when their preferred item is SHO, as well as full substitution increase inventory under certain conditions (such as short lead times and frequent replenishment). Considering the occurrence of a SHO as a network disruption like natural disasters, Wu et al. (2013) found substitution decreases product and retailer market share, whereas other studies find decreasing manufacturer market share and no significant effect on store market share (Rosales, Whipple, & Blackhurst, 2018). Studies on designing products for substitutability show that using product variety to capture market segments is more profitable than customer switching unless items are basically fully substitutable (Rong, Chen, & Shen, 2015). Other studies show that increasing substitutability decreases profits (Hsieh, 2011), and leads to increasingly

worse performance, or a “spiral-down effect” (Cooper, Homem-de-Mello, & Kleywegt, 2006) when the retailer adjusts its model of demand according to data that is constrained to sales. Demand estimates have been shown to be extremely biased with substitutable products, (Anupindi, Dada, & Gupta, 1998) and over-estimate demand for items that aren’t SHO often and are desirable substitutes. The store-observable realized demand is the “effective demand” (Rajaram & Tang, 2001) and is the combination of customers who originally preferred an item (*initial customers*) as well as those who switched to it as a substitute (*alternate customers*). Various attempts to find unconstrained demand—how many customers would buy an item if the item had limitless supply—is reviewed by Guo et al. (2012).

Contradictory findings about the effects of customer switching may be based on two differing views on consumer research which influence research method choice. There are five methods in consumer research: analytical modeling, empirical and statistical modeling, systems dynamics modeling, consumer behavior experiments, and agent-based modeling (Rand and Rush, 2011, p. 183). Rand and Rush (2011) recommend mixed methods where agent-based models (ABM) complement one of the other four methods since it allows for both views of consumer research to be considered in a single study. ABM provides a “fine-grained resolution,” with a bottom-up view of consumer research. It posits that a phenomenon can be parsimoniously explained in terms of its smallest components and may include interactions between those components so that an overall structure emerges from the process itself (Odell, 2002). While consumer experiments have a similar perspective of individual consumer behavior forming “emergent system properties,” ABM is said

to be more realistic and can complement this method in terms of scaling up to populations larger than the experimental group and analyzing interaction effects (Rand & Rust, 2011). Both ABM and consumer experiments have knowledge of initial and alternate customers, either by post-experimental survey or simulation design and thus enable a deeper more granular view of consumer behavior at the individual level.

The other view on consumer research is the basis for the remaining three types of research models. In this top-down view the phenomenon involves a system too complex to reduce into its smallest components, so that overall attributes are studied at the aggregate level. In analytic modelling this may take form as assuming a joint demand distribution for the substitutable products or assuming the switching rate a priori instead of modelling customer behavior to reach final sales outcomes. With empirical research, after assuming a functional form of the customer demand the substitutability between two products may be estimated as correlation coefficients. Based on observation of extant data (effective demand) such methods assume homogenous customers, imposing a single switching rate representative of all customers. In other words, the aggregate attributes of the system capture an average customer's behavior. Last of all, systems dynamics modelling can test the impact of different levels of correlated demand, generating a range of possible sales or SHO outcomes due to different degrees of substitutability or over time. Referred here as only systems dynamics, this umbrella method includes discrete event simulation where the processes under study are already known and Monte Carlo simulation where the distribution is known or assumed for a sensitivity analysis. While all of

these methods increase knowledge of the relationship between substitutability and firm outcomes, they do not provide information about the process: an "explanation of why and how the observed variables correlate in the observed way" (Held et al., 2013, p. 12).

Using ABM as a complement specifically to analytical modeling "add(s) a layer of realism" using "individual-level theories of behavior" (Rand and Rust, 2011) which affect the retailer more so than other supply chain members. For example, analytical models assume substitution requires "the joint probability of shortage of one product and surplus of another product" (Bansal and Moritz, 2015, p. 67). In other words, alternate customers purchase an item only after all of the initial customers have already purchased the item during the inventory cycle. While a reasonable assumption for catalog retailers (Anderson, et al., 2006) or manufacturers (Li et al., 2006) who may suggest customers buy an alternate item in place of a SHO item or who may automatically impose the switch themselves, in the self-service environment of a store the reality is that both initial and alternate customers arrive to the shelf space of an item in random sequence. According to Gilland and Heese (2013, p. 886), "the sequence of customer arrivals affects how product is allocated to customers," since an item may be stocked out earlier than expected due to alternate customers buying the item before arrival of some of the expected initial customers during that time period. Although this random customer arrival sequence may exist with upper supply chain members as well, they are aware of prior period effective demand breakdown and can better prepare for random arrival sequence.

Additionally, compared to inventory further up the supply chain, a store's product category consists of a larger quantity of different suppliers providing substitutable items. Even when a store replenishes its stock from a retail distribution center, it is less likely that it will order and receive every item in a product category at once as analytical models generally assume with joint replenishment (Federgruen & Zheng, 1992). Simultaneously replenishing all of the shelves in a product category would decrease the duration of time that one item is in stock while another is SHO, making customer switching less likely to occur. In reality, not only do customers arrive to store shelves in random sequence but items also arrive (shelf replenishment) at different points in time. Indeed, it is this interaction of demand (customer arrival) and supply (shelf inventory) factors which sets the environment for customer switching. The substitution matrix of analytic models may have switching rates that ignore any interaction effects of random product and customer arrival. Especially considering a retailer's product category management software normally does not incorporate alternate demand in its calculations (Hubner and Kuhn, 2012), the assumptions in analytical models may poorly reflect store practices.

Most customer switching studies use analytic, empirical, or systems dynamics modeling methods. While some use consumer experiments to study customer switching, these studies normally focus on the antecedents of the phenomenon (Zinn and Liu, 2001) with few studies on the impact of customer switching on the firm, beyond finding an increased frequency of SHO occurrence (van Woensel et al., 2007). Some studies using analytic models either provide numerical examples showing when order quantities would increase (Gerchak & Mossman, 1992) with

increased substitution or how much profit could increase when considering customer order arrival and switching (Gilland and Heese, 2013), or which heuristic provides the closest values to the analytical model's optimal results (Rong et al., 2015). Other analytic work can include a Monte Carlo (Yang and Schrage, 2009) or systems dynamics (McGillivray and Silver, 1978) simulation to explore when partial or full substitution leads to increased profits. In such analytic works where substitution may increase or decrease inventory (such as order size in Rajaram and Tang, 2001) or may not necessarily increase profits (Khouja, et al., 1996; Hsieh, 2011) the results depend largely on the relative values of underage and overage costs. Since on-hand inventory restarts at the beginning of every inventory cycle in these newsvendor models (Smith and Agrawal, 2000), the financial penalty of holding too much or too little stock at the end of a period often drives the study's conclusions on whether customer switching helps or hurts the firm. Instead of assumed amounts of underage and overage costs, empirical work on customer switching struggles with measuring when an item is SHO, which is necessary to split effective demand into initial and alternate purchases. Since on-hand inventory is usually unknown, sales patterns help researchers identify when SHO may have taken place (Campo et al., 2008), so that empirical studies mainly focus on the antecedents to customer switching, instead of its impact on sales. Anupindi et al. (1998) are able to capture SHO events, assume a substitutability matrix for the items in their product category, and find that using sales as a way to determine the amount of initial demand leads to biased estimates.

A limited number of studies use only ABM to study customer switching. The most recent (Rosales et al., 2018) explores how much market share a focal retail store

and a manufacturer are able to capture given different distribution channels (direct store delivery or through a retail distribution center) and customer loyalty (high loyalty for the manufacturer's brand versus low loyalty). Their experimental design assumes low loyalty customers overall drop their likelihood of initially preferring the focal item at a greater rate than high loyalty customers, when faced with a SHO. They find "regardless of brand loyalty and/or distribution channel strategy…manufacturers' lost sales are even higher for low brand-loyal products using traditional distribution under repeated OOS [SHO] conditions" (Rosales et al, 2018 p. 152). While customer switching from repeated SHO of focal item hurts focal manufacturer performance, the focal store has no significant change in market share despite its competitor store having been designed to have no SHO occurrence. Wu et al. (2013) also have a similar ABM design but SHO duration of items and initial market share of each item and of each store varies, with customers reacting to SHO in different ways (leave store, switch to alternate item, switch stores, and delay purchase) in initial proportions according to prior empirical research and product type. They find that the proportion of customers willing to switch between items or to delay purchase affect a loss of market share of the SHO item at a greater rate than other customer behaviors, and that the market share loss is greater with longer SHO durations and for stores with larger initial market share.

Studies that combine the ABM framework with other research methods focus on substitution or inventory management with SHOs but do not usually study customer switching. Heppenstall et al. develop an ABM that recreates the competition between gasoline stations by considering customers who substitute because of price

and are able to "replicate spatial pricing trends on both idealized and more complicated real data…to successfully predict whether stations are still viable over a long-term period of ten years" (2013, p. 678). In this case the researchers have compared ABM outcomes with patterns in extant data sets to get an enhanced understanding of the impact of pricing decisions on customers switching gas stations. Signorile (2002) find increased sales, as compared to the economic order quantity model, linked to modelling a store's inventory decisions by looking at the behavior of the following 5 types of people who determine how many consumer units of an item are supplied to stores: forecasting agents, inventory control, replenishment, department, and management agents. Their focus is on inventory control methods given sales will be lost due to stockout and not the customer switching mechanism itself. Regardless, ABM allows researchers to replicate the individual decisions that affect an item's supply and compare it with an analytic (EOQ) model. In other words, ABM is not merely a numerical analysis for testing an analytical model but uses a different theoretical view than the analytic model to approach the same phenomenon or firm decision being studied. According to Rand, ABM "requires rethinking the problem in a computational framework instead of a mathematical framework" (Rand, 2013, p. 383).

This study uses both theoretical frameworks using the methods of ABM along with a customer switching heuristic (FRH) to gain a better understanding of when and why a retailer's attempt at taking advantage of risk pooling may sometimes not pay off. The FRH provides a theory of what proportion of customers will be able to purchase which item, using traditional probabilistic views on the average consumer

and average inventory availability. The ABM recreates the retail environment from the bottom-up, allowing for individual customers with different initial preferences and willingness to substitute to arrive at random sequence. The customers each individually face different levels of inventory on store shelves depending on the purchases of other customers before them and the relative occurrence of shelf replenishment. The framework under the heuristic method views the customer switching environment as so complex that it needs to be measured by mean attributes of expected sales and SHO events which an average customer may face. The ABM framework says the system can be broken down simply into the basic components of different types of customers and shelf availability levels to capture the actual customer switching environment. Comparing the theoretical sales and SHO outcomes from the FRH to the actual outcomes from a simulation of the ABM brings individual-level theory to the discussion on risk pooling through stockout-based substitution.

This study has three main contributions. While the research question centers around how stockout-based substitution impacts fill rates, the use of both research frameworks allows for an understanding of when a heuristic rule of thumb is satisfactory and how and when it does not sufficiently explain the effects of the customer switching process. The study reveals how the practitioner's measures for stocking and controlling the performance of its inventory management can be misaligned and suggests how to better make use of customer switching's risk pooling effect. Last of all, this study builds on the dual framework findings to study the

central question of how the supply and demand drivers of customer switching have different effects on various measures of fill rate.

A literature review in the next section focuses on how dynamic demand lays the foundation for the individual-level theory used in ABM, and on how limited literature is on customer switching's impact on fill rates. The review is followed by sections on the theoretical framework of customer switching, including discussion on how the ABM framework differs from the FRH, and on research design. The statistical analysis and discussion of results is in two parts: in reconciling contradictory findings stemming from the different views on customer switching outcomes, and in studying the impact of the phenomenon on different fill rate measures. Last of all, the conclusion presents managerial implications, study limitations and opportunities for future study.

2.2 Literature review

1.1.1 Dynamic demand in substitution studies

Theory development on customer switching has mirrored theory development about consumer demand in general. General models on consumer demand were static, where consumers could be grouped into different classes which defined their spending behavior for a given environment. In the early 1980s, several studies showed "the continuing lack of accord between the postulates of demand theory and empirical static demand functions estimated on time series data" (Anderson & Blundell, 1984). Consumers were increasing or decreasing expenditure depending on

their spending activity during the prior year. This observed temporal quality of demand opened up avenues of theory on consumers shifting from their prior preferences relative to prior purchases and depending on the point in time that they were making their purchasing decision instead of depending on any supply-based factors. In terms of modeling substitution, dynamic demand leads to the notion that customers may switch from one product to another at a point in time with different rates or for different reasons instead of a single static switching rate (elasticity or substitutability) between two products for all consumers. In other words, there is value in looking at the disaggregate time-based purchase behavior instead of just averaging customer behavior over time.

Dynamic demand and the need to consider customer behavior finds its basis in an analytical work by Smith and Agrawal (2000). In it they posit that customers can exhibit three different types of substitution behavior: random, adjacent, and one-item substitution. They state that the random substitution matrix is analogous to the logit model, because customers have a probability of substituting from one product to the other in the same proportion as each product's original market share. Then they introduce to this static demand model "adjacent substitution," which assumes customers have a ranked preference among products and will only switch to products that are adjacent or one step different in price or quality. This is a dynamic model. Last of all, in one-item substitution customers all prefer the same item as their substitute when their original preference is stocked out. This is an inventory-based dynamic model. By calculating profit levels using these 3 models, the researchers establish that assortment choices in retail inventory management must consider the

pattern of customer behavior before trying to optimize assortment size. This is because the profit obtained at different levels of product variety depend on how the customer behaves when faced with stockout and choice set options. In their study Smith and Agrawal (2000) find a greater increase in profits with substitution for the one-item and the adjacent substitution behavior than the random model. They call for empirical work in retail research that studies substitution behavior in greater detail.

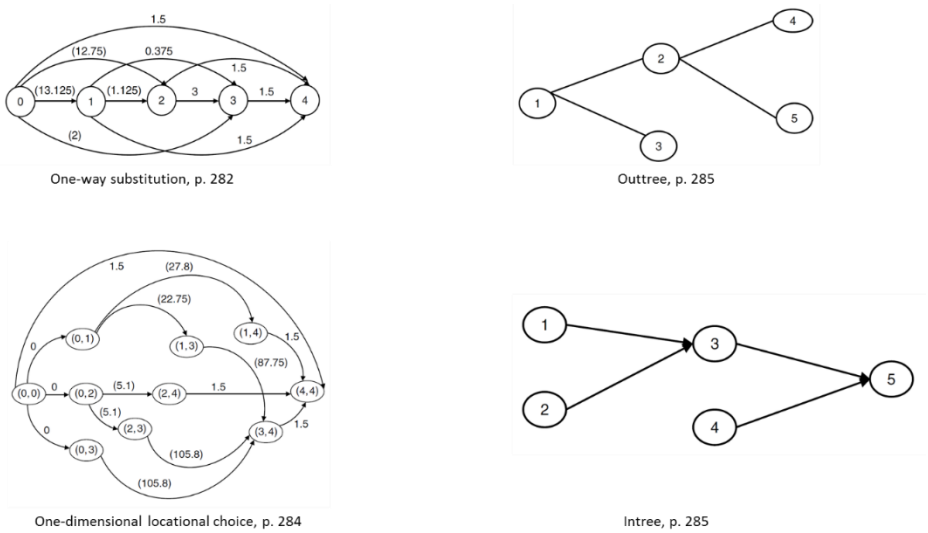


Figure 2-1 Different types of customer behavior (Honhon et al., 2012)

The most recent analytic work building on Smith and Agrawal (2000) is Honhon et al., (2012) which develop methods for efficiently finding optimal assortments in four choice models (Figure 2-1): one-way substitution, where customers switch only (up) down to a (higher) lower quality or priced item; locational choice, where customers switch horizontally to items varying slightly by a single trait; outtree models, where customers all have the same first preference but differ in alternate preferences; and intree models. Intree models are similar to Smith and Agrawal (2000) where customers have different first preferences but the final

substitute item they all prefer is the same, while the locational choice model parallels the adjacent substitution model. The authors also consider costs (holding, goodwill, stockout) and wish to maximize profit, supporting the Smith and Agrawal finding that a retailer's assortment does not have to include the most profitable product if there is a less profitable product which is a preferred substitute by all customers. Honhon et al (2000) contribute to the dynamic demand literature by approximating the optimal solution for the four models listed.

Model	Description	Example (Honhon et al, 2012)
One-way substitution	Substituting between vertically differentiated products	Steel beams of different strengths
One-dimensional locational choice	Substituting between horizontally differentiated products.	A different flavored yogurt.
Outtree	Nested preference of vertically differentiated products	Versions of all-purpose cleaner: antibacterial, heavy-duty, grease-cutting, deodorizing, etc.
Intree	Nested preference of horizontally differentiated products	Pet food for different age and type of pet: indoor vs. outdoor, etc.

Table 2-1 Current analytical demand models (Honhon et al., 2012)

Examples of the four models are provided by Honhon et al (2012) and are listed in Table 2-1. In the intree model, if cat food for a senior cat with hairball and urinal tract issues is stocked out, a customer may switch to a more general type of food until reaching the final choice of cat food for any type of cat. On the other hand,

outtree models describe customer behavior when a general product such as an all-purpose cleaner is stocked out. Then some customers may opt for antibacterial cleaners whereas others may choose a deodorizing version. With all of these models substitution occurs in one direction only and a restriction of customer types is required because there are otherwise too many combinations of customer movement to figure out which products to choose for a product category to have the highest profits. Instead of the best solution, a "good enough" solution is found with approximations (heuristics) and other approaches (Honhon et al., 2012, p. 280). This stream of research following Smith and Agrawal (2000) maintains that considering customer behavior allows for different solutions to assortment problems than traditional analytical models.

Also considering costs and profit, Gilland and Heese (2013) study consumer behavior in terms of the random arrival sequence of heterogeneous customers. They develop an analytical model with numerical examples showing that taking arrival sequence into account when making stocking decisions ("Optimal Policy") results in the greatest financial gain overall. Compared to the optimal policy, stocking products without considering profitability, substitution, or sequence of customer arrivals results in a 4.9% loss in profit. Their study assumes that item purchases by initial customers are always more profitable than purchases by alternate customers, and that there is a shelf space constraint. They state that they are the first to consider the sequence of order arrivals and that simply considering stockout-based substitution in

a static manner with substitution rates generated from past sales outcomes “tends to underestimate actual substitutions” (2013, p. 885).

The implications of inventory-based dynamic demand from the Smith and Agrawal (2000) and Gilland and Heese (2013) studies are that a customer-based view is more than a research methodology, but an opportunity to consider demand characteristics that are ignored or lost in traditional static views. Indeed, the concept of random arrival sequence as a demand driver has already improved models on warehouse-order picking (Diaz, 2016) and certain types of shrinkage (Zhou & Piramuthu, 2015). The dynamic demand paradigm takes a step back from “existing consumer choice models, such as the locational choice model or multinomial logit model,” (Honhon et al., p.280). It does not require a substitution matrix or the assumption of independence from irrelevant alternatives (IIA). IIA assumes customers facing a SHO will switch to alternate items in numbers proportional to each alternate item’s overall market share. Nested models which avoid IIA may still provide “unreasonable” results according to Gowrisankaran and Rysman (2012) who find elasticities (price and product attributes) in a dynamic demand model varies greatly from traditional static models. They conclude the biased elasticities in static models result from considering an average individual customer instead of a set of customers whose composition changes over time.

A changing composition of customers over time requires dynamic instead of static customer demand modelling. While the Gowrisankaran and Rysman (2012) study centers around substitution types outside of stockout-based substitution, it has implications for the customer-switching environment because of splitting and

combining demand streams for items. The composition of customers demanding an available item certainly changes during the inventory cycle as soon as alternate items SHO and their customers switch to the available item. Items with overlapping lead time intervals (where on-hand inventory is low but there is a replenishment order on its way) may have overlapping demand streams. Especially if the lead time periods do not overlap, then the demand for the SHO item may switch over to an item that is not in its lead time period. Its inventory would now have to serve both initial and alternate customers and the additional demand may result in the item SHO earlier than expected, serving fewer initial customers than without switching. In periodic review inventory systems especially, the item could stay SHO longer than planned. The switching that resulted from a SHO would be driving a SHO of an alternate item. With changing composition of customers over time, the fill rate can become complicated to measure.

1.1.2 Fill rate as a performance measure

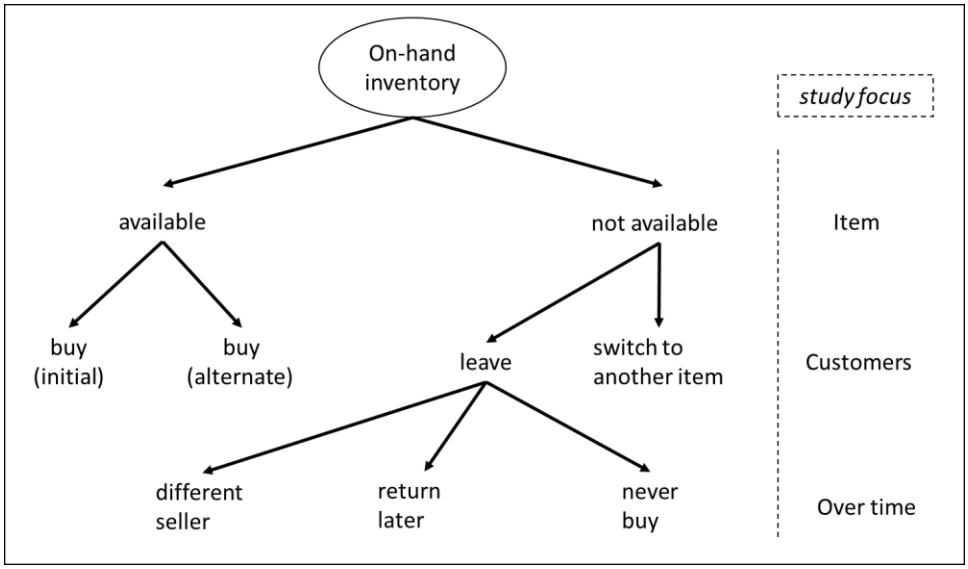


Figure 2-1 Fill rate generation schematic

While fill rates typically measure how well on-hand inventory (supply) meets cycle demand its operationalization depends on the study's focus, as in Figure 2-1. When the focus is on an item, the topic of the study involves *product availability* when there is customer demand. An item is either available or not when the customer wants it, and it may not be available because it is completely stocked out, is still being processed, or there aren't enough units of it to complete the customer's order. When the focus is on customers, the topic of study involves *customer service* (service quality) as the fill rate often becomes a proportion of customer orders or quantity of customers served instead of units available for purchase. When the focus of studies is customer reaction to unavailability over a period of time, the unsuccessful initial transaction (visit to a store) is a precondition to a second phase of behavior at a later time. During this second phase the customer may later return to the same seller, switch to another seller or decide there is no need for the item (Zinn and Liu, 2001).

Item-centered studies generally revolve around supply-based decisions and *operational efficiency*. As such the research questions are generally based in inventory-focused decisions affecting product availability (supply) and may have nothing to do with substitutability. One of the earliest papers using the term "fill rate" mentions the measure as one of four Air Force controls for their transportation function (distribution activities) where it is "defined as the number of items shipped divided by the number of items requisitioned. The actual rate is compared with an 85 per cent goal, and any poor performance is analyzed to determine the cause" (Hauk, 1964, p. 16). Referring to Figure 2-2, the fill rate is generated by taking the

“available” units as the numerator of the rate and the sum of “available” and “unavailable” as the denominator.

Supply-based decisions in item-centered studies include but are not limited to: how often to control inventory levels (Sezen, 2006), production capacity of products of varying complexity (Closs, Nyaga, & Voss, 2010), how many different items to carry and distribute (Wan et al., 2012), whether to have a centralized inventory or a decentralized one with emergency transshipments (Evers, 1997), and the appropriate shipping quantity (case pack size) of items to stores (Waller, et al., 2008). These and other decisions impact the level of inventory investment since the decision affects how many units are held where. A review by Moussaoui et al. (2016) provides an overview of drivers of product availability. Their review centers around product unavailability instead, which is simply looking at product availability in terms of either SHO lost sales units or lost sales rates (Gruen and Corsten, 2008, p. 3). The lost sales rates take the “not available” units as the numerator instead of the “available” value in Figure 2-2. Such studies measure the impact of SHO occurrence instead of customer switching.

Customer-centered studies operationalize service quality (*customer service*) using fill rates. The focus is on the customer experience or response and as such the research questions generally revolve around marketing or customer information-based decisions to alter demand behavior. Customer behavior can include substitution even without stockout occurrence. Demand-based decisions in such studies include: whether or not to move consumer demand information up the supply chain (Boone et

al., 2002), how to estimate order timing (Ali and Pinar, 2016), and what type of after-sales support to offer (Cohen, Agrawal, & Agrawal, 2006). Instead of adjusting supply to adequately meet demand, such studies focus on how to make better use of existing supply by reshaping demand with discounts or promotions (Laroche et al., 2003) or using demand information to affect supply planning higher up in the supply chain among the retailer's vendors. Outside of the billboard and scarcity effects (Xue et al., 2017), where customer preference or demand changes for items with a greater or lower on-hand inventory, reshaping demand involves inventory management decisions other than inventory investment (how much inventory to have on hand at a specific location).

Studies considering product availability and customer behavior together generally do not use fill rate measures as outcomes although specific types have been proposed. Also known as "customer-based availability measures", Zinn et al. (2002) propose capturing the proportion of customers able to buy either initial or alternate items. In other words, assuming that the fill rate generation tree in Figure 2-2 is for a single focal item, customer-based availability measures would take multiple such trees for multiple items. The numerator of the fill rate would be generated with the "buy (initial)" numbers of the focal item as well as the "buy (alternate)" numbers of any other fill rate generation trees of alternate items. While the sum of these purchases would constitute total customers served from available inventory, the number of customers who face product unavailability could be double counted, as both the initial and alternate item could be SHO. Double counting also includes

customers who face an initial item SHO being counted as not being served even if they buy an alternate item.

Customer-based availability measures still focus on supply-based decisions although the fill rate is generated using demand (customer) movement instead of the item. Zinn et al. (2002) refer to a "trigger-based system" of inventory control, where the reorder point (the trigger) for an item may increase when considering a specific (key) customer subset's demand during lead time characteristics (mean demand, standard deviation, and safety factor for that specific customer). This trigger point ensures that the item has enough units on hand during that inventory period for this key subset of demand. "When the trigger is reached, managers have the option of (1) halting sales to non-key customers until the inventory is replenished, (2) extending the delivery time to non-key customers, (3) negotiating with the key customer the amount actually needed until the inventory is replenished, or (4) arranging for an emergency transshipment" (Zinn et al., 2002, p. 33). Options (1), (2) and (4) are all supply-based decisions, while option (3) refers to reshaping demand of key customers. In the retail environment it is possible to consider initial customers of an item as "key customers," since ensuring their purchase of their initially preferred item avoids loss of goodwill or customer loyalty.

Customer-based availability measures are difficult to use both in retail practice and research due to the unique role of the retail store in the supply chain. In the self-service environment of the store, how will the manager be able to stop sales to non-key customers (suggestion (1) in Zinn et al., 2002) when customer initial

preference is unobservable? The store's unique service of instantly providing goods on purchase through customers transporting their purchases themselves means that extending delivery time (2) would require drop-shipping or other transportation costs when the retailer would normally have none. Unlike demand higher up in the supply chain, customers at a store often only purchase one unit of an item, leaving little negotiating room (3) on adjusting order size. Unlike apparel, consumer packaged goods are largely substitutable and the retailer's unit profit is usually prohibitively too low for emergency transshipments to be a reasonable option for the store manager. In terms of research, scholars may be able to approximate initial customer demand (Vulcano, Van Ryzin, & Ratliff, 2012) for the inventory period or during lead time using sales data, but that would mean their research would have to focus on the supply-based drivers of availability since sales is being used for estimation purposes. Fill rates in this case would still involve research questions on operational decisions and efficiency (supply), although they may measure the impact of these decisions on customers.

Studies considering product availability and customer behavior in terms of multiple store visits are illustrated as "over time" in Figure 2-2, and do not normally use fill rates as outcome measures. The focus of such studies is usually retaining customer demand and the long-term consequences of product unavailability. Customers stop going to stores where they face repeated SHO (Turk, 2012), with such stores losing market share (Wu, et al., 2013), as well as losing sales of items that are available but would have been purchased in the same "basket" (store visit) as the SHO item (Anderston et al., 2006). Indeed, all retailers carrying an item that is SHO

at only one store can potentially lose sales of that item as 9% of customers facing SHO simply decide to never purchase that item (Gruen and Corsten, 2008, p. 10). Since this "over time" branch in Figure 2-2 falls under product unavailability, the studies in this sphere do not consider what proportion of an item's sales were by initial or alternate customers. Customer switching is merely one among several possible outcomes of SHO, instead of being the focal phenomenon of study.

Finally, target fill rates are sometimes considered in substitution studies that use various performance measures. A *target fill rate* refers to the proportion of customers arriving to the store who management intends to serve, acknowledging that the remaining customers will leave without purchasing the item. In the Hauk (1964) example it was stated as a goal of 85%. The target fill rate is an interchangeable term with *target service level* (TSL) and is the practitioner's "subjective management judgment" (Tersine, 1994, p. 211) recognizing it is not possible to hold a limitless supply of an item, and the store's space constraints make holding enough customer units needed for expected inventory cycle demand too costly. This TSL is a measure of inventory investment, or how much of an item to have on hand. Higher TSLs involve higher inventory investment as the store will either hold more units on its shelves or in the store (higher safety stock) or replenish more frequently (higher reorder point or shorter cycle periods), or both.

In many studies on substitution the TSL is a constraint (Tan and Karabati, 2013) and the outcome studied is profit instead of fill rate (Shin et al., 2015). The TSL constrains the model—which incorporates holding costs—to a minimum level of

required inventory investment so that the service level is achieved. In their review of customer switching with two-way substitutability literature, Shin et al. find that scholars look to either maximize expected profit, expected payoff, or likelihood (2015, pp. 689-690). Payoff takes the profit of each consumer unit of an item and multiplies it by the number of units purchased of that item. This is a useful way of incorporating price differences between substitutable items. Likelihood in this study refers to "the proportion of brand n_i's demand that spills-over to brand n_k when n_i is unavailable" (Anupindi et al., 1998, p. 416). It is the only paper found in the Shin et al. (2015) taxonomy which uses likelihood (substitutability rates, in this case) as outcomes. It does not consider target fill rates at all and focuses on how biased demand estimates can be when using realized sales to estimate the "core demand" (the size of initial customers) of an item. Shin et al. (2015) cite one study that uses TSL and a service rate outcome but it is for a study on price- and assortment-based switching and about centralized inventory decisions.

Outside of such rare instances of using likelihood as a performance outcome, studies on management of substitutable items usually measure the impact of substitution with some financial measure (revenue, cost, profit) (Shin et al., 2015). Indeed, in their substitution taxonomy, Shin et al. (2015) show that regardless of modelling objective (assortment planning, capacity planning, inventory decision, price decision) the formulation of the substitution model uses financial outcomes. That is because the models consider balancing having too many units (overage costs) with too few units (underage costs) while still being able to capture alternate demand to benefit from risk pooling. In practice, however, "the most common situation is

when an organization does not know its stockout (underage) costs or feels very uneasy about estimating them," (Tersine, 1988, p. 211), so that a TSL is used. The TSL is used to stock an item based on the expected proportion of arriving demand it aims to serve. The TSL is then compared to a fill rate calculated with the effective demand (units sold) divided by the arriving demand. The resulting rate is the likelihood that an item will be available when a customer demands it. But which customers? Both initial and alternate customers? In practice, the TSL for an item generally does not consider substitution (Hubner and Kuhn, 2012).

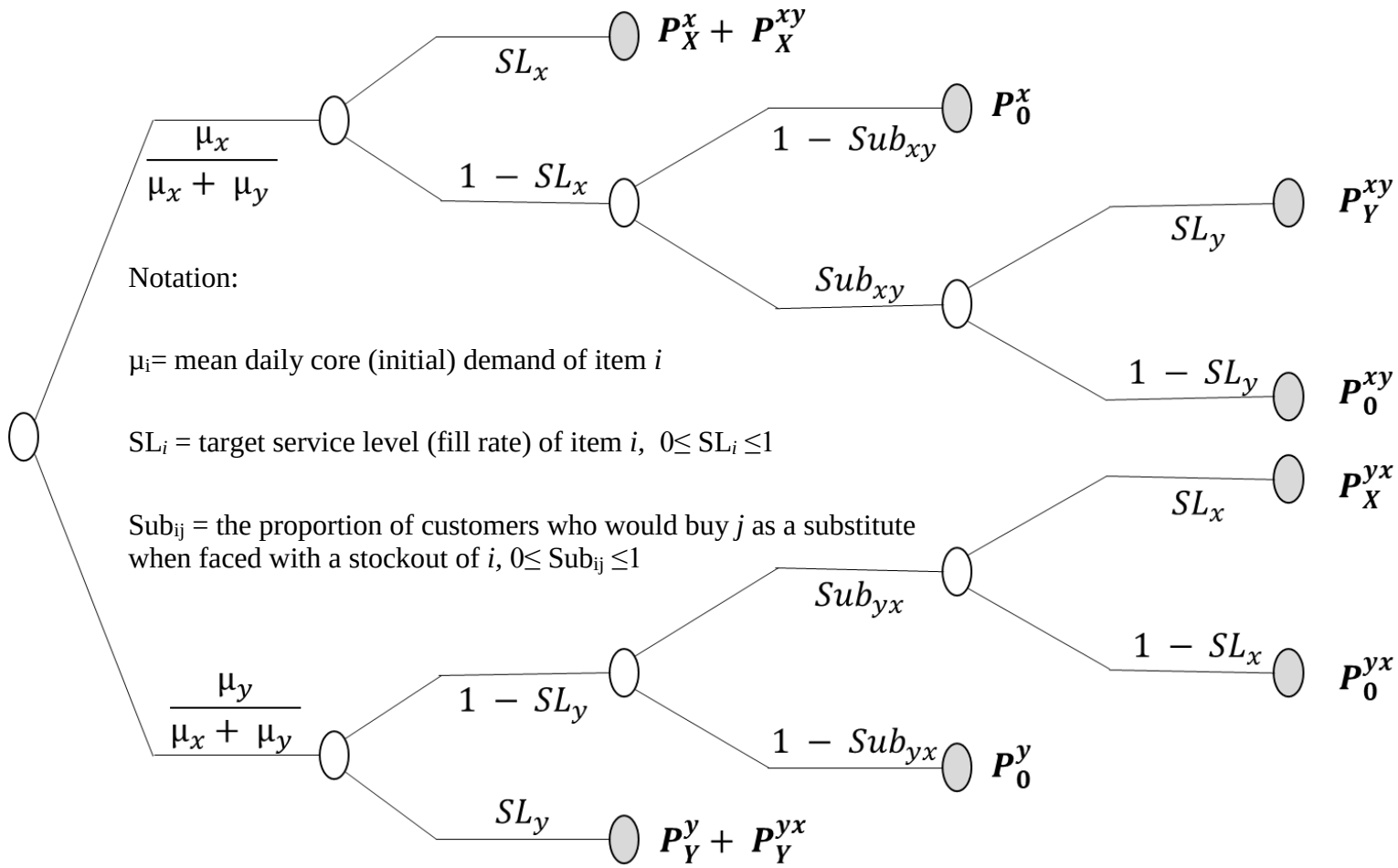
There is one study in the literature on the impact of customer switching on fill rates. Kale et al. (2017) simulate the lead time period demand of two substitutable items with "desired fill rates" (TSLs) to study customer switching's effect on unit fill rates for each item's safety stock. While demand, like lead time, is stochastic and individual orders can be for 1, 2, or 3 units each, it is not dynamic. In other words, the number of units available at the start of each day first serves initial customers which arrive that day. If the initially preferred item is SHO, then a fraction of customers will switch to the alternate item at the substitution rate the researchers impose (by simulation inputs) on the system. It is not clear how the TSLs are used in the simulation, but t-tests of the results show that the achieved unit fill rate is at least as high as the TSL for each item and that the degree of improvement depends on the TSL. Low customer service level targets benefit more from increasing substitution than higher TSLs. Their simulation also shows that increased alternate sales are not linked to decreased initial sales of the same item. In the worst case, a scenario with substitution is not significantly different than a scenario without substitution, so that

customer switching never results in a unit fill rate that is lower than TSLs. All of these findings make sense with the assumption of initial customers being able to buy the item before any alternate customers attempting to do so.

This study brings together TSLs and multiple types of performance fill rate (item, customer, initial, category fill rates) by isolating the customer switching phenomenon in a retail environment of two products and random arrival sequence of heterogeneous customers. The main assumption is that the retailer has a TSL for each item which only focuses on the item's availability (unit fill rate) instead of its substitutability. This not only reflects practice but is often assumed in research as well (Gurler & Yilmaz, 2010). The substitutability is accounted for in the theoretical framework (FRH) by conditional probabilities and the actual environment is created by a simulation program (an agent-based model simulation or ABMS) where customers with different initial preferences arrive in random sequence to store shelves, may face SHO and may be willing to switch to the alternate item. This study avoids demand reshape from price differences, promotions, or inventory billboard effect, and only considers the customer's behavior in purchasing an item or after facing SHO. There is no consideration of how the same individual customer would respond to SHOs over time, so this study incorporates only the "item" and "customer" tiers of Figure 2-2, and not "over time". Differing from prior research, fill rate measures of both item availability and of customer service are in this study of customer switching.

2.3 Theoretical framework of customer switching

Consider two different items, called "X" and "Y," which are in the same product category but could be sized or packaged differently or from different brands or suppliers. Item "X" is the focal item, but not all customers arriving to the store shelf initially want to buy item X. Initial customers of item X are focal customers while alternate customers are the initial customers for item Y. Initial customer size for each item, or core demand, is noted in Figure 2-3 as $\mu_{i,}$ and is the average daily units demanded for each item. This is a value that is known to the retailer and used to determine order size for an inventory period. Whatever the reason (limited shelf space, insufficient funds to invest in more inventory, insufficient vendor capacity, etc.) the retailer plans to serve only a certain proportion of μ_i initial customers, with SL_i as the item's TSL. In terms of customer switching this theoretical framework assumes only a portion (Sub_{ij}) of customers facing SHO is willing to switch to their alternate item. This is different than viewing Sub_{ij} as the proportion of customers arriving to the store (μ_i) and are willing to switch if faced with initial item SHO. There is no guarantee that every customer who faces SHO will be able to purchase their alternate item because that item too has only a certain level of inventory investment as determined by its own TSL. In this way, there are 8 customer outcomes for 2 items.



Notation:

μ_i= mean daily core (initial) demand of item *i*

SL_i = target service level (fill rate) of item *i*, $0 \leq SL_i \leq 1$

Sub_{ij} = the proportion of customers who would buy *j* as a substitute when faced with a stockout of *i*, $0 \leq Sub_{ij} \leq 1$

Such that: $P_X^x + P_0^x + P_0^{xy} + P_X^{xy} + P_Y^{xy} + P_Y^y + P_0^y + P_0^{yx} + P_X^{yx} + P_Y^{yx} = 1$

Figure 2-2 Customer switching outcomes

2.3.1 Customer switching outcomes

The tree chart in Figure 2-3 shows the 8 possible customer switching outcomes as darkened ovals, marked with specific notation containing sub- and superscripts. The subscript refers to which item (X or Y) the customer has purchased, or 0 if the customer leaves without purchase. The superscript with lowercase letters refers to the customer type. If there is only an "x" or a "y" the customer is not willing to substitute if the initially preferred item is SHO. In contrast, a superscript of "xy" refers to customers who initially prefer item X but would switch to item Y if faced with SHO of X. Similarly, "yx" are initial customers of item Y whose alternate item is X. Customers leave without purchase in 4 of the 8 possible outcomes: P_0^x, P_0^{xy}, P_0^{yx}, and P_0^y. The remaining 4 outcomes all involve customer purchases. In two of the purchase outcomes, the item is an initial purchase ($P_X^x + P_X^{xy}$ and $P_Y^y + P_Y^{yx}$) and in the other two purchase outcomes the items are alternate purchases(P_Y^{xy} and P_X^{yx}). While the FRH doesn't differentiate between customers buying their initial item by whether they are willing to switch or not, it is included in the outcome to better link it to the ABM framework outcomes presented later.

Each customer outcome is a likelihood, or a rate which takes the denominator of the total size of initial item demand, μ_i, (or total demand, $\sum \mu_i$ for all i, if considering the product category overall) and various numerators depending on outcome. For example, outcome P_0^x is the fraction of μ_x customers who leave the

store without purchase because their initial item (X) is SHO and they are not willing to switch to the alternate item (Y). The core (initial) demand for each item can be expressed in terms of outcomes as the sum of those who are able to buy an item (initial or alternate item) and those who leave without purchase. For example, if the demand for item X is half of the overall demand for both items: $\frac{\mu_x}{\mu_x + \mu_y} = 0.5 = P_0^{xy} + P_X^{xy} + P_0^x + P_X^x + P_Y^{xy}$. The fraction of total demand served from existing supply of both items is the sum of purchase outcomes: $P_X^{xy} + P_X^x + P_Y^{xy} + P_X^{yx} + P_Y^{yx} + P_Y^y$. The numerator of each outcome is a measure of supply because customers are able to buy the item if it is in stock. The denominator of each outcome is a measure of the core demand (initial customers) of both items. The sum of all *P* likelihoods is 1 (or 100%) so that no other outcomes are possible for these two items. If considering only one item at a time then: $\mu_i = 1 = P_0^{ij} + P_I^{ij} + P_0^i + P_I^i + P_J^{ij}$, where *i* is the initially preferred item, *j* is the alternate item, *I* is purchase of the focal item and *J* is purchase of the alternate item.

The outcomes are expressed in proportions instead of absolute (unit or dollar) counts sold or lost. There is a total count of customers arriving to the store, and each customer demands only one consumer unit of an item, so there is a total number of units that need to be supplied by the retailer. However, the retailer does not maintain that level of inventory investment so some customers are going to leave without purchase. The outcomes are tracked with rates instead of counts because of the overarching relationship of the outcomes ($\sum P$=1) mentioned earlier. As the count for

one of the 8 outcomes increases, there must be an equivalent decrease at one or more of the other counts. Having all of the counts at each outcome divided by a total number of units demanded expresses each outcome as a fraction of a whole that a simple count value cannot.

Although each outcome is expressed as a stockout or fill rate, these switching outcomes are not the same stockout or fill rates used in practice or in research. For example, the outcome P_0^i is the proportion of customers who are not willing to substitute and face SHO while P_0^{ij} are the proportion of customers who face SHO of both initial and alternate items. The current availability and customer switching literature does not go to this level of granularity because of issues retailers face in capturing SHO occurrence or distinguishing between initial and alternate customers of an item. This is why the second part of this study uses these outcomes for various fill rate measures which have been proposed or are used in research. The first part of this study looks at the outcomes themselves to see if there is a difference between the FRH and ABM frameworks of customer switching.

2.3.2 The FRH and ABM frameworks

In the theoretical framework or the fill rate heuristic (FRH) the 8 mutually exclusive outcomes in Figure 2-3 can be estimated in one of two ways. In terms of whether customers will be able to buy their initial item, this is simply the store's

target service level for the item. For those who face SHO instead, the outcome is expressed as a conditional probability. Since there is a probability that an item will not be available for a certain proportion of initial customers and a proportion of those facing SHO are willing to substitute, the likelihood of buying the substitute would be: $P_J^{ij} = (1 - SL_i)(Sub_{ij})(SL_j)$ where the last term is included because there is a chance the alternate item will not be available either.

If considering only one customer demand stream (following only one of the first branches in Figure 2-3) there are only 4 outcomes to consider instead of 8: buy the initially preferred item, buy the alternate item, or leave after facing one or two SHO. In that case the four outcomes equal 1 where: $\frac{\mu_i}{\mu_i} = 1 = P_0^{ij} + P_I^{ij} + P_0^i + P_I^i + P_J^{ij}$. Note that considering one customer demand stream includes availability outcomes for both products, as Zinn and Liu (2001) propose with their customer-based availability rates. In contrast, all of the outcomes for item X (instead of customer X) are: 1 = $P_0^{xy} + P_X^{xy} + P_X^x + P_0^{yx} + P_X^{yx}$. This considers all the instances that item X is demanded, either by initial or alternate customers, and the proportion of the time X is available to meet that demand. Retailers generally set TSLs according to these item (unit) fill rates (Hubner and Kuhn, 2012) instead of the customer-based availability rates.

Outcome	FRH expression of outcome
$P_X^{xy} + P_X^x$	SL_x
P_0^x	$(1\text{-}SL_x)(1\text{-}Sub_{xy})$
P_Y^{xy}	$(1\text{-}SL_x)(Sub_{xy})(SL_y)$
P_0^{xy}	$(1\text{-}SL_x)(Sub_{xy})(1\text{-}SL_y)$
$P_Y^{yx} + P_Y^y$	SL_y
P_0^y	$(1\text{-}SL_y)(1\text{-}Sub_{yx})$
P_X^{yx}	$(1\text{-}SL_y)(Sub_{yx})(SL_x)$
P_0^{yx}	$(1\text{-}SL_y)(Sub_{yx})(1\text{-}SL_x)$

Table 2-2 Customer switching outcomes according to FRH

Following the branches of the tree to each outcome provides the likelihood of that outcome, as listed in terms of items X and Y in Table 2-2. In the FRH, each item is available at level SL_i, and when it is unavailable, only Sub_{ij} proportion of customers facing SHO will switch to the other item. If customers switch, then the demand for the alternate item is greater than the store's on-hand inventory stocked according to its initial-demand-based TSL. In other words, expected outcome P_J^{ij} calculated in this manner may overestimate the proportion of customers actually able to buy item *j* as a substitute. Zinn et al. (2012) note the reorder point for the substitutable item would have to be higher to accommodate the additional demand from customer switching and still maintain the service level originally intended when investing in that item. Indeed, the proportion of initial demand served ($P_X^x + P_X^{xy}$, for item X) of an item may also be less than its expected service level, as random arrival sequence (Gilland and Heese, 2013) may result in alternate customers purchases proceeding initial customers for whom the inventory is held. The same rate of availability (SL_i) that is

planned for initial customers is also imposed upon alternate customers over the entire inventory period. The first part of this study investigates to what degree this assumption in the theoretical framework of the FRH is in accord with actual outcomes computed by simulation under the ABM framework.

In contrast, the ABM framework avoids imposing a priori values on the actual proportion of customers who have switched. While the total demand ($\mu_x + \mu_y$) is the same for both views of customer switching, the process of obtaining customer outcomes differs between views. In ABM, the customer switching phenomenon is reduced to demand and supply components or drivers. The demand is made up of two customers types for each item, those who are willing to switch and those who are not. Some of the demand characteristics (mean and standard deviations of the inventory period and the mean lead time period) are incorporated into calculations to determine supply. In terms of supply, the TSL is not used to directly calculate conditional probabilities of who will buy the item (as in FRH) but is instead used to determine reorder points of each item, as listed in Table 2-3. Items with higher TSLs ("SL" in the table) have greater inventory investment. Their reorder point, B, is higher because there is greater safety stock, S, to hold in order meet a greater proportion of incoming demand. Items with higher TSL keep more units on hand, which is why the TSL is a measure of inventory investment. Other two-way customer-based substitution studies focus on other measures of inventory investment including: base-stock levels (Chiang, 2010) (Nagarajan & Rajagopalan, 2008), order quantity (Khouja et al., 1996; Huang et al, 2011; Stravulaki, 2011), centralized versus competitive inventory locations (Netessine & Rudi, 2003), and order-up-to levels (Tan and Karabati, 2013).

Notation	Description	Variables used in calculations
D	Daily demand	Normally distributed with mean, μ_D, standard deviation, σ_D
Q	Order quantity when on-hand inventory hits reorder point	Inventory period in days * μ_D
L	Lead time	Constant, in days
M_i	Average demand during lead time	μ_D*L
σ_{Li}	Standard deviation of lead time demand	$\sigma_D*\sqrt{L}$
Notation	**Description**	**Calculations**
E(M>B)	Expected lead time stockout in units	$\frac{(1-SL)*Q}{SL*\sigma_L}$
E(Z)	Partial expectation function, standardized stockout for standard normal distribution	NORMDIST(Z, 0, 1, FALSE) - Z*(1-NORMDIST(Z, 0, 1, TRUE)) in Excel
Z	Standard normal deviate	Use Excel's Goal Seek function to find Z where E(M>B) = E(Z)
S	Safety stock	Z*σ_L
B	Reorder point in units	S+M

Table 2-3 Supply calculations for ABM

The customer switching outcomes in the ABM depend on looking at customer behavior instead of effective demand outcomes. The level of detail in the notation in the model in Figure 2-3 incorporates the 5 ABM outcomes within the same figure of the FRH framework. The first customer type is one who is not willing to substitute and will leave if faced with SHO.

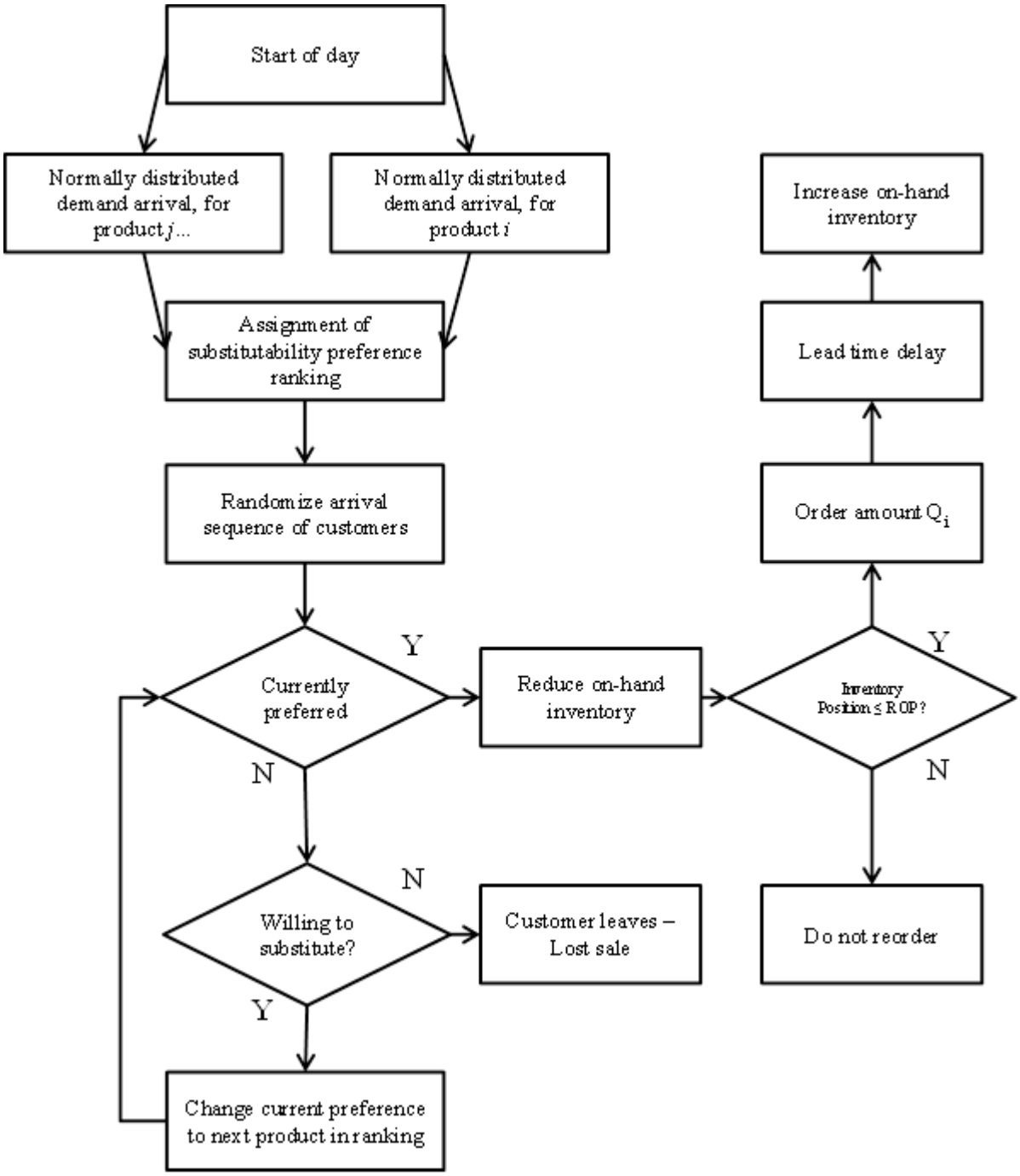


Figure 2-3 Mapping the ABM process

The two possible outcomes for this type of customer is buying the initially preferred item (P_I^i) or leaving without purchase (P_0^i). The second customer type is willing to switch and are Sub_{ij} fraction of the core demand, μ_i. The 3 possible outcomes for this customer type is: buying the initially preferred item (P_I^{ij}), buying the alternate item (P_J^{ij}), or leaving the product category without a purchase (P_0^{ij}). Unlike the FRH framework's conditional probabilities, ABM is a computational method where customers fall into one outcome over another over based on their interaction with one another and the on-hand inventory. In other words, each individual agent is allocated into one of either two or three outcomes depending on customer type (willing or not willing to substitute), and the current inventory levels upon arrival to the store shelf. The on-hand shelf inventory level is a running total of replenishments to the shelf (adding units) and prior customer picking from the shelf (subtracting units). In the ABM framework, the components—customer types and reorder points—of the customer switching phenomenon are defined and allowed to interact, instead of making assumptions about the phenomenon's outcomes as in the FRH framework.

The final breakdown or proportions of each outcome in the ABM builds up organically from the process mapped out in Figure 2-4. In this process a store day begins with a certain level of inventory on hand. Customers, who have a set of preference rankings before arriving to the store, arrive in random order sequence. They check the store shelf to see if their initially preferred item is available and buy it if it is. If SHO, they switch if they are willing to and buy the alternate item or leave without purchase if the alternate item is also SHO. The customer not willing to switch

and facing SHO of the initial item leaves without purchase. As soon as a store visit is complete the inventory position is checked to see if it has dropped to the reorder point. If so, an order is placed of the predetermined order quantity for the inventory period. That order arrives to store shelves after a certain lead time. The customer switching outcomes are portrayed in Figure 2-4 as the rectangles "customer leaves – lost sale" and "reduce on-hand inventory." The process is computational in that there is a tally at each rectangle accumulating how many customers of each type fall into which outcome. In this way, the ABM compiles the customer switching outcomes unit by unit (or customer by customer) instead of assuming overall proportions for an average customer as in the FRH.

Certain relationships are expected to take place in the FRH framework, between the outcomes and the supply (TSL) and demand (willingness to switch) factors. First, increasing inventory investment of an item (increased SL_i) by definition should be linked to an increased proportion of customers (initial and alternate) being able to buy the item as compared to lower TSLs. A positive relationship between inventory investment and unit fill rates has previously been shown (Closs et al., 2010). Secondly, an item's alternate demand size increases as the willingness for alternate customers to switch increases (increased Sub_{ij}). Of the customers facing a SHO of an item, a greater proportion would switch to the alternate item as compared to a lower willingness to substitute. Risk pooling should have a greater effect on performance with increased substitutability.

Units of demand or of sales	Impact of increasing focal item TSL	Impact of increasing alternate item TSL	Impact of increasing focal customer willingness to switch	Impact of increasing alternate customer willingness to switch
Initial sales of focal item	1st ↑, more units of supply to meet demand	4th ↑, fewer units of supply taken away by alternate customers	6th ↓, more alternate sales	3rd ↓, more alternate sales
Alternate sales of focal item	1st ↑, more units of supply to meet demand	3rd ↓, alternate demand size smaller	5th ↑, alternate demand size larger	2nd ↑, alternate demand size larger
Alternate demand size of alternate item	2nd ↓, supply meets greater proportion of initial demand	5th ↓, supply meets greater proportion of initial demand	1st ↑, more (greater proportion) of initial customers facing SHO willing to switch	4th ↑, more initial customers of focal item face SHO
Initial sales of alternate item	4th ↑, fewer units of supply taken away by alternate customers	1st ↑, more units of supply to meet demand	3rd ↓, more alternate sales	6th ↓, more alternate sales
Alternate sales of alternate item	3rd ↓, alternate demand size smaller	1st ↑, more units of supply to meet demand	2nd ↑, alternate demand size larger	5th ↑, alternate demand size larger
Alternate demand size of focal item	5th ↓, supply meets greater proportion of initial demand	2nd ↓, supply meets greater proportion of initial demand	4th ↑, more initial customers of alternate item face SHO	1st ↑, more (greater proportion) initial customers facing SHO willing to switch

Table 2-4 The interconnectedness of customer switching outcomes

On the other hand, every customer switching outcome involving sales may not increase with increasing TSL or willingness to switch, because of the complex interconnectedness or interaction of the outcomes, as illustrated in Table 2-4. Every row of this table lists the size of demand or of sales outcomes that are affected by TSL or willingness to switch, which are listed as table columns. The initial demand size for each item is not included in the table. It is exogenous and is not affected by changing the inventory investment levels or the changing proportion of initial customers willing to switch. If 100 customers initially want item X then the difference between none of them willing to switch to item Y and all of them willing to switch to item Y does not change the fact that there are 100 customers who initially want item X. The alternate demand size, however, can change in one of two ways: directly by changing the proportion of customers willing to switch or indirectly by changing the supply of an item. Since there is an overall number of units of supply and of demand, when something changes in one outcome (labelled "1st" in the table) it is linked to subsequent changes in the others (2nd, 3rd, 4th, or 5th). In this study, subsequent changes are referred to as steps, whereas other studies may refer to similar changes as "indirect dynamics of substitution" (Schlapp & Fleischmann, 2018).

The impact of increasing either the focal or alternate item TSL is expected to take place in 5 steps. First of all, increasing the supply of an item should be linked to a higher likelihood of initial or alternate customers being able to find the item in stock (1st ↑). As more initial customers are able to buy their initial item, the number of initial customers who face SHO and are willing to substitute should be smaller (2nd

↓). Fewer customers approaching the alternate item should be linked to a smaller number of alternate sales of that alternate item (3rd ↓). This in turn should leave more supply of the alternate item to its initial customers (4th ↑). Since the initial customers of the alternate item are able to buy their initial preference in greater numbers, the number facing SHO and willing to switch should be lower (5th ↓). Overall, 3 of the sales outcomes increase with increasing TSL, while one of the sales outcomes decreases.

The relationship between increasing willingness to switch and demand and sales outcomes is expected to take place in 6 steps. Increasing the willingness to substitute for a group of customers should directly increase the number of customers willing to switch of the group of customers who face SHO (1st ↑). This increased demand size means more customers will approach their alternate preference, more likely resulting in greater sales (2nd ↑). Since a greater number of units of each item is being purchased by alternate customers, there will be fewer units of the item available for its initial customers (3rd ↓). Next the number of customers willing to switch will be greater, not because the proportion has increased (as in step 1) but because the number of customers facing SHO has increased (4th ↑). This also should be linked to more alternate sales (5th ↑) which in turn means fewer initial customers to that item are able to purchase it (6th ↓). Overall, 2 of the sales outcomes increase with increasing substitutability, while two of the sales outcomes decrease.

Presenting the drivers of customers switching and switching outcomes in this manner can answer the question of the usefulness of the FRH. The FRH framework does incorporate steps 1-3 (Table 2-4) for changes to TSLs. It also incorporates steps 1 and 2 for changes to customer willingness to substitute. It does not expect that changing the supply of one item will indirectly change the relative supply of another item as well as changing the alternate demand size of that other item, as listed in steps 4 and 5, respectively of columns 1 and 2 in Table 2-4. Similarly, the FRH framework does not incorporate (in columns 3 and 4) a relative change in the supply of one item (step 6) with a change in demand (willingness to switch) of another item. This type of interaction between all customers and the available supply of both items does take place in the ABM framework, not because it is modelled in with a priori assumptions but because each component (a unit of demand and a unit of supply) has its own myopic rules of behavior which are left to interact with one another. That means when comparing the FRH and the ABM it is possible that there will be changes to switching outcomes in ABM that are not in FRH.

Even if a comparison of the customer switching outcomes from the two frameworks shows when the FRH sufficiently captures the phenomenon, it does not provide any measure of switching's impact on performance fill rates. None of the sales rates among these switching outcomes are the actual fill rates used in practice. In practice, fill rates are different combinations of these outcomes. Since the sales outcomes are not expected to all move in the same direction with an increase in TSL

or substitutability, it is possible these customer switching drivers have differing relationships to different fill rate measures.

2.3.3 Fill rate measures

The fill rate measures used in the second part of the study include various combinations of the 8 outcomes outlined in Figure 2-3. The components consist of original (core) demand, initial purchase, and alternate purchase, where demand (in the denominator of a fill rate measure) refers to the size of initial customers intending to buy an item, and purchase (numerator) refers to the store's supply. Since there is an assumption that the supply will not fully meet the demand (TSLs are not 100%) the number of units an item has sold is actually a measure of item supply instead of customer demand. The fill rates also include a component that is not explicitly stated using its own notation in Figure 2-3, and that is the alternate demand size after the initial customers have faced SHO. This alternate demand size depends on the TSL of the initially preferred item and the level of customer willingness to substitute. Initial and alternate demand as well as purchase size components in different combinations constitute the fill rates in this study.

Every row in Table 2-5 is for a different performance fill rate measure used in this study: item, customer, initial customer and category fill rates. The second column of the table (Description) describes the components of each fill rate. In this column, every fill rate numerator includes initial sales of the focal item and every denominator

includes the size of focal item initial demand. The remaining components are what differentiate one fill rate measure from another. Without customer switching, the item, initial and customer fill rates would be indistinguishable because there would be no alternate customers demanding or buying items. Table 2-5 lists the focal item fill rate and the initial and overall customer fill rates for the focal customer and excludes the same fill rates for the alternate-item and -customer fill rates to save space and avoid redundancy. The category fill rate is the only performance measure in this study which considers total customer demand size (sum of initial or core demand of each item) and both initial and alternate sales of both items.

The category and initial fill rates are included to serve a specific purpose. First, the category fill rate is included because it is an overall performance measure which should increase with customer switching because of fewer units of lost sales (McGillivray and Silver, 1978). If this performance measure shows a significant link to customer switching drivers while the other fill rates do not, then it illustrates how a single phenomenon can be linked to performance outcomes with contradictory findings. On the other end of the spectrum is the narrowest fill rate measure, initial fill rate. The retailer's main goal is to supply an item that a customer wants so increasing its supply (focal item TSL) enables increased sales to those initial customers. While initial fill rate has no components of alternate demand or sales, increasing an item's supply allows for alternate customers to also purchase it if their initial item is SHO so that alternate TSLs and willingness to switch could be correlated to the initial fill rate.

Fill rate	Description	Impact of increasing focal item TSL	Impact of increasing alternate item TSL	Impact of increasing focal customer willingness to switch	Impact of increasing alternate customer willingness to switch
Item	$\frac{initial\ sales\ of\ focal\ item + alternate\ sales\ of\ focal\ item}{initial\ demand\ of\ focal\ item + alternate\ demand\ of\ focal\ item}$	$\frac{\uparrow + \uparrow}{\approx + \downarrow}$	$\frac{\uparrow + \downarrow}{\approx + \downarrow}$	$\frac{\downarrow + \uparrow}{\approx + \uparrow}$	$\frac{\downarrow + \uparrow}{\approx + \uparrow}$
Customer	$\frac{initial\ sales\ of\ focal\ item + alternate\ sales\ of\ alternate\ item}{initial\ demand\ of\ focal\ item}$	$\frac{\uparrow + \downarrow}{\approx}$	$\frac{\uparrow + \uparrow}{\approx}$	$\frac{\downarrow + \uparrow}{\approx}$	$\frac{\downarrow + \uparrow}{\approx}$
Initial	$\frac{initial\ sales\ of\ focal\ item}{initial\ demand\ of\ focal\ items}$	$\frac{\uparrow}{\approx}$	$\frac{\uparrow}{\approx}$	$\frac{\downarrow}{\approx}$	$\frac{\downarrow}{\approx}$
Category	$\frac{initial\ sales\ of\ both\ items + alternate\ sales\ of\ both\ items}{initial\ demand\ of\ both\ items}$	$\frac{\uparrow + \uparrow + \uparrow + \downarrow}{\approx + \approx}$	$\frac{\uparrow + \uparrow + \downarrow + \uparrow}{\approx + \approx}$	$\frac{\downarrow + \downarrow + \uparrow + \uparrow}{\approx + \approx}$	$\frac{\downarrow + \downarrow + \uparrow + \uparrow}{\approx + \approx}$

Table 2-5 Fill rate measures in terms of customer switching outcomes

The customer and item fill rates are included in the study for two reasons. First, the retailer stocks items according to effective demand but considers only initial customers. This stocking procedure uses a fill rate that has the numerator of the item fill rate and the denominator of the customer fill rate in Table 2-5. This results in two different sources for skewed fill rates, one from using a numerator of "initial and alternate sales of focal item" and the other from using a denominator of "initial demand of focal item." By using effective demand in the numerator of the fill rate it includes alternate sales of an item, requiring an inflated supply of that item by a degree that would normally not occur if the *alternate* item was sufficiently stocked. By ignoring demand from switching in the denominator of the fill rate, it overlooks the combined demand that the supply must meet, from both initial and alternate customers. This makes the denominator smaller than it should be. The fill rate in practice thus misguides the retailer because it is inflated from larger numerators and smaller denominators than the item and customer fill rates used in this study. These two fill rates are included to better capture the proportion of an item's supply which meets the total demand for the item (item fill rate) and the proportion of initial customer size able to buy any item (customer fill rate).

Secondly, the item and customer fill rates allow for separately viewing the impact of supply versus demand drivers to switching. Since the customer fill rate (Zinn and Liu, 2001) follows how well a customer group's demand is served by the store carrying multiple items (through switching), a higher customer fill rate with the same amount of inventory is evidence of the risk pooling effect. The item fill rate

measures how well multiple customer groups are served by the store's supply of a single item, so that higher fill rates are evidence of the "store-level fill rate effect" (Waller et al., 2008) where having more inventory at the store allows for greater sales. Incorporating both of these measures into one study brings together the operations mindset of product availability with the marketing mindset of customer service. If item fill rate is more of a product availability measure, as discussed earlier (section 2.2.2), then the impact of inventory investment (TSLs) should be larger than it is for the customer fill rate. Similarly, if the customer fill rate captures risk pooling effects from switching, then then the customer willingness to switch (a demand driver) should have a greater impact on it than on item fill rates.

The remaining columns of Table 2-5 illustrate how the impact of customer switching drivers may vary between fill rate measures because of their different components. It is beyond the scope of this research to hypothesize or test the direction of impact, whether the change in each fill rate measure overall is positive or negative. Fill rate measures, as this table shows, are complicated because of their components, and can mislead our understanding of performance improvement. This study explores how each switching driver can have a different size of impact on each different fill rate measure. Billboard and scarcity effects (Xue et al., 2017) are also out of the scope of this research since the ABM framework clearly defines a customer as switching to an alternate item only when the initial item is SHO, and not because the alternate item has higher on-hand inventory than the non-SHO initial item. Therefore an increase in inventory (increased TSL) will not increase the size of initial or

alternate customer demand. This lack of association with initial demand size is portrayed throughout the table with a "≈." As stated earlier, increased customer willingness to substitute also does not change the size of initial demand. Each expression of a fill rate measure in the table represents the main effects of increased TSLs or substitutability, a main effect which is comprised of multiple customer switching outcomes which sometimes vary in direction.

Performance fill rates may also be affected by a combination or the interaction between the four drivers of customer switching in addition to their main effects. It is possible that performance improves at a greater rate when the inventory of the focal item increases and the willingness of alternate customers to buy the focal item also increases. Indeed, it may increase even more so if the inventory investment of the alternate item is lowered. On the other hand, if alternate customers buy the focal item at greater numbers this may result in the focal item stocking out at a greater rate than planned with TSLs, and a greater proportion of focal customers may leave without any purchase since the alternate item would also have stocked out. This would mean worsening customer fill rates for the focal customers. Since all of these interactions may pull different fill rates in different directions, it is unclear which factor or set of factors would have the greatest impact on each type of fill rate.

2.4 Experimental design

2.4.1 Design assumptions

For the most part the overall design is a typical two-product symmetric problem (Huang et al, 2011). First, the inventory system is under continuous review (Sivakumar, 2008) so that the store is always aware of on-hand inventory levels. Secondly, there is normally distributed (truncated) demand for each item (Rajaram and Tang, 2001). Third, demand arrives one unit at a time (each customer wants one unit only). This means in the continuous review system the inventory position will never drop below the reorder point. Fourth, there is equal mean and standard deviation of demand and of demand during lead time (Shin et al., 2015). The order quantity of each item, which is for a period of 10 days, is also equal for both items.

Assumptions that differ from the two-product symmetric problem of prior studies help set the foundation for this study's contributions. The substitutability is not symmetric as generally assumed (Nagarajan and Rajagopalan, 2008). It is possible that a greater (or smaller) proportion of initial customers for one item are willing to switch to the alternate item than in the other direction. Additionally, lead time is not zero and if there is unsold inventory it is carried over to the next period unlike substitution studies using newsvendor models (Khouja et al., 1996; Huang et al., 2011). There is instantaneous replenishment of shelves as soon as the item arrives at the store but it takes 5 days from placing the replenishment order to store arrival.

This nonzero lead time period is important because it is during this time interval that alternate demand from other items which have SHO are able to switch to the available item during its lead time period. If these SHO items had instant replenishment (zero lead time) then their initial customers would not need to switch, so a period of SHO is necessary for the switching to occur. The item could be SHO because it has a lead time instead of instantaneous replenishment. But it could also be SHO because its demand arrived faster (reorder point reached earlier) than expected due to additional alternate demand, putting it out of sync with the reorder cycle of the rest of the items. Having switching in both directions at different rates allows for an item to be alternately demanded and thus SHO before it is in its own lead time interval period. The assumption recognizes the store may replenish one or some of the items in a product category without simultaneously replenishing all of them.

There are also assumptions which further isolate the switching phenomenon to the only two factors of store's decision of how much inventory to hold and the customers' ability to buy an item or leave without purchase. There is no lag or error on the part of store personnel in putting the right items in the right place, or if there is a backroom, in realizing a product is out of stock and refilling the shelves from excess stock in the backroom. Indeed, there is no backroom effect or inventory overflow (Waller et al., 2008). There are no time, space or human labor effects in terms of the supplier. There are no errors in order receipt or delivery in terms of quantity, quality or timing. There is no backordering; if an item is not available to the customer upon arrival, it is not provided to that customer at a later time. Orders are not combined

since the products come from different suppliers. There is also no order crossover since the lead time is constant.

	Description	Values of each level
Target Service Level	The target fill rate the retailer chooses for the focal item X	SL_X={50, 75, 95}
	The target fill rate the retailer chooses for the alternate item Y	SL_Y={50, 75, 95}
Customer Willingness to Substitute	Proportion of focal item's customers who would buy substitute Y if X is stocked out	Sub_{XY}={0, 50, 100}
	Proportion of the alternate item's customers who would buy item X if Y is stocked out	Sub_{YX}={0, 50, 100}

Table 2-6 Experimental factors

The first part of the study is more restrictive in its assumption of target service levels. The experimental factors in this study's design are listed in Table 2-6. While in the second part of the study each item can have a different level of inventory investment within a scenario, in the first part of the study there is one TSL for both items in each scenario (SL_X=SL_Y). Depending on the scenario, both items can have a TSL of 50%, 75% or of 95%. Symmetric inventory investment removes any differences from supply-based decisions so the focus can be on customer behavior and the on-hand inventory resulting from that behavior. The items may each reach the reorder point at slightly different times but have approximately the same level of availability, being reordered at the same level and almost same point in time. In this

way the customer's behavior is isolated and the switching outcomes can be aggregated to see what, if any, interaction effects there are because of random customer arrival sequence.

In the second part of the study, the focus shifts to how this customer switching behavior impacts performance fill rates. By picking one of the two items as a focal item, the TSLs are then allowed to be different from one another so that SHOs occur earlier for the item with lower TSLs. In this way there are different levels of customer switching, because there are different levels of customer willingness to switch and different levels of inventory investment. The focal item can be stocked more (less) than the alternate item and have a greater (smaller) proportion of customers willing to switch to the alternate item. Varying these amounts attempts to capture a wider spectrum of switching than symmetric TSLs and substitutabilities can. It would be difficult to ascertain the effect of the alternate item's TSL on the focal item's fill rate if both the supply and demand drivers of switching were symmetric.

2.4.2 Notation for customer switching outcomes

To better describe the different customer switching outcomes of the first part of this study, additional sub- and superscripts, shown in Figure 2-5, are added to the notation introduced in the theoretical framework (section 2.2.3). After the subscript for purchase outcome and superscript for customer type is a bracket, "]". To the right of the bracket are the experimental conditions, the drivers to customer switching. The supply driver is the superscript, which identifies what proportion of customers who

initially prefer item X or item Y are willing to switch when faced with SHO. The subscript is the TSL of both of the items which are stocked at the same level of inventory investment. In the example in Figure 2-5, none of the focal customers (*x*) are willing to switch to item Y, while half of the alternate customers (*y*) are willing to switch to item X. In this example X and Y are both stocked such that the retailer expects to serve 75% of core demand (initial customer size), with the remaining 25% facing SHO. Overall, the example in this figure is the notation for the likelihood that alternate customers, half of whom are willing to substitute, are able to buy the focal item, where both items are stocked with a TSL of 75%.

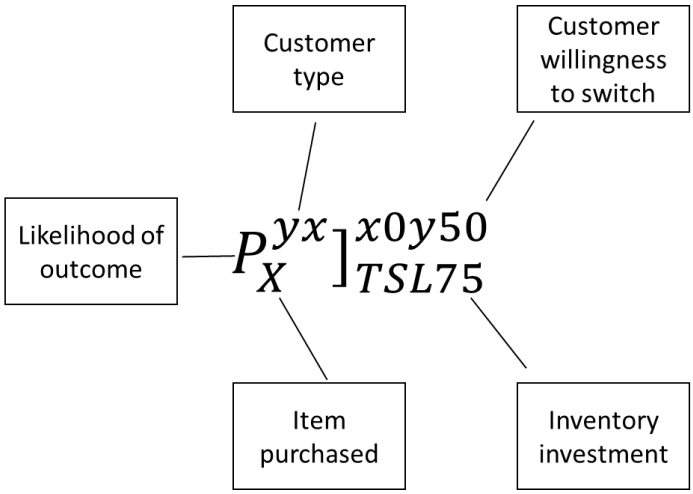


Figure 2-4 Notation scheme for switching outcomes

2.4.3 Computational generation of outcomes for the ABM framework

Although simulation has been used for more than half a century in inventory management research (Lucas & Moody, 1957), it is still considered one of the "valid

but unfamiliar methods" (Waller & Fawcett, 2011, p. 209) in supply chain research. Simulation allows for "complete control over the system without having to make potentially restrictive assumptions and… [to] isolate the results by avoiding confounding factors" (Evers & Wan, 2012, p. 82). The problem considered here is complicated given heterogeneous customer preferences, random order arrival sequence and non-symmetric substitutabilities. Because of this complexity, simulation is chosen to generate switching outcomes.

Of the three types of simulation methods, the most suitable approach for a individual-level view of customer switching is discrete-agent or agent-based (ABMS) modeling and simulation. The other two methods are discrete event (DES) and system dynamics simulation methods. Both are top-down models of computation which assume that the aggregate behaviors of processes or events are simply a sum (and are no different) than individual occurrences. While system dynamics simulation is appropriate for finding hysteresis (lag) and acceleration effects (bullwhip) (Christopher, 1976, p. 29), DES is suitable for processes where the underlying structure is already known. On the other hand, ABMS allows for "patterns, structures, and behaviours emerge that were not explicitly programmed into the models, but arise through the agent interactions" (Macal & North, 2014, p. 11). Mour et al. (2013) provide a further differentiation and background between the different simulation methods.

Agent-based simulation models (or agent-based model simulation, ABMS) have four characteristics that specifically suit stockout-based customer switching: a

bottom-up perspective, boundedly-rational agents, networked direct interactions, and temporal aspects (Robinson, 2015). Necessary for finding the underlying structure of phenomena, a bottom-up perspective has "strongly *heterogenous* agents living in complex systems that evolve through time…[so that]…aggregate properties are interpreted as emerging out of repeated interactions among simple entities" (Windrum, Fagiolo, & Moneta, 2007). Boundedly-rational agents have "myopic optimisation rules" (Windrum, Fagiolo, & Moneta, 2007) where agents have to react and adapt to endogenously changing environments. Additionally, the agents in ABMS "interact directly because current decisions directly depend, through adaptive expectations, on the past choices made by other agents in the population" (Windrum, Fagiolo, & Moneta, 2007). These networked direct interactions can be overlooked in the FRH framework presented earlier (Figure 2-3). Last of all, Rand and Rust state that temporal aspects are "almost a necessary condition for the ABM approach" (2011, p. 185). Starting out with two items that have equal initial demand size and inventory levels, this experimental design relies on the temporal aspect of customer switching, where the timeline for supply is altered by splitting demand streams over time.

Of the two types of agent-based modeling this study uses an active walker model. A random walk model does not interact with the environment so that "the space in which the random walker walks is always passive and time independent" (Freimuth & Lam, 1992). An active walker better reflects customer-based switching because agents, according to Freimuth and Lam (1992), have deterministic or

probabilistic rules that tell them what step to take next when faced with their current environment. In this study each agent has a deterministic set of items ranked according to the agent's initial and alternate preferences, as listed in Table 2-6. The interaction with the environment occurs when a customer arrives to the shelf for the initially preferred item or alternate items which may not be available due to prior customer purchases. Items drop in on-hand inventory levels and also reach their reordering point after customers take items from the shelf. In this way, the agents interact both with one another and the environment.

The customers in the ABM simulation (ABMS) in Table 2-7 have myopic rules depending on their customer type and they directly interact with one another because of product availability. While customers at a store may never see one another at the product category shelf, the resulting SHO from a prior purchase of one customer determines what a subsequent customer will do. There are 4 different customers (4 cells in Table 2-7), with customers who initially prefer item X ("Initial item: X") having the same structure of decision rules as those customers who initially prefer item Y. Mapping these agent definitions to the outcomes described in the theoretical framework an initial customer of item X who is not willing to substitute has two possible outcomes (P_X^x or P_0^x), whereas those willing to substitute have three outcomes (P_X^{xy} , P_0^{xy}, P_Y^{xy}). For each customer type, their outcome is determined entirely by product availability at the time of store visit, with the product availability depending on prior customer purchases and store replenishment timing.

	Initial item: X	Initial item: Y
Willing to substitute	Buy X if in stock, else Buy Y if in stock, else Leave without purchase	Buy Y if in stock, else Buy X if in stock, else Leave without purchase
Not willing to substitute	Buy X if in stock, else Leave without purchase	Buy Y if in stock, else Leave without purchase

Table 2-7 Customers in the ABM framework of customer switching

The customer types are simulated using ARENA 14.5 (Rockwell Automation, 2013). Since ARENA is not ABMS software, the agents are created by using existing program entities in an unconventional way. The "job.sequence" entity attribute, which is normally used for items to move from one machining step to another after a certain period of processing time, is used to move customers instead, on condition of product availability (instead of time) using the decision rules in Table 2-7. The job sequence thus becomes the customer preference ranking: "[f]or example, a customer of type (1,2,4) has product 1 as his first choice, product 2 as his second choice, product 4 as his third choice, and he never buys products 3 or 5 to *n*" (Honhon & Seshadri, 2010, p. 1367). For every new day generated in the simulation, the program pulls that day's quantity of customers from a truncated normal distribution of daily demand ($\mu_D = 8$, $\sigma_D = 2$) for each item. Whatever proportion of these customers are supposed to be willing to substitute are assigned by chance to each item's core demand and given the job sequence of 3 steps. The remainder of the core demand of each item are given the two-step job sequence, as listed in Table 2-7.

After defining the agents (the demand) in the simulation program, the store shelves (supply) also need to be simulated. Supply is defined by a reorder point corresponding to each TSL in the experimental design: B_{50}= -40, B_{75}= 13, and B_{95}= 36 (Table 2-3 provides step-by-step directions). The process of calculating the ROP are "mere approximations to the optimal solution" (Tersine, 1994, p. 227) so they are lower bounds to the TSL so that actual fill rates may be higher, even without customer switching. While the calculations for expected stockout, E(M>B), had been developed for the case of complete backordering, it can be applied to this study without backordering because "the service level fraction for units demanded is insensitive to backorders or lost sales and gives essentially similar results in either case or any mix of these two extremes" (Tersine, 1994, p. 220). Along with defining the reorder points in the program, the initial on-hand inventory is also entered. Both items have Q order size units (80 units in this study) of on-hand inventory at the start of the simulation. For replenishing supply, whenever the inventory position of either product has been reduced to its reorder point, the retailer sends an order of quantity Q to the supplier. Replenishment occurs after a constant lead time, L (5 days in this study). It does not occur at the beginning of a newly simulated day, but at a time exactly L days after the customer purchase which has triggered the order so that the two products are not reordered or replenished at exactly the same time unless by chance.

The simulation program is then run for 3 x 3 x 3 x 3 factor levels (81 scenarios, though the first part of the study uses outcomes from 27 scenarios). Each

scenario is replicated 30 times for a length of 5,000 days per replication, which is double the suggestion of Law and Kelton (2000, p. 512) who suggest 15 replications per scenario for estimating the unit fill rate with an absolute error of 0.05 and confidence level of 90%. The first 500 days of each replication is dropped as transient data from the simulation's initial bias. This warm-up period is determined using a graph of on-hand inventory over time (sawtooth model) and truncating a period twice as long as the time at which the sawtooth model enters a steady pattern. Fill rate statistics, thus, are calculated from the consumer responses from day 501 through 5,000. The simulation run time for each scenario is about 6 minutes, for a total processing time of 8.1 hours.

2.4.3.1 Model verification and validation

Rand and Rust (2011) outline a rigorous process of verification and validation for ABM research, comprised of 7 steps. The first 3 steps are verifying the model: documentation, programmatic testing, test cases and scenarios. Documentation is provided as the sections on developing the ABM framework and on the experimental design. Regarding programming documentation, ARENA uses a front-end drag-and-drop modelling system for its backend code. All sections of the code are split into their respective modules automatically, and descriptions are added within each module of the function or calculations that take place at every step. The initial model is completely deterministic in nature, with a known number of customers being created for each product type and no substitutability between them. This establishes

the fundamental logic of the inventory tracking system, as outlined in the process map. Upon validating initial simulation results from this model matched expected (manually calculated) values, the first phase of building the model is complete with unit testing. Additional phases include assigning customers a preference ranking, assigning a normal distribution for product demands, and randomizing customer arrival sequence. Every phase of development follows the same verification-validation procedure until the model matches the process map presented in the ABM framework. Code and debugging walkthroughs also take place at each phase and when the programming is done. Formal testing of the final model, due to its stochastic nature, is qualitative in nature. For example, if the TSLs are lowered (raised) the resulting initial sales of the product also decrease (increase). Using logic in this step as formal testing is the norm for ABM (Rand and Rust, p. 187) and relative value testing examines the relationships between inputs and outputs in this manner. The corner cases were considered in terms of the "no customer switching" scenario, following the unit testing process.

The remaining 4 steps are for validating that the ARENA model corresponds to reality: micro- and macro-face validation, empirical input and output validation. For micro-validation, the components of customers switching (customers and items) need to correspond to real-world properties. In this simulation program, individual customers each have initial preferences and will automatically switch to the alternate item if they were already willing to substitute and face SHO. In realty, customers have been shown (Zinn and Liu, 2001) to substitute, when faced with SHO, if they

stated such an intention beforehand. With regard to the supply (the items in the simulation), stores in reality do use reorder points and target service levels. They do face lead time, although that lead time may be stochastic in nature.

In terms of macro-face validation, this study focuses on the patterns that will emerge from the customer switching environment, and thus cannot compare simulation outcomes with known process outcomes beyond the existing contradictory findings that switching sometimes increases fill rates and sometimes does not. For empirical input validation, the range of substitutability between 0 and 100% reflects a realistic range as does target service levels at or below 50% (Anderson, et al., 2006) and up to 95%. For empirical output validation, this study rests mainly on stylized facts of inventory management of substitutable items. There is no real-world data testing, as no predictor model is being developed, and no cross-validation, as it is "an optional validation process that compares the new model against another model that has already been validated" (Rand and Rust, p. 189). The comparison of ABM with FRH is not for validation purposes but to see how the interaction between customers and items may yield scenarios contrary to risk pooling theory and the FRH.

2.5 Analysis

2.5.1 Comparing FRH and ABMS outcomes

The simulation output generates the experimental cross-sectional data consisting of counts. There are 2,430 data points (81 scenarios times 30 replications,

as counts) for every customer switching outcome, over all of the scenarios, for a total of 19,440 different outcome outputs from the ARENA simulation. In addition to these counts, the software also provides, for every replication of every scenario, the counts of the randomly generated initial demand for each item and the number of initial customers willing to switch. Since simulation values are all counts, they are converted into rates so they can be compared to the FRH outcomes. Figures 2-6 and 2-7 provide an example of the values to be compared for both views of a single scenario. In the example, a random scenario has been picked with a TSL of 75% for both items, with 100% of focal customers and 50% of alternate customers willing to substitute.

The FRH outcomes are calculated in Figure 2-6 according to the conditional probabilities described earlier. In the example in Figure 2-6, of all of the focal customers, 18.75% should be able to buy the alternate item ($P_Y^{xy}]_{TSL75}^{x100y50} = 0.1875$). In contrast, only about 3% of alternate customers ($P_X^{yx}]_{TSL75}^{x100y50} = 0.03125$) are able to buy item X as a substitute. Since the core demand for each product and TSLs are equal in this study, for easier interpretation of the results the calculations for each product branch is taken separately, $\left(P_X^x + P_X^{xy}\right) + P_0^x + P_Y^{xy} + P_0^{xy} = 1$ for focal item X, and $(P_Y^y + P_Y^{yx}) + P_0^y + P_X^{yx} + P_0^{yx} = 1$ for alternate item Y. The sum of all of the likelihoods for each customer do indeed equal 1: (0.75 + 0 + 0.1875 + 0.0625 = 1 for focal customer, and 0.75+0.125+0.09378 + 0.03125=1, with a small rounding error—0.00003--for alternate customer).

The corresponding rates in the ABM view are calculated using the simulation outputs. In the example provided, the average is taken of the 30 replications of the scenario to get a single mean count for each outcome and demand. The example scenario's average values for the alternate (*y*) customer are: 18,014 customers not willing to substitute and 18,014 customers willing to substitute, for a total of 36,028 customers. The average counts of the outcomes are: 24,855 for initial purchases, 1,110 for buying the substitute, 5,570 and 4,464 for customers leaving without a purchase. The sum of these is 36,029 instead of 36,028 since these are averaged count values. These counts are then turned into rates by dividing each count by the average generated demand of 36,028.

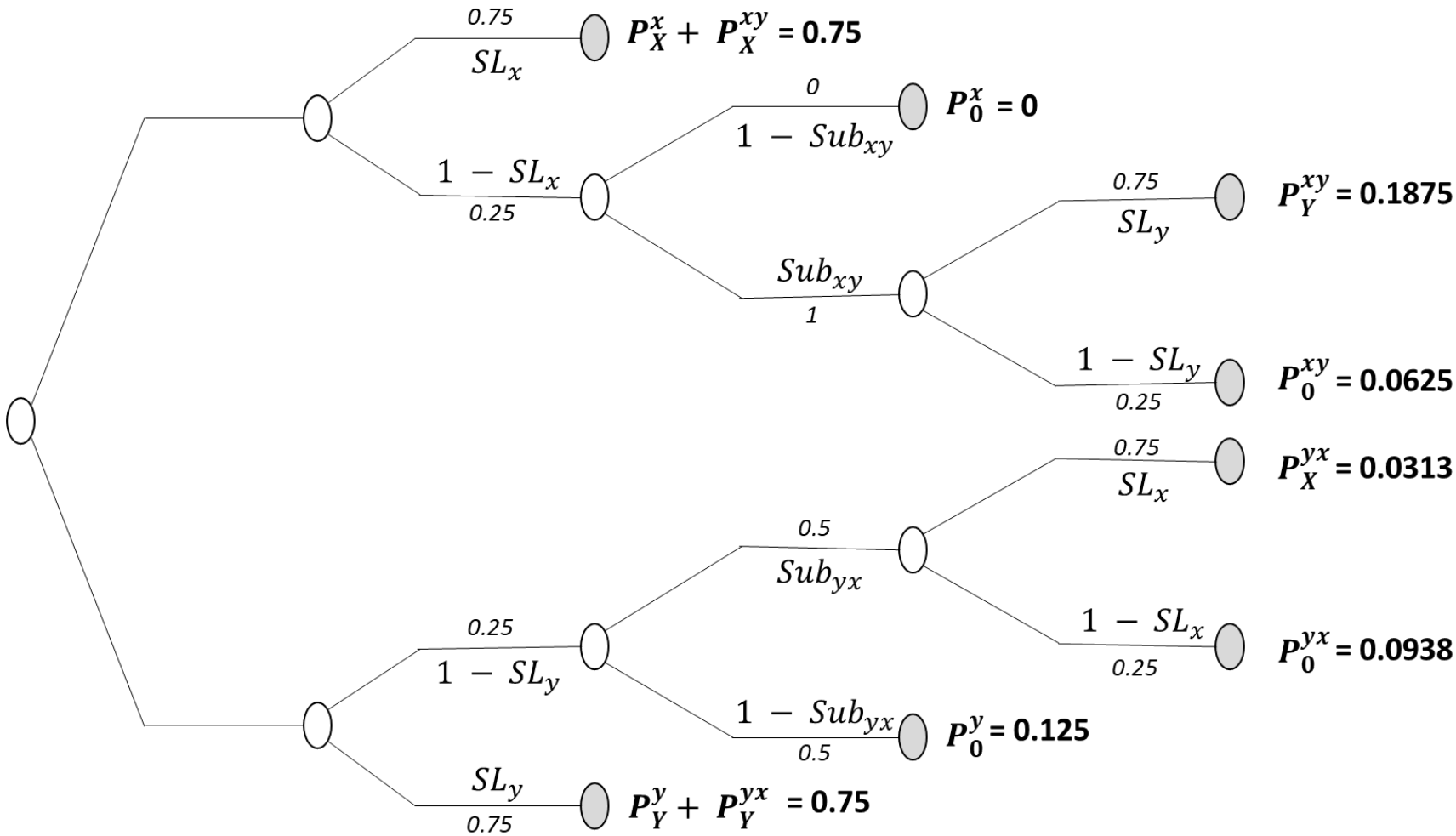


Figure 2-5 Example scenario calculations for FRH framework

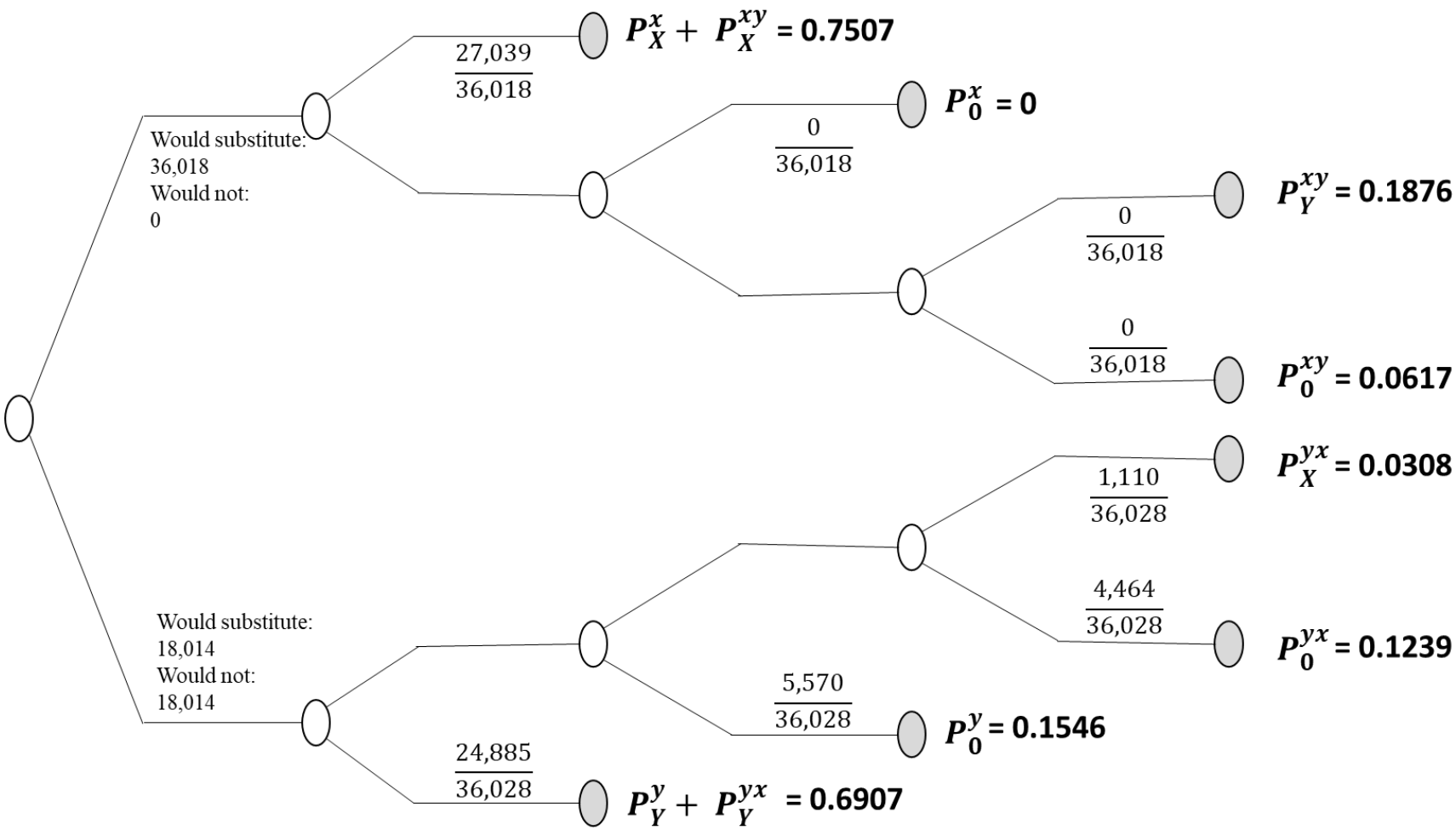


Figure 2-6 Example scenario calculations for ABM framework

2.5.2 Statistical analysis

This section presents the statistical analysis of 27 scenarios for the focal customer, computed in Stata 13.1 (StataCorp, 2016). Since the demand and TSLs are equal in this part of the study, the tables for alternate customers are provided in the appendix (Appendix III). A one-way ANOVA with a 0-1 independent variable (where 1 is the focal customer) is not statistically significant over any of the 27 scenarios (maximum F = 3.90, p = 0.0532 for $(P_X^x + P_X^{xy})]_{TSL50}^{x50y100}$). For each customer switching outcome there are two charts. The top chart lists the FRH value followed by the mean ABM value. The bottom chart lists the difference between the two followed by the t-statistic in parenthesis. In this bottom chart, positive values of the difference reflect scenarios where the ABM sold or stocked out at a rate greater than FRH, whereas negative values are when the FRH has sale or SHO outcomes occurring more often. In every chart, outcomes for scenario $]_{TSL75}^{x100y50}$ are darkened with a shaded box as that was the scenario used as an example in section 2.2.3.

Due to demand distribution assumptions, the statistical analysis compares outcomes from the two frameworks but also compares between scenarios for a better understanding of what drivers (and their interactions) to customers switching the FRH might have missed. ANOVA tests control for type I error in comparing mean scenario values with one another, since there are 8 outcomes for 27 scenarios and repeated t-tests are prone to type I error. A one-way ANOVA with Bonferroni correction ($\alpha = 0.0019$) using the 8 outcomes as dependent variables and the scenario number as the independent

variable indicate significant mean differences between scenarios (minimum F = 15,153, p < 0.0001). Tukey's Studentized Range (HSD) is conducted as the post hoc test since all pairwise comparisons are required and sample size is equal, with 30 simulation replications in every scenario.

Tukey groupings are a robust way of looking at the differences between the means of the scenario outcomes, without getting caught up in the individual mean value of any one outcome. Since the safety stock in the ABM was calculated using a continuous (normal) distribution even though units of inventory are discrete, its outcome values will differ by a few units from the FRH, even when there is no customer switching. This absolute difference in the values of both views might be important in a predictor model instead of a descriptive model that explores how the two views differ in looking at the customers switching phenomenon overall. That is why Tukey groupings are more helpful in this study to compare any differences in patterns across scenarios between the FRH and ABM.

The Tukey groupings are listed as letters in the top part of each of the 4 t-test tables below. The letter "A" refers to the highest scenario mean, "B" is the second highest mean, and so forth. Scenarios with the same letter do not have significantly different values. For example, in the first table of results (Table 2-8), $(P_X^x + P_X^{xy})]_{TSL95}^{x0y50}$ has Tukey grouping values of AB with mean 97.45, while $(P_X^x + P_X^{xy})]_{TSL95}^{x100y100}$ has Tukey groupings BC with mean 97.19. Since both scenarios are in Tukey group "B," they are not significantly different from one another. However, $(P_X^x + P_X^{xy})]_{TSL95}^{x0y50}$ is significantly greater than $(P_X^x + P_X^{xy})]_{TSL95}^{x50y100}$ while $(P_X^x + P_X^{xy})]_{TSL95}^{x100y100}$ is not

significantly different. Similarly, $(P_X^x + P_X^{xy})]_{TSL95}^{x100y100}$ is significantly less than $(P_X^x + P_X^{xy})]_{TSL95}^{x50y50}$ while $(P_X^x + P_X^{xy})]_{TSL95}^{x0y50}$ is not significantly different.

Initial sales rates of focal customers in the FRH differs from ABM. In the FRH, this rate of sales is equal to the TSL, regardless of customer willingness to switch. In terms of Tukey group letters, the nine TSL 95% scenarios should have all group "A," while 75% should have 9 "B" scenarios and 50% should have "C" grouping. Instead, as alternate customers are increasingly willing to switch to the focal item, the sales rate to initial customers significantly drops at the 95% TSL from A to B to C, where all alternate customers are willing to switch to X. When looking at the rate of sales of alternate customers of item X (the 3rd set of charts in the Appendix III), it does not show the same grouping pattern. There, $P_X^{yx}]_{TSL95}^{x100y100}$ has a significantly higher rate of sales than any of the other 8 scenarios at that TSL. Indeed, the Tukey grouping pattern of alternate sales of the focal item matches that of P_Y^{xy} charts in Table 2-10, where focal customers buy the alternate item. Since $(P_X^x + P_X^{xy})]_{TSL95}^{x50y50}$ is greater than $(P_X^x + P_X^{xy}]_{TSL95}^{x100y50}$ but $P_X^{yx}]_{TSL95}^{x50y50}$ is not significantly different than $P_X^{yx}]_{TSL95}^{x100y50}$ that means that the drop in the initial sales rate of the focal item with increasing alternate customer willingness to switch does not appear as an increase in the alternate sales rates of the focal item.

As the ABM varies in mean rates while the FRH does not, it also exhibits a different pattern of rates within itself, depending on TSL. While TSL 95% consists of 3 Tukey groupings (A, B, C), TSL 75% consists of 6 groups (D, E, F, G, H, I) and TSL 50% consists of 5 groups (J, K, L, M, N). There is greater variation of outcomes at the

middle level of inventory investment rather than at either end. The suggests that the combination (interaction) of customer willingness to substitute for focal and for alternate customers have greater impact on initial sales at the middle level of inventory rather than when SHOs rarely occur (TSL 95%) or occur extremely frequently (TSL 50%). When inventory investment is at its lowest (TSL 50%) the initial sales rate is at its highest at $(P_X^x + P_X^{xy})]_{TSL50}^{x100y100}$followed by $(P_X^x + P_X^{xy})]_{TSL50}^{x50y50}$, in other words, with increased symmetric substitutability. However, at the middle range (TSL 75%) the lowest initial sales rate occurs in full single direction substitutability $(P_X^x + P_X^{xy})]_{TSL75}^{x0y100}$ with the highest initial sales at no substitutability $(P_X^x + P_X^{xy})]_{TSL75}^{x0y0}$. At this TSL, increasing substitutability decreases initial sales. Depending on the level of inventory investment customer switching has a positive or negative effect on initial sales.

The reorder point as the measure of inventory investment in this study may be linked to unintended supply factors when customer switching is involved. Indeed, the reorder point at TSL 50% is negative (-40), which means the store is not going to reorder until 40 initial customers face SHO. With a lead time that is half the time of an inventory period and allowing for a SHO length that is also half of the inventory period (since mean demand is 8) this means focal item X is going to be fully stocked every other inventory period. With increasing alternate *and* focal customer substitutability, the reorder point of the alternate item may occur sooner than without focal customer willingness to switch. This in turn may push alternate customers to the focal item to cause it to SHO (after it has been replenished) even earlier than halfway through the inventory period. This shortened period of availability would mean increased frequency of replenishment even if the SHO

$P_X^x + P_X^{xy}$ Tukey group FRH ABMS	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	M 50 51.35	M 50 51.35	M 50 51.21	D 75 77.82	D 75 77.82	D 75 77.82	A 95 97.72	A 95 97.72	A 95 97.72
Half	N 50 49.30	K 50 54.25	L 50 51.98	F 75 74.58	F 75 74.60	E 75 75.07	AB 95 97.45	A 95 97.48	A 95 97.55
All	N 50 49.42	L 50 52.23	J 50 57.62	G 75 72.59	I 75 69.07	H 75 70.23	C 95 96.01	C 95 97.04	BC 95 97.19
	None	Half	All	None	Half	All	None	Half	All

Focal customers: buy Y if X is out (Sub_{XY})

$P_X^x + P_X^{xy}$ Difference (t stats)	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	1.35 (94.19)	1.35 (93.49)	1.21 (49.30)	2.82 (91.25)	2.82 (91.25)	2.82 (91.25)	2.72 (94.34)	2.72 (94.34)	2.72 (94.34)
Half	-0.70 (-17.67)	4.25 (106.90)	1.98 (39.53)	-0.42 (-8.57)	-0.40 (-10.43)	0.07 (1.99)	2.45 (77.03)	2.48 (83.63)	2.55 (102.20)
All	-0.58 (-5.51)	2.23 (13.00)	7.62 (102.80)	-2.41 (-48.60)	-5.93 (-96.88)	-4.77 (-130.0)	2.01 (63.22)	2.04 (55.94)	2.19 (63.31)
	None	Half	All	None	Half	All	None	Half	All

Focal customers: buy Y if X is out (Sub_{XY})

Table 2-8 Initial item sales rate for focal customers

duration and lead time remain the same. The increased replenishment may mean higher levels of on-hand inventory overall than the store expects to hold at such TSLs, because of the alternate demand of customer switching.

The SHO duration may even decrease because more than the expected D_X customers try to buy the item, with the addition of alternate customers, D_Y*(1-SL_Y). Thus the 4,500-day length of the simulation may have a greater number of replenishments at TSL 50% with increased substitutability than no substitutability. This increased SHO frequency and replenishment may allow for a greater proportion of arriving focal customers over the entire time period to buy the focal item. While it has been said that more frequent replenishment is a factor in the inventory anomaly (Yang and Schrage, 2009), the increased replenishment frequency itself may be a result of a combination of inventory investment and willingness to switch. At higher inventory investment levels, however, the initial sales with full substitutability ($(P_X^x + P_X^{xy})]_{TSL75}^{x100y100}$ or $(P_X^x + P_X^{xy})]_{TSL95}^{x100y100}$) are significantly lower than with no substitutability ($(P_X^x + P_X^{xy})]_{TSL75}^{x0y0}$ or $(P_X^x + P_X^{xy})]_{TSL95}^{x0y0}$).

The P_0^x charts in Table 2-9 list the proportion of initial customers who are not willing to substitute and leave when faced with SHO. It is a SHO frequency for initial customers of the focal item which avoids double counting, since the customers who are willing to switch are not included in these rates. In the FRH, these rates are simply the conditional probability depending on the TSL and the focal customer willingness to switch. As such, when focal customers are all willing to switch, the P_0^x outcome is 0 while at other times outcomes do not vary by alternate customer willingness to switch,

P_0^x Tukey group FRH ABMS	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	B	E	Q	H	K	Q	N	P	Q
	50	25	0	25	12.5	0	5	2.5	0
	48.65	24.35	0	22.18	11.12	0	2.28	1.14	0
Half	A	G	Q	D	J	Q	M	P	Q
	50	25	0	25	12.5	0	5	2.5	0
	50.70	22.91	0	25.42	12.74	0	2.55	1.26	0
All	A	F	Q	C	I	Q	L	O	Q
	50	25	0	25	12.5	0	5	2.5	0
	50.58	23.89	0	27.41	15.47	0	2.99	1.48	0
	None	Half	All	None	Half	All	None	Half	All
	Focal customers: buy Y if X is out (Sub_{XY})								

P_0^x Difference (t stats)	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	-1.35	-0.65	-	-2.82	-1.38	-	-2.72	-1.36	-
	(-94.19)	(-20.12)	-	(-91.25)	(-56.62)	-	(-94.34)	(-86.64)	-
Half	0.70	-2.09	-	0.42	0.24	-	-2.45	-1.24	-
	(17.67)	(-52.09)	-	(8.57)	(7.62)	-	(-77.03)	(-72.12)	-
All	0.58	-1.11	-	2.41	2.97	-	-2.01	-1.02	-
	(5.51)	(-11.44)	-	(48.60)	(66.66)	-	(-63.22)	(-47.04)	-
	None	Half	All	None	Half	All	None	Half	All
	Focal customers: buy Y if X is out (Sub_{XY})								

Table 2-9 Initial item SHO rate for focal customers

according to the FRH. However, the ABMS does show varied outcomes depending on alternate customer willingness to switch, with increasing SHO rates as substitutability increases—in contrast to traditional risk pooling theory.

An interesting outcome is that initial customers face more SHO at TSL 75% if alternate customers are willing to switch ($P_0^x]_{TSL75}^{x0y100}$ and $P_0^x]_{TSL75}^{x0y50}$) than at TSL50% when half of the focal customers are also willing to switch (($P_0^x]_{TSL50}^{x50y100}$, $P_0^x]_{TSL50}^{x50y50}$, and $P_0^x]_{TSL50}^{x50y0}$). In other words, the store loses up to 5% ($P_0^x]_{TSL75}^{x0y100} - P_0^x]_{TSL50}^{x50y50}$) more focal customers with higher inventory investment and one-way substitution as opposed to two-way substitution and lower inventory. In the FRH, the SHO rate is 25% for all of these scenarios (Tukey groups C, D, H and F, G, E in Table 2-9). The direction of the difference between FRH and ABMS outcomes (whether SHOs are above or below 25%) depends on the direction of the substitutability. Table 2-4 suggested that initial sales of the focal item would go down as alternate customer willingness to substitute increased because alternate demand sales increase. When alternate customers switch to the focal item at TSL 75% (the bottom two rows of the chart) there are fewer units of the focal item for the initial customers to be able to purchase. For these scenarios the FRH underestimates SHO rates. The outcome at $P_0^x]_{TSL50}^{x50y0}$ shows a significantly higher SHO rate than $P_0^x]_{TSL75}^{x0y0}$ (Tukey group E is greater than Tukey group H), but increased SHO rates (Tukey groups C and D are greater than F and G) when there is one-way substitution towards the focal item. This supports Gilland and Heese (2013) findings that customer order arrival sequence affects product allotment, as the initial customers face SHO more because of prior purchases by alternate customers.

At the highest level of inventory investment, the SHO rates in Table 2-9 appear to be much smaller in the ABMS than in the FRH. Indeed, that is the general pattern in this chart outside of the scenarios mentioned above. Looking at the bottom chart in this figure and the bottom chart in Table 2-8 of initial sales, it is possible to separate scenarios where the difference is not due to customer switching. All 3 scenarios where no customer switching is possible have mirroring differences. For example, the difference between FRH and ABMS for $(P_X^x + P_X^{xy})]_{TSL50}^{x0y0})$ is the same as the difference at $P_0^x]_{TSL50}^{x0y0})$ in opposite directions (1.35 and -1.35 respectively). These are normal distribution assumption differences and can be ignored.

The focal customer's purchase of the alternate item is given as rates in Table 2-10. Naturally, both the FRH and ABMS have the value 0 whenever the focal customer is not willing to substitute. In the FRH the greatest proportion of focal customers who buy the substitute item is when the focal item has minimal inventory investment (TSL 50%) and all focal customers are willing to switch. In fact, the order for the scenarios from largest substitute sales to smallest non-zero rate are: $P_Y^{xy}]_{TSL50}^{x100y*} > P_Y^{xy}]_{TSL75}^{x100y*} > P_Y^{xy}]_{TSL50}^{x50y*} > P_Y^{xy}]_{TSL75}^{x50y*} > P_Y^{xy}]_{TSL95}^{x100y*} > P_Y^{xy}]_{TSL95}^{x50y*} > 0$, where the asterisk (*) refers to all levels of customer willingness to switch for the alternate customer. In contrast, the FRH varies in substitute fill rates depending on the alternate customer's switching behavior. Indeed, the outcomes vary the most at TSL 50% (7 Tukey groups) and the least at TSL 95% (5 Tukey groups).

Generally speaking the FRH and ABM frameworks are in accordance, where higher levels of inventory investment have smaller proportions of focal customers buying

the alternate item, but there are a few notable differences. First of all, the highest two substitute sales rates occur at the medium level of inventory investment (TSL %75) which again implies that there is such a thing as too few SHOs (at TSL 95%) and too many SHOs (at TSL 50%) for the switching phenomenon to occur. Second, a very high inventory investment with full two-way substitutability, $P_Y^{xy}]_{TSL95}^{x100y100}$, results in the same fill rate as a very low inventory investment with two-way substitutability when only half of the customers are willing to switch $P_Y^{xy}]_{TSL50}^{x50y*}$. This also implies that low switching can be due to too low SHOs or too high SHOs. Third, at the very low inventory investment scenario ($P_Y^{xy}]_{TSL50}^{x50y50}$) the FRH overestimates how much substitution will occur by about 10%. With full two-way substitutability $P_Y^{xy}]_{TSL50}^{x100y100}$ the difference is almost 20%. The FRH overlooks the impact of switching behavior on driving SHOs. In other words, the FRH framework does recognize the supply driver of SHO, the inventory investment level. It does not recognize the demand driver of SHO, prior purchases from customers who have switched. This coincides with the earlier discussion about Table 2-4, where the FRH accounts for the first few steps of increasing inventory or substitutability, but does not incorporate how interconnected the outcomes are, and ignores the subsequent steps.

P_Y^{xy} Tukey group FRH ABMS	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	M	E	B	M	H	F	M	L	J
	0	12.5	25	0	9.38	18.75	0	2.38	4.75
	0	10.17	18.95	0	6.31	9.63	0	1.09	2.15
Half	M	I	C	M	D	B	M	KL	J
	0	12.5	25	0	9.38	18.75	0	2.38	4.75
	0	2.75	14.72	0	12.12	18.76	0	1.21	2.37
All	M	G	H	M	D	A	M	K	I
	0	12.5	25	0	9.38	18.75	0	2.38	4.75
	0	6.98	6.23	0	12.39	29.06	0	1.44	2.73
	None	Half	All	None	Half	All	None	Half	All

Focal customers: buy Y if X is out (Sub_{XY})

P_Y^{xy} Difference (t stats)	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	-	-2.33	-6.05	-	-3.06	-9.12	-	-1.29	-2.60
	-	(-51.03)	(-140.0)	-	(-50.22)	(-120.0)	-	(-78.53)	(-81.50)
Half	-	-9.75	-10.28	-	2.74	0.01	-	-1.16	-2.38
	-	(-200.0)	(-79.51)	-	(60.37)	(0.10)	-	(-76.14)	(-90.93)
All	-	-5.52	-18.77	-	3.02	10.31	-	-0.94	-2.02
	-	(-42.17)	(-180.0)	-	(32.05)	(295.30)	-	(-46.30)	(-56.21)
	None	Half	All	None	Half	All	None	Half	All

Focal customers: buy Y if X is out (Sub_{XY})

Table 2-10 Alternate item sales rate for focal customers

The chart on SHO outcomes of focal customers who were willing to substitute (Table 2-11) illustrates the role of inventory investment on the customer switching phenomenon. At the highest inventory investment (TSL 95%) the scenarios all belong to Tukey group L. That is the same group as the scenarios with no focal customers willing to switch where FRH and ABMS are both 0. Naturally, if the store has enough inventory of each item it will not only be able to meet the demand of initial customers but also alternate customers, who likely approach their alternate item in very small numbers anyway. On the other end of the spectrum, the SHO rate with the highest difference between the two perspectives is $P_O^{xy}]_{TSL50}^{x100y100}$ where about 10% more customers face their second SHO than the FRH expected. This is the scenario discussed earlier, where ABMS showed a higher rate of initial sales than expected, from which it was reasoned that quicker or earlier SHO lead to more frequent replenishment of the focal item. There is also a much smaller rate of alternate sales in this scenario in the ABMS outcome; the focal customers simply cannot buy item Y in proportions that the FRH assumes they will.

The design of the experiment assumes that the mean daily demand of each item is equal, as is the inventory investment level and initial on-hand inventory, so that there should be no difference in retailer supply. In other words, at TSL 50% both items should run out of inventory in 5 days, since only that much inventory is ordered for the expected demand over 10 days. At day 5 if the focal item stocks out since all focal customers are willing to switch any time after the focal item is SHO all the focal customers check the shelf for the alternate item. The alternate item has stocked out at this time too, or due to random order arrival sequence a few focal customers are able to buy the alternate item, while the rest leave without sale. The retailer cannot take advantage of risk pooling

P_0^{xy} Tukey group FRH ABMS	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	L	F	C	L	I	G	L	L	L
	0	12.50	25	0	3.13	6.25	0	0.13	0.25
	0	14.13	29.84	0	4.75	12.56	0	0.05	0.12
Half	L	D	B	L	K	H	L	L	L
	0	12.50	25	0	3.13	6.25	0	0.13	0.25
	0	20.08	33.30	0	0.55	6.17	0	0.04	0.08
All	L	E	A	L	J	K	L	L	L
	0	12.50	25	0	3.13	6.25	0	0.13	0.25
	0	16.90	36.16	0	3.07	0.72	0	0.04	0.07
	None	Half	All	None	Half	All	None	Half	All

Focal customers: buy Y if X is out (Sub_{XY})

P_0^{xy} Difference (t stats)	Target Service Level								
	50%			75%			95%		
Buy X if Y is out (Sub_{YX})									
None	-	1.63	4.84	-	1.63	6.31	-	-0.08	-0.13
	-	(42.56)	(105.90)	-	(26.31)	(76.33)	-	(-24.38)	(-17.14)
Half	-	7.58	8.30	-	-2.58	-0.08	-	-0.08	-0.17
	-	(160.90)	(95.54)	-	(-71.38)	(-0.58)	-	(-22.26)	(-31.29)
All	-	4.40	11.16	-	-0.06	-5.53	-	-0.08	-0.18
	-	(76.51)	(141.80)	-	(-0.87)	(-190.0)	-	(-26.40)	(-26.58)
	None	Half	All	None	Half	All	None	Half	All

Focal customers: buy Y if X is out (Sub_{XY})

Table 2-11 Alternate item SHO rate for focal customers

because the inventory investment is too low to allow for an alternate purchase. If some focal customers are able to buy the alternate item, that item's reorder point is reached earlier because the rate of demand arrival has increased with the alternate demand of these focal customers. Since lead time is deterministic the earlier reorder point results in the earlier replenishment where newly arriving focal customers within the 5- to 10-day period are able to buy the newly replenished focal item. Alternate customers are also able to buy the newly replenished item at this time, before their own item is replenished. Over time the two items which initially have the same exact supply characteristics may be pushed further and further out of sync so that the composition of the customers demanding each item changes even further.

Although the question of whether replenishment orders are in or out of sync because of differing inventory levels is beyond the scope of this study, different levels of substitutability could be considered akin to different inventory investment levels and are considered in this study. With asymmetric substitutability the reorder point could be reached sooner for the item that customers are more willing to switch to than the item that has much lower alternate customer demand. In Table 2-11, the outcomes at TSL 75% and 50% each vary by scenario--there are 6 Tukey groups at each TSL, plus the non-switching group outcome L. The FRH assumed there would be 2 different mean outcomes, instead of 6. At TSL 50% there is no clear difference between the two scenarios where customer willingness to switch are equal and the two scenarios where more customers switch in one direction than another. All 4 scenarios at this inventory investment level show a stark difference between the FRH and ABM perspectives so any role of asymmetric substitutability cannot be distinguished.

Observing all of the outcomes for focal customers by TSL is necessary since it appears that the FRH is not capturing the outcomes emerging from interaction effects and the way it differs from ABMS depends on the TSL. The following 3 illustrations (Figure 2-4, 2-5, and 2-6) illustrate how the FRH has significantly over- or underestimated switching outcomes for all focal customers. In the illustrations there are circles and arrows. The "+" and "-" symbols inside the circles are to emphasize the direction of the arrows. The arrows go from where the FRH overestimates an AMBS outcome to where it underestimates an ABMS outcome. Whenever a scenario outcome is impossible, there is a circle with an X in it. For scenario outcomes where the FRH and ABM are not significantly different, there is no circle. Furthermore, the different levels of significant difference are identified by the circle outlines. Differences below 2% are illustrated with a tight dotted line and are negligible differences since even the non-switching scenarios yield different outcomes due to the continuous distribution assumption mentioned earlier. Differences above 2% have different patterns. Those below 5% have a dotted-dash pattern, below 10% but at least 5% difference is illustrated with a dashed line, and below 15% but at least 10% difference is a solid line. For a considerably large difference between the FRH and ABMS outcomes, the circle has a thick solid line (less than 20% difference, but at least 15%). Scenarios with no significant difference between FRH and ABMS outcomes are illustrated with a thin solid circle and the letters "ns."

When the inventory investment is very high and SHO infrequently occurs, the FRH is a useful rule of thumb for retailers. Figure 2-4 illustrates that overall, the FRH generally underestimates the proportion of focal customers able to buy their initial item. Even if all alternate customers are willing to buy the focal item (row 3 of Figure 2-4),

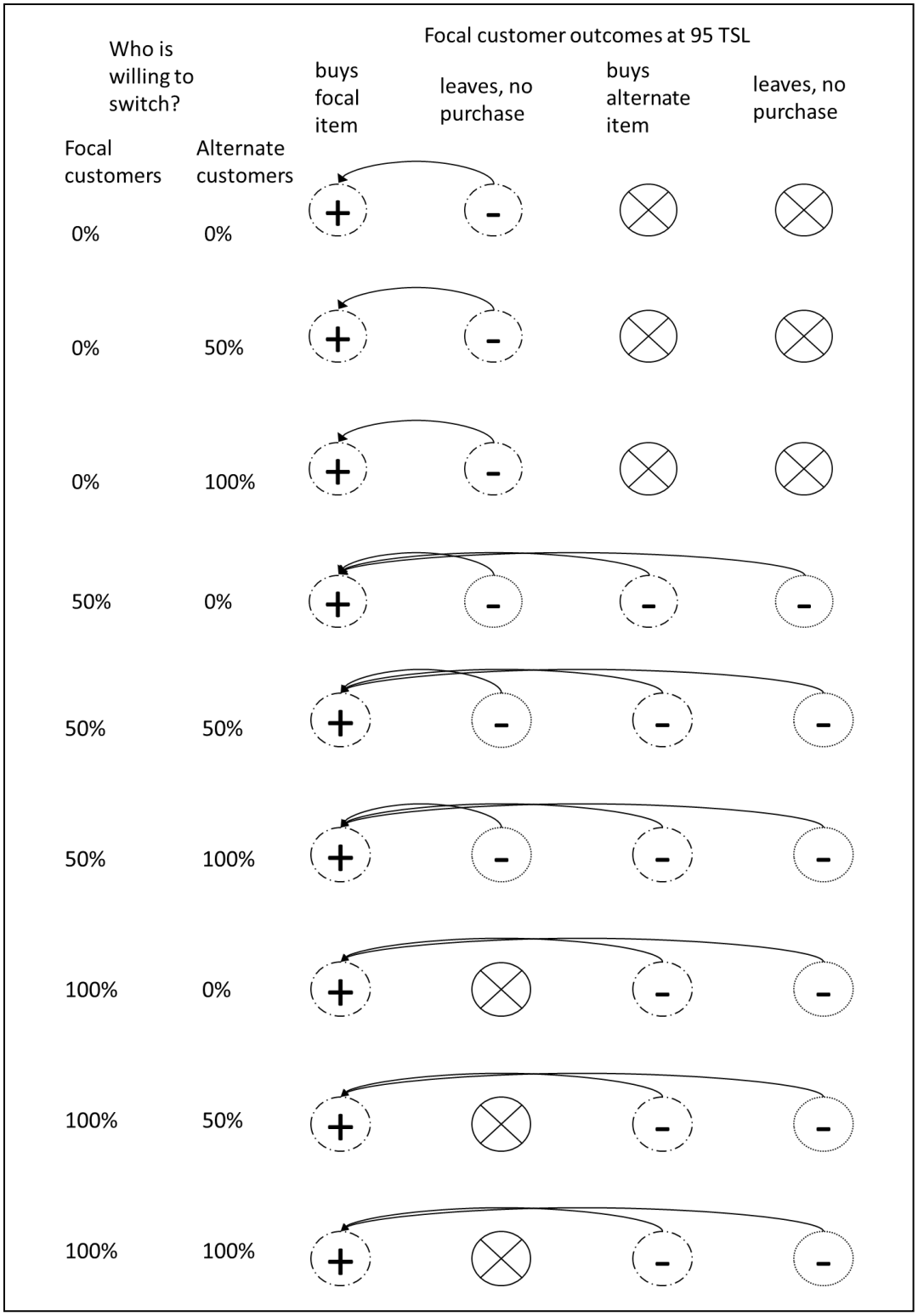


Figure 2-7 Difference between ABM and FRH at 95% TSL

the difference between the FRH and the ABMS is similar to the difference when no switching is possible (row 1). When the focal customer is willing to substitute, then the FRH overestimates how many focal customers will face SHO, as evidenced by the last two columns where the circles both have negative signs. Since the second column (leaves, no purchase) is also negative whenever there are any focal customers who are not willing to substitute, it follows the overall pattern at this TSL that the FRH underestimates the initial sales of the focal item, and overestimates the alternate sales and SHO rates. Alternate customers are still able to buy the focal item as a substitute, just not at FRH rates. For the retailer at this level of inventory investment, using the FRH is useful because the proportion of focal customers able to buy their initial item will be at least what they can expect by using the FRH. They need not be too concerned of customer willingness to substitute in either direction and can still take advantage of risk pooling to capture otherwise lost demand.

At the other extreme, when the inventory investment is very low and SHO frequently occurs, in using the FRH the retailer must consider the substitutability between items, but especially the alternate demand of the focal item. Figure 2-5 illustrates that the FRH underestimates initial sales rates, just like at 95% TSL, in all but two cases (rows 2 and 3). If there are any alternate customers willing to buy the focal item but no focal customers willing to buy the alternate item as a substitute, then the FRH overestimates how many initial customers will actually be able to buy the focal item. The FRH overlooks the interconnectedness of customer switching outcomes when it comes to the impact of increasing alternate customer willingness to switch (3rd step in last

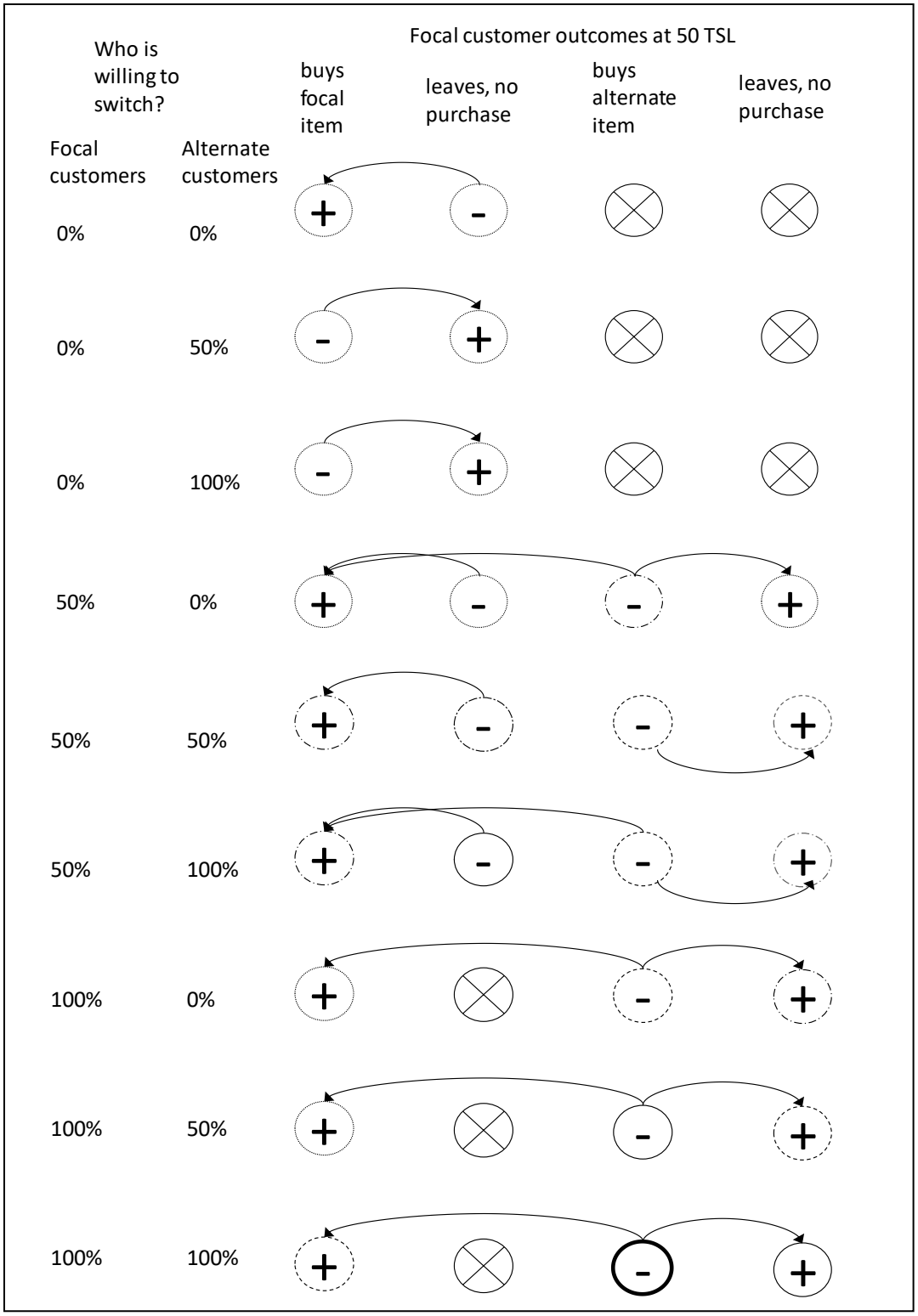


Figure 2-8 Difference between ABM and FRH at 50% TSL

column of Table 2-4). The FRH also overestimates how many focal customers will be able to buy the alternate item. A comparison of the bottom two thirds of this table with the table at 95% (the fourth row and below of each table), show a similar underestimation of buying the focal item and overestimation of buying the alternate item. In contrast to the 95% TSL, however, the FRH underestimates how many focal customers will face SHO twice and leave with no purchase even though willing to switch to the alternate item. With full substitutability in both directions (bottom row of Figure 2-5) the FRH overestimates the number of alternate purchases by over 15%. There simply is not enough stock for switching to occur as the FRH expects, so even if the alternate demand size is for the alternate item is larger (Table 2-4, step 4 of last column) it does not mean that sales too will increase (Table 2-4, step 5 of last column). Indeed, the proportion of that alternate demand size that faces SHO is larger than according to FRH. At this level of inventory investment overall, increasing alternate customer willingness to substitute leads to worse sales outcomes (since both initial and alternate sales of the focal customer are lower than expected), contrary to risk pooling theory.

Between the two extremes is when inventory investment is at a medium level and the observed pattern in differences between the FRH and ABMS outcomes is unique for scenarios with two-way substitution (Figure 2-6). In one-way substitution, the observed differences are similar to those at 50% TSL. When alternate customers are the only ones willing to switch, the FRH overlooks how alternate demand for the focal item reduces its supply for its initial customers. When focal customers are the only ones willing to switch, the same dynamic results in fewer focal customer alternate purchases than expected. For two-way

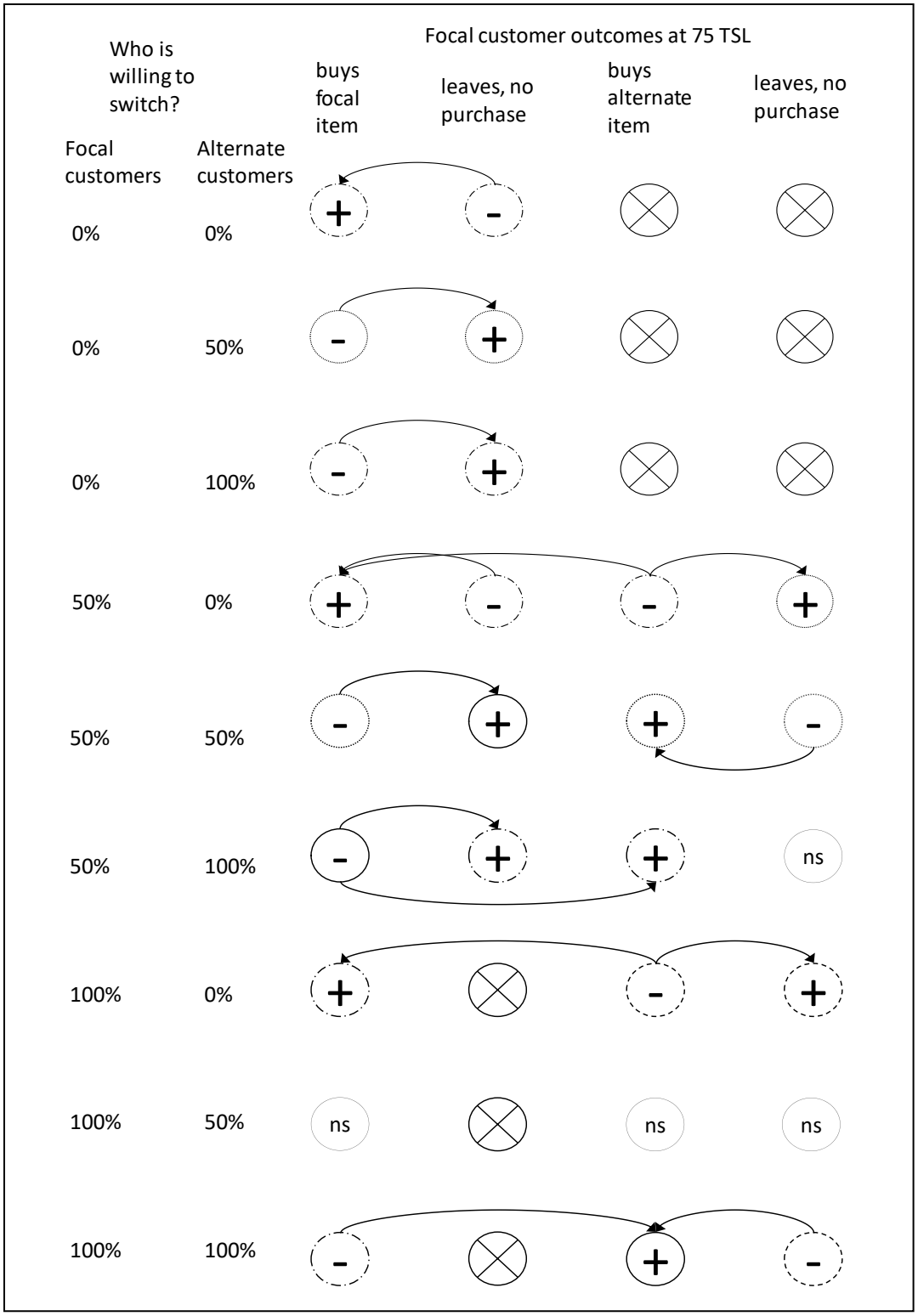


Figure 2-9 Difference between ABM and FRH at 75% TSL

substitution the observed pattern of differences between FRH and ABM are not observed at the other two TSL levels. At this medium-level of inventory investment, the FRH does not consider the combined increase of alternate sales of alternate item and decrease of initial sales of focal item (Table 2-4, steps 3 and 5 in last column, steps 6 and 2 in second-to-last column). Since there is a compound impact of increased substitution in both directions that the FRH does not account for, it overestimates initial sales of the focal item and underestimates alternate sales of the alternate item.

There is one scenario at this level of inventory investment where there is no statistically significant difference between the FRH and the ABMS outcomes. This occurs when all focal customers are willing to switch while half of the alternate customers will switch (Figure 2-6, second from bottom row). This is the scenario that was used for the example on calculating the theorized fill rates in the FRH and the actual fill rates from the ABMS. The mirror opposite of this scenario, where half of focal customers and all of alternate customers would switch, does have significant differences. This implies that an illustration similar to Figure 2-6, but focusing on alternate customer behavior instead would have a scenario where initial and alternate purchases would be sufficiently expected by the FRH when all of alternate customers and half of focal customers would switch. While looking at all possible outcomes for the focal customer is already complicated, the question remains of how to simultaneously focus on customer demand and on item supply at once (all 8 outcomes instead of 4 for the focal customer or for the focal item). This is more of a theoretical question as the practical applications of a scenario with no significant difference from the ABMS outcomes means the FRH is a

useful rule of thumb for the retailer, especially if the initial customers for one item are more "key" or focal to the retailer than the other.

The difference between focusing on items or focusing on customers is again highlighted when looking at the 4 switching outcomes of the focal customer to numerically see where differences in one outcome is absorbed by the others. This can be done by looking at the numerical differences between the two perspectives (the bottom chart in each of Tables 2-8 through 2-11). For the scenario where the TSL is 50% and half of both customers are willing to switch, the differences between perspectives are: $\Delta(P_X^x + P_X^{xy}) = +4.25$; $\Delta P_0^x = -2.09$; $\Delta P_Y^{xy} = -9.75$; $\Delta P_0^{xy} = +7.58$. The sum of differences is about 0.01% and this fraction is likely due to rounding error. Summing the two SHO differences as well as summing the fill rate differences shows that overall the focal customer stocked out about 5.5% more than the FRH estimated and sold less than expected to focal customers in about the same percentage. This scenario is an example of when customer switching hurts retailer performance. Raising the inventory investment for this "*x50y50*" level of two-way substitutability from TSL 50% to TSL 75% results in a 10% greater rate of substitute sales ($P_Y^{xy}]_{TSL75}^{x50y50} - P_Y^{xy}]_{TSL50}^{x50y50}$) while raising it to TSL 95% results in a lower substitution rate. But over all of the four outcomes the differences between the two views drops from 5.5% to 2.3% and then to 1.32 %, giving the impression that the FRH sufficiently captures customer switching as inventory investment is higher. However, it is capturing demand instead of customer behavior. The customer behavior of greater switching rates than expected is lost when aggregating the SHO and fill rates. Furthermore, these four outcomes do not really reflect how items are stocked and focus only on the outcomes customers reach by falling into different paths.

Looking at an item's fill rate as well as the customer's fill rate at the same time is necessary but not what this analysis accomplishes. That is the goal of the second part of this study.

2.5.3 Assessing the impact of customer switching on fill rates

The analysis of drivers to customer switching provided a close look on how the FRH and ABMS differ in terms of customer outcomes (of capturing switching) but it did not provide a look of item outcomes or overall sales and SHOs. Certainly, a combination of the 8 outcomes can allow for looking at switching with an item focus or with an overall focus, but those would all be individual analyses similar to that of the last section on customer outcomes. Additionally, none of the outcomes are at the aggregate level of fill and SHO rates which are used in practice and research. The close look was necessary to see where the FRH could overlook interaction effects even when inventory investment levels were equal and replenishment initially occurred at basically the same time.

2.5.3.1 Fill rate calculations

The fill rate calculations (Table 2-12) use almost identical notation as the customer switching outcomes, but use the letter *C* instead of *P* and D instead of μ. This is because the calculations take the count values from the ABMS instead of the rate outcomes from the first part of the study. The counts for demand are: focal customers who are not willing to substitute (D_x), focal customers willing to substitute (D_{xy}), alternate customers not willing to substitute (D_y) and alternate customers willing to substitute (D_{yx}). The calculations for item Y are in smaller font and italicized because

they are not included in the multivariate analysis. This is because the design experiment has every combination for X and for Y, just like in the switching outcomes tables, they would yield identical results. They are still included in the chart to show the full range of fill rate types and give a more complete overview of which outcomes are used where.

Fill rate	Description	Calculation
Item	The fraction of initial and alternate demand able to buy the focal item	For item X: $\frac{C_X^x + C_X^{xy} + C_X^{yx}}{D_x + D_{xy} + D_{yx}}$ *For item Y:* $\frac{C_Y^y + C_Y^{yx} + C_Y^{xy}}{D_y + D_{yx} + D_{xy}}$
Initial	The fraction of customers able to buy their initial item	For customer X: $\frac{C_X^x + C_X^{xy}}{D_x + D_{xy}}$ *For customer Y:* $\frac{C_Y^y + C_Y^{yx}}{D_y + D_{yx}}$
Customer	The fraction of customers able to buy their initial or alternate item	For customer X: $\frac{C_X^x + C_X^{xy} + C_Y^{xy}}{D_x + D_{xy}}$ *For customer Y:* $\frac{C_Y^y + C_Y^{yx} + C_X^{yx}}{D_y + D_{yx}}$
Category	The fraction of total demand able to buy any item	$\frac{C_X^x + C_X^{xy} + C_X^{yx} + C_Y^{xy} + C_Y^y + C_Y^{yx}}{D_x + D_{xy} + D_y + D_{yx}}$

Table 2-12 Fill rate calculations using ABMS-generated counts

2.5.3.2 Statistical analysis

This second part of the study uses Stata 13.1 (StataCorp, 2016) in analyzing the size effects of the customer switching phenomenon in terms of four different fill rates: item, initial, customer, and category. There are 2,430 observations from the simulation, 4 outcomes (dependent variables) and 4 experimental parameters (independent variables). As the null hypothesis is that there are no differences between the mean values of the four fill rates over the switching factors, the MANOVA provides a global test of whether they do vary. After the MANOVA separate univariate ANOVAs are listed for the main and interaction effects that are found to be significant in the MANOVA. The results are listed in Table 2-13, where CWS stands for "customer willingness to switch."

While there are four different test statistics for MANOVA (Wilk's Lamda, Hotelling's Trace, Roy's Largest Root, and Pillai's Trace), this study uses Pillai's trace. Pillai's trace can be used when group sizes are equal (and larger than 20 per group), as in this experiment and is the most robust of the statistics as it does not require equality of covariance matrices (Box's M test rejects the null assumption of homoscedasticity with (χ^2 (6) = 15,699.30, $p < 0.001$)). With Pillai's Trace, there is no need for an insignificant Box's Test to confirm homoscedasticity (Mertler & Rachel, 2005, p. 124). Roy's Largest Root has the most power of the 4 tests when there are highly correlated dependent variables (the fill rates are all comprised of a combination of the same customer switching outcomes with lowest correlation between initial and category fill rate at 0.6626 and highest correlation at 0.9926 between initial and item fill rates). As an upper bound test statistic Roy's Largest Root also showed significance (l_1=69,775, F (15,2414)

=11422.84, p<0.001), just as all four test statistics show the same results in terms of level of significance. Additionally, instead of Wilks' Lambda, which is the most often reported test statistic in MANOVA research (Mertler and Vannatta, 2005, p. 121), this study uses partial eta square (η_p^2) measures to see how big the relationship is between the fill rates and the customer switching drivers. Wilks' Lamda is an inverse criterion which means as the number gets smaller, the effect gets larger. Eta squared is obtained by (1-Wilk's Lamda) so that the η_p^2 value increases as the size effect increases.

Using Pillai's trace, all interaction and main effect variables are significant (p<0.001) except for SL_X * SL_Y, which are the item inventory investment levels (F(4, 2411)=1.67, p=0.1540, $\eta_p^2 = 0.0028$). No univariate analysis results are provided for this two-way interaction term in the ANOVA post hoc, which explains the corresponding incomplete row in the results Table 2-13. The univariate ANOVA for the category fill rate shows that every term included in the calculations has a significant effect on this aggregate term. The largest of the effects on category fill rates are equal in size ($\eta_p^2 =$ 0.30, p<0.001) and show the 60% of the variance of category fill rates can be explained by the inventory investment of both item X and Y. The next largest effect on category fill rates is the two-way interaction term of customer willingness to switch, SUB_{XY} * SUB_{YX} ($\eta_p^2 = 0.03$, p<0.001). The 3-way interaction of these two switching factors with an additional term for inventory investment (SUB_{XY} * SUB_{YX} * SL_i) also have the same effect size.

Switching drivers	Multivariate results	Item (X) fill rate	Customer fill rate	Initial fill rate	Category fill rate	Test statistics
Focal item TSL (SL_X)	0.87 4120.24*** 0.14	0.86 6890.49*** 0.74	0.88 2895.73*** 0.76	0.80 7816.20*** 0.76	0.25 1054.89*** 0.30	Pillai's trace F-ratio Partial η^2
Alternate item TSL (SL_Y)	0.84 3206.00*** 0.12	3.13x10-4 2.51 1.0x10-3	2.5x10-3 8.24*** 3.4x10-3	8.3x10-5 0.81 3.3x10-3	0.25 1055.92*** 0.30	P F η_P^2
Focal CWS (Sub_{XY})	0.18 130.69*** $2.6x10^{-4}$	3.2x10-5 2.51 1.1x10-4	4.9x10-4 1.60 6.6x10-4	3.2x10-4 3.11* 1.3x10-3	0.01 57.58*** 0.02	P F η_P^2
Alternate CWS (Sub_{YX})	0.23 183.85*** $7.0x10^{-4}$	0.02 141.68*** 0.06	0.04 140.37*** 0.05	0.03 278.51*** 0.10	0.01 55.42*** 0.02	P F η_P^2
$SLx*SL_Y$	0.00 1.67					P F
SL_X * Sub_{XY}	0.12 84.29*** 5.2x10-5	1.9x10-5 0.15 6.2x10-5	1.5x10-4 0.50 2.1x10-4	1.7x10-4 1.68 1.1x10-3	8.4x10-3 36.20*** 0.01	P F η_P^2
SL_X * Sub_{YX}	0.62 985.86*** 0.04	0.11 854.06*** 0.26	0.12 395.94*** 0.14	0.10 958.22*** 0.28	1.5x10-3 6.5*** 2.7x10-3	P F η_P^2
SL_Y * Sub_{XY}	0.50 605.47*** 0.02	7.2x10-5 0.58 2.4x10-4	0.18 588.83*** 0.20	1.7x10-4 1.68 7.0x10-4	1.4x10-3 5.88** 2.4x10-3	P F η_P^2
SL_Y * Sub_{YX}	0.17 120.33*** 2.1x10-4	0.01 87.84*** 0.04	0.03 93.93 0.04	0.02 183.34*** 0.07	8.1x10-3 34.77*** 0.01	P F η_P^2
Sub_{XY} * Sub_{YX}	0.11 78.24*** 3.7x10-5	5.9x10-6 0.47 2.0x10-4	0.02 61.52*** 0.02	0.03 270.09*** 0.10	0.02 52.61*** 0.03	P F η_P^2
SL_X * SL_Y * Sub_{XY}	0.38 372.42*** 5.3x10-3	4.3x10-5 0.34 1.4x10-4	0.11 363.81*** 0.13	1.4x10-4 1.34 5.6x10-4	8.4x10-4 3.63* 1.5x10-3	P F η_P^2
SL_X * SL_Y * Sub_{YX}	0.50 592.10*** 0.02	0.06 501.77*** 0.17	0.07 245.59*** 0.09	0.06 590.26*** 0.20	9.3x10-4 4.02** 1.7x10-3	P F η_P^2
SL_X * Sub_{XY} * Sub_{YX}	0.10 63.58*** 2.5x10-5	5.4x10-6 0.04 1.8x10-5	0.02 62.99*** 0.03	0.02 210.70*** 0.08	0.02 75.87*** 0.03	P F η_P^2
SL_Y * Sub_{XY} * Sub_{YX}	0.09 56.00*** 1.6x10-5	4.9x10-4 3.93** 1.6x10-3	0.02 50.35*** 0.02	0.02 166.23*** 0.06	0.02 75.75*** 0.03	P F η_P^2
SL_X * S_{LY} * Sub_{XY} * Sub_{YX}	0.06 40.40*** 3.2x10-6	2.5x10-4 2.02 8.4x10-3	0.01 41.31*** 0.02	0.01 121.61*** 0.05	0.01 5.43*** 0.02	P F η_P^2

Table 2-13 Multivariate and univariate test results

Since initial and item fill rates are defined by similar customer switching outcomes (as illustrated in Table 2-5, section 2.1.1.1 description of fill rates), their results are compared with one another. For item fill rate, 5 of the terms have a significant effect, while for initial fill rate there are 9. The two differ in terms of: focal customers willing to substitute (Sub_{XY}), two-way substitution interaction (Sub_{XY} * Sub_{YX}), focal item inventory investment and both customers' willingness to switch (Sub_{XY} * Sub_{YX} * SL_X), and the four-way interaction term of all of the customer switching drivers (Sub_{XY} * Sub_{YX} * SL_X * SL_Y). In each one of these cases the initial fill rate is significantly affected while item fill rate is not. It would be wrong to infer that customer switching does not impact item fill rate as the effects of alternate customers willing to substitute are significant as a main effect (Sub_{YX} η_p^2= 0.06) with even greater interaction effects (SL_X * Sub_{YX} η_p^2= 0.26 and SL_Y * Sub_{YX} η_p^2= 0.04 as well as SL_X * SL_Y * Sub_{YX} η_p^2 = 0.17). Together these effects can explain more than half of the variance (0.06+0.26+0.04+0.17=0.53) of the item fill rate and the SL_X * Sub_{YX} interaction effect on its own is considered large since it is above 0.25 (University of Cambridge, 2018). Comparing the two rates, the item fill rate is affected by alternate customers willing to switch to the focal item when their initial preference is SHO, and the initial fill rate is affected by not only alternate customers but also focal customers willing to substitute.

The customer fill rate, which is a combination of purchases of the focal and alternate items by the focal customer, differs in the significance of effects as compared to the item and initial fill rates, by one main switching driver. The inventory investment of the alternate item has a significant main effect (SL_Y η_p^2 = 0.0034) and interaction effects

(SL_Y * Sub_{XY} η_p^2 = 0.20 and SL_Y * Sub_{XY} * SL_X η_p^2 = 0.20) on the customer fill rate, while the effect on item and initial fill rate are not significant. This makes sense since the focal customer has to be willing to switch to the alternate item which has to be in stock for the substitute purchase to occur. Significant effects in the following illustrate that customer fill rates are closer to initial fill rates than to item fill rates: the two-way interaction of customer willingness to switch, and the three-way interaction of these two substitutabilities and the inventory investment of the focal item. It is not surprising that the customer fill rate is impacted in a way that is more similar to the initial fill rate than the item fill rate because the numerator of the customer fill rate also includes the initial sales which is the numerator of the initial fill rate. Furthermore, since their denominators are the same (demand for the focal item by focal customers) the customer fill rate should always be larger than the initial fill rate because it includes alternate sales in the numerator as well.

Finally, looking at the univariate results by the factors instead of by the dependent variables can provide additional insight. For the inventory investment of the focal item, the smallest impact appears to be for the category fill rate (SL_X η_p^2 = 0.30) while it can explain most of the variance in item (η_p^2 = 0.74) and initial fill rates (η_p^2 = 0.76). The alternate item's inventory investment explains 30% of the category fill rate as well, but its effect on item and initial fill rates are not significant, while its effect on customer fill rates is significant but negligible (SL_Y η_p^2 = 0.0034). Interestingly the main effects of customer switching, even if significant, are negligible compared to their 2- and 3-way interaction effects. For example, focal customer willingness to substitute can, at most,

explain 2% of the variance of the category fill rate, but it can explain 20% of the variance of customer fill rate when considered in combination with the inventory investment of the alternate item (SL_Y * Sub_{XY} η_p^2 = 0.20). Similarly, alternate customer willingness to substitute can significantly explain 2-10% of the variation of all the fill rates, but its impact increases to 26-28% when combined with focal TSL and up to 17-20% when combined with both TSLs on item and initial fill rates. Though it was theorized that the impact of inventory investment (TSLs) would be larger for item fill rate than it is for the customer fill rate, the largest effect size of η_P^2 =0.76 is for the customer fill rate while η_P^2 =0.74 is for the item fill rate. Similarly, the customer willingness to switch has a greater impact on item fill rate (η_P^2 =0.06) than on customer fill rate (η_P^2 =0.05).

2.6 Conclusions

The first part of this study examines customer switching by comparing the outcomes of the phenomenon through FRH and ABM frameworks. The FRH is simply the conditional probabilities of the outcomes, assuming independent initial demand streams for each item. The ABM uses a simulation (ABMS) to recreate the store environment without making any assumptions at all about the outcomes and also assuming independent initial demand streams for each item. Comparing and contrasting the outcomes from the FRH and ABM perspectives shows that while the FRH on its own may not be able to completely account for the customer switching phenomenon, it can be a reasonable guide whenever its sales outcomes are a lower bound. A lower bound is

whenever the sales outcomes from the ABM are at least as much as what the FRH predicts they will be.

FRH versus ABMS	Focal item sales	
	To focal customers ($P_X^x + P_X^{xy}$)	To alternate customers (P_X^{yx})
Sold significantly more	19*	3
No significant difference	1	10*
Sold significantly less	7	14

Table 2-14 Overview of FRH and ABM differences

A tally of the outcomes is in Table 2-14, where the 27 scenarios (columns) of this part of the study are split into whether the ABMS outcomes show significantly greater sales or not, when compared to the FRH. If sales through ABMS are more than FRH probabilities, or there is no significant difference, then using such a simple rule of thumb (FRH) is useful and warranted. Scenarios where no customers are willing to substitute are marked with an asterisk (*) in the table. For focal customer purchases of the focal item, the FRH is useful 20 out of 27 or 74% of the time. For focal item sales to alternate customers, the FRH is useful only 13 out of 27 or 48% of the time. The same pattern can be seen for the alternate customer, where the FRH is more useful for initial sales than it is for alternate sales. Instead of customers, if the retailer is focusing on items, then the performance of the FRH is somewhere between its views on outcomes for initial and alternate customers. For either item, 33 times out of 54 scenarios or 61% of the time the FRH's framework is useful for approximating sales outcomes for that item.

When does the FRH overlook customer switching outcomes? The FRH overestimates how many initial customers will buy their item when inventory investment of both products is not high (SL_I = 50% or 75%) and alternate customers are willing to buy that item (Sub_{IR} = 50% or 100%). In other words, if there is one-way substitution towards the focal item, the FRH overlooks the possibility that focal customers may not be able to buy the item because of alternate customer purchases. This is to be expected from the top-down view of the FRH where substitution takes place on the condition that all initial demand is met and then there is excess inventory for subsequent alternate demand. The FRH also overlooks the interconnectedness of customer switching outcomes (Table 2-4) with two-way substitution at the medium inventory investment level. Focal customers are unable to buy the focal item at FRH rates because: (1) alternate customer willingness to switch takes away units of the focal item from initial customers, (2) focal customer willingness to switch takes away alternate item units so that a greater number of alternate customers switch to the focal item, decreasing supply for initial customers. The FRH overlooks this combined decrease in initial sales only at the medium level of inventory.

How does customer switching impact a store's customer service levels? It depends on the type of fill rate the retailer chooses to measure as the fraction of demand its inventory investment is capable of meeting. If looking at the overall category fill rate, all of the drivers of customer switching—both main and interaction effects—have a significant effect on outcomes. When setting target service levels goals on one item, the TSL of the other item and both directions of willingness to switch should be considered in tandem. In contrast the (unit) fill rate for a focal item mainly depends on its own level

of inventory investment and what proportion of alternate customers are willing to buy it if they are faced with SHO. If the retailer has a high-priority customer, initial or customer fill rates would be the appropriate performance measure. For both measures, the largest impact is simply the inventory investment (TSL). For the customer fill rate, the retailer needs to consider the TSL of the alternate item in relation to what proportion of that item's initial customers will switch to the focal item. For the item fill rate, the retailer must consider the TSL of the focal item in relation to what proportion of alternate customers may switch to the focal item. Customer switching impacts the different fill rate measures through different drivers of the phenomenon.

- A fill rate heuristic (FRH) using the 4 drivers of stockout-based substitution can sufficiently capture switching outcomes most of the time.
- Retailers should use the FRH with caution when there is a medium-level of inventory investment in both items and two-way substitutability.
- None of the 8 outcomes in the FRH on their own are the sales or SHO rates used in research or practice.
- There are 4 fill rate measures which are a combination of customer switching outcomes: initial, customer, item, and category fill rates.
- Category fill rate is the only measure that is significantly linked to all drivers of switching and their interactions.

Table 2-15 Summary of study findings

To summarize (Table 2-15), stockout-based switching's relationship with performance fill rate is complicated. There are four drivers to the customer switching phenomenon between two items: the inventory investment of each item and their customers' willingness to switch to the alternate item. The impact of each driver on performance fill rate depends on which fill rate measure is used. Category fill rate is the only measure that is significantly linked to all drivers and their interaction. Focal customer service levels can be measured by initial, item or customer fill rates. Changes in inventory investment of the focal item has the greatest effect on these fill rates, whereas

the inventory investment of the alternate item only has a direct effect on customer fill rates. All four fill rate measures are a combination of various customer switching outcomes.

Whether a sales or SHO outcome improves or worsens in performance with substitutability depends on the level of inventory investment of the items. With high inventory investment, increased substitutability is linked with increases sales and decreased SHO outcomes, supporting risk pooling theory. With low inventory investment, increased substitutability increases SHO outcomes and decreases sales rates, contrary to risk pooling theory, because customers face SHO of the alternate item as well. At the medium level of inventory, increasing two-way substitutability is linked to greater substitute purchases, because initial and alternate customers purchase an item in random order, contrary to analytic models which assume alternate customers buy an item only after all initial customers have attempted to buy it.

2.6.1 Limitations and future research

This study is just an initial look at the relationship between inventory investment and performance fill rate when customer switching is involved. Future research can lead in four different and possibly overlapping directions: (1) performance fill rate measures, (2) product availability (SHOs), (3) customer characteristics, (4) a deeper understanding of the customer switching phenomenon. Additionally, relaxing any of the experimental design assumptions could link this stream of research to that of product variety, operational efficiency, inventory record inaccuracy, inventory review method, backroom

effect studies, lead-time issues, attribute-based substitution, and so on. Such changes in design assumptions would also improve generalizability.

This work can be used to build on performance fill rate studies in two main ways. A taxonomy on the different fill rate measures used in practice and research across various streams of research is paramount. The taxonomy would lead to the first way research could build on this study in terms of fill rates. Different measures of performance fill rate can be incorporated into a study such as order fill rate or the item fill rate measures used more generally in practice. Fill rates could be compared against one another in terms of whether they were generated using ABM, regression, or analytical models on stockouts, product availability, or on customer switching. The comparison could be made by looking at the relative size of each fill rate, whereas here only switching outcomes were compared in size and fill rates were studied in terms of size effects of switching drivers. The second way of focusing on fill rates is expanding on this study's design. Here there are only three levels each of two supply and two demand factors. A study that looks at smaller intervals between target service levels or between willingness to switch would allow for more descriptive modelling on the impact of these factors on fill rates. This deeper study would also allow for more rigorous statistical methods beyond those used here. Perhaps a cut off can be found for when SHO events occur "too often" or "too rarely" and what other antecedents to customer switching make it "just right" for maximum risk pooling benefits of increased fill rates.

Future research can also focus on product availability. In this study SHO was merely a precondition to the phenomenon of study. Instead it could be the focus of a

study by looking at various SHO attributes. Here only product availability was considered in terms of fill rates. SHOs were only two of four possible switching outcomes for a customer in a two-item product category. Those SHO outcomes were unit-based rates, which measured the impact of the SHO in terms of number of units of sales lost. Future research can also look at the duration of a SHO either in terms of absolute number of days or as a proportion of the length of time the inventory is held for. For example, in this study, the order size was based on providing a supply for 10-day demand. Studies on SHO can see how much of that 10-day period actually has on-hand inventory. Similarly, studies on SHO frequency or using multiple attributes in a single study could better link the impact of SHO on the customer switching phenomenon or on performance fill rates.

Another direction for future studies is with a focus on customers. In this study customers over time are not identified as the same individual; repeat visits are not included. None of the outcomes for the customer after the initial visit to the store are considered in this study. In contrast, the fill rates in this study were generated using the outcomes over the entire simulation period. With customer learning the fill rates would also change because the willingness to switch or even the average daily demand may change. Such studies would follow the stream of research by Wu et al. (2013). Another study focused on customers could be one where their arrival to the store varies by time of day and day of week. Trauzettel (2014) looks at how fill rates are affected by demand fluctuating throughout the order-up-to periodic inventory review system. If the arrival of customers relative to order replenishment affects performance fill rates, it may also be a

demand characteristic that drives customer switching in addition to willingness to switch and the supply driver of TSL.

Last of all, a deeper look into the customer switching phenomenon is warranted. Building directly on this study, additional measures of inventory investment would be helpful in defining the supply driver of switching. Overall and average measures like the total units ordered from the supplier, the remaining units left on-hand at the end of the experiment or average on-hand inventory could better tease out the switching phenomenon's ability to pool risk. In this study, the TSL assumed the role of inventory investment, because more safety stock is held at higher TSLs. However, a retailer could have two items at the same TSL, but one could have an order size for 3 weeks while the other has an order quantity for 3 days' worth of demand. This leads to another two supply characteristics to be considered: synced supply and replenishment frequency. Items that are held on hand in smaller quantities are replenished more frequently which leads to more possible instances of SHO occurrence, especially if there is stochastic lead time. Synced supply refers to whether the items are replenished jointly or at different times with non-overlapping lead time periods. It is possible for customer switching to have a greater effect on fill rates under more complicated supply characteristics. Whether such supply characteristics make customer switching result in higher or lower fill rates, or SHOs, would be the focus of such studies.

The switching phenomenon could also be further studied in terms of demand characteristics. This study assumes normally distributed demand. How does this demand assumption affect fill rate outcomes compared to Poisson, negative binomial, or other

distributions? How about joint distributions for substitutable items? Is there an agent-based perspective on jointly-distributed demand? What about non-symmetric demand? In this study the items had the same average daily demand. What if one had twice the demand of the other? What would the impact of willingness to switch be on fill rates in any of these cases? Finally, the FRH was a parsimonious model consisting of conditional probabilities. How does it compare to more complicated models of demand? How does the ABM compare to other analytic models beyond FRH? Perhaps more sophisticated approximations are nearer to the outcomes found in ABM. If the components of effective demand could be observed for empirical work, how does the ABM compare to an empirical predictive model? The FRH is not significantly different than the ABM at a medium level of inventory investment when the focal customers were all fully willing to switch while only half alternate customers would switch. However, the FRH did overlook the interconnectedness of all of the outcomes, where sales and lost sales outcomes shift in different ways. Perhaps more sophisticated heuristics or analytical models could account for these as well.

Chapter 3 Can instore logistics postponement improve product availability?

Unlike Chapter 2, this chapter incorporates the store's shelf capacity in its efforts of dealing with SHO and introduces the concept of instore logistics postponement instead of using product substitution as a risk pooling tool. Instore logistics postponement is based on the use of two different inventory types in store. Stores grant access to customers at shelf spaces where items have a dedicated location in the front room of the store. Customers do not have access to items in the backroom space, which the store uses as a shared space area for items that have been delivered to the store. Instead of customer service and inventory investment, this chapter considers operational efficiency (labor) and inventory investment (shelf space). The study presents a framework for instore logistics postponement where the shared space can be used to store safety stock and cycle stock depending on the strategy and emphasis of the store as well as item-based demand characteristics. Instore logistics postponement is found to be linked to decreased category SHO duration and frequency, but further research is needed to test the theory on items that can be classified into the different areas of the presented framework.

3.1 Introduction

As a consumer it is not uncommon to reach a store shelf to see one item fully stocked while a similar adjacent item of a different brand, flavor or size is completely stocked out. Such a stark difference in product availability of similar items points to item-based supply and demand drivers within the store, product category or brand. These item-based differences contribute to overall category and store performance (availability)

measures. Store availability globally is widely cited as having an 8% stockout rate (Gruen, Corsten, & Bharadwaj, 2002, p. 13), meaning that 8% of a store's assortment is not available at its dedicated shelf space at any given time. Studies of how consumers respond to this temporary product unavailability find that the stocked out item may lose sales 54% of the time (with an additional 15% of consumers delaying purchase) and that the store may lose sales 40% of time (Corsten & Gruen, 2003). Annually retailers lose over $600 billion dollars in revenue due to stockouts (Buzek, 2015). Furthermore, since the average global stockout rate has not decreased from the time scholars began studying product availability in the 1960s (Aastrup & Kotzab, 2010) the need for further research into various item-based supply and demand drivers of retail product availability (Ettouzani, Yates, & Mena, 2012) is still relevant.

Postponement and speculation (Alderson, 1950; Bucklin, 1965) theories provide a system-wide perspective (Yang, Yang, & Wijngaard, 2007) (Garica-Dastugue & Lambert, 2007) on improving product availability by changes to distribution channels, but have seldom been studied as methods for retail firms (Jafari, Nyberg, & Hilletofth, 2016) or completely within a single firm. Standard postponement-speculation (P/S) theory involves a "decoupling point" (Maister, 1976) which is where a supply chain member receives final customer demand information from downstream partners and adjusts the physical processing of goods (in terms of its form, identity or location) to better meet that final downstream demand. Demand is better met because a system which supplies goods "to inventory" can switch to a system that supplies goods "to order," at this decoupling point, without substantially affecting any other member in the chain, beyond information sharing activities. Product availability is the outcome of the ability to

meet demand with appropriate supply yet few studies have tested the postponement-product availability relationship empirically (Boone, Craighead, & Hanna, 2007), much less in retail. Retail studies so far have been limited to the online retailing context (Kim, 2014). Rabinovich and Evers (2003) assess postponement's effect on inventory management performance which is measured as the percentage of a manufacturer's production built to inventory. Building "to inventory" or speculatively processing and holding goods makes goods more available (increasing product availability) and they find a positive relationship between postponement and inventory management performance. Online retail is about 9% of total consumer retail spending and 71% of American shoppers prefer to buy items which are also offered online from brick-and-mortar stores instead (Stanley, 2016).

Brick-and-mortar retailers face issues, including product availability, which other members of the supply chain do not (Defee, Randall, & Gibson, 2009). Retail customers approach a store shelf themselves to look for an item at the shelf location for which the store has dedicated a certain capacity of space. Regardless of whether the item is in the store or not, it is unavailable to the customer and referred to as a shelf stockout (shelf-out or SHO) if that dedicated store shelf space is empty. If the item is in the store but at some shared inventory space instead of its dedicated shelf space (an instore replenishment stockout or IRO), it is similar to work-in-process inventory for members higher up in the supply chain in that the item is almost available but needs further processing of some kind. In the store's case the further processing refers to changing the item's location. The inventory stockouts that are more akin to unavailability for upper supply chain members is when the SHO item is also not in the store (a store stockout or STO). It is therefore

necessary to consider a finished good within a retail store as not being in its final state until it is at its customer-accessible dedicated shelf space. Any IRO type of SHO for a store item means that item is in shared space and not available when the customer arrives at its dedicated space.

The distinction between shared versus dedicated space and their purposeful combined use to improve product availability through postponement is seldom considered in this manner in research or in industry (Hubner & Schaal, 2017). In industry over 90% of the time (Gruen & Corsten, 2008), shelf space is dedicated to an item by the size of the shipment of that item to the store, while shared inventory space is often in the backroom of the store and generally seen as a location for "inventory overflow," which is the unintended phenomenon of items delivered to the store not fitting on dedicated shelf space (Eroglu, Williams, & Waller, 2013). Limited shelf space for a product category is often not incorporated into software retailers most often use (Hubner & Kuhn, 2012) for making space allocation decisions so that there is no consideration of how the dedicated space for one item may limit the availability of another. Assortment research (Shin, Park, Lee, & W.C., 2015) does consider this relationship between dedicated space of items to decide which products to carry on store shelves, but product availability in this stream is considered only in relation to customer reactions to stockout (Honhon, Johnalagedda, & Pan, 2012), focusing on which items in the assortment to carry as a buffer to pick up customer demand that originally preferred a stocked out item. Some studies acknowledge shared space exists but posit it does not directly relate to managing dedicated shelf space: "...units in the backroom are not immediately available to customers. Hence, a backroom acts like a warehouse" (Cachon, 2001). Similarly, work

which includes backroom and front-room (cash registers) store operations are in the context of a queuing or personnel management problem (Terekhov & Beck, 2009) and not with regard to product availability outcomes. A case study (Hubner & Schaal, 2017) of a retailer using both dedicated shelf space and customer inaccessible shared space develops an analytic model of the costs involved with replenishing within a store to be able to optimize use of both types of space. Their study, however, does not consider this type of postponement's effects on product availability, nor do they frame the use of both spaces within the store through the theoretical lens of postponement. Product availability is the focus in the backroom effect (Waller, Heintz Tangari, & Williams, 2008) which suggests that "when retailers carry inventory in two locations" (Eroglu, Williams, & Waller, 2011) items that are in the backroom (shared space) remain there indefinitely, decreasing product availability through increased IRO. Instead of simply being an unintentional but necessary factor for IRO to occur as stated in extant work, this study postulates and empirically tests whether purposeful use of both types of retail space (termed "instore logistics postponement") can improve product availability through decreasing SHO, without having to differentiate between IRO and STO.

This work contributes to product availability and postponement theory in four ways. First, it answers the call for more postponement research in retail (Jafari, Nyberg, & Hilletofth, 2016) by theorizing that postponement activity can take place *within* a retail store and develops a new measure for logistics postponement based on concepts of shared and dedicated shelf space. This postponement activity is linked to demand-related antecedents and product availability outcomes. Product availability is measured at various dimensions simultaneously, as encouraged by other scholars (Gruen & Corsten,

2008). Multiple measures are possible because SHO events are captured in their entirety, from the moment the item's dedicated shelf space is empty until it is replenished. The shelf liner technology that captures these SHO events enables sales data (demand drivers) to be incorporated into the model since they are no longer needed to estimate when SHO may have occurred (Grubor & Milicevic, 2015). The next section develops the instore logistics postponement theory and hypotheses, and is followed by methodology, results and discussion.

3.2 Theory and Hypotheses

Postponement (Alderson, 1950) and speculation (Bucklin, 1965) theories aim to provide firms with methods to better match their supply to uncertain customer demand. Demand uncertainty according to Alderson, can be reduced by delaying firm activities on the form, identity or inventory location of an item to as late a time as possible. This time postponement (Zinn & Bowersox, 1988), allows for the firm or supply chain to gather more demand information and make less costly decisions about which products to supply when. The location of this supply can also be delayed (its forward movement to the customer (Oeser, 2015)), referred to as place postponement (van Hoek, 1998), by Bucklin's speculation addendum to postponement theory. Bucklin puts forward (1965, p. 31) the concept of creating intermediate inventories which form "whenever its additional costs are more than offset by net savings in postponement to the buyer and the seller." In other words firms may choose to reduce demand uncertainty risks by delaying the commitment of goods to a specific downstream location until there is more certainty the customer demands it there, or by bringing together the demand of multiple locations at

some new inventory point to achieve lower demand variability overall through risk pooling. A review by Boone et al. (2007) mentions prior research findings of excess inventory reduction by postponing or changing the order of steps in a manufacturing process, reducing shipping costs by delaying shipment of a product until it is purchased, and increasing operational efficiency in postponing final assembly of goods.

Building on postponement-speculation theory, logistics postponement focuses on manipulating the distribution channel by adjusting the location or time of availability of final products or finished goods only (Yang, Burns, & Backhouse, 2004). In logistics postponement finished goods are made to inventory and are either kept in centralized inventories (place postponement) or directly distributed to the end of the supply chain (time postponement) (Twede, Clarke, & Tait, 2000, p. 106). A review of distribution network design (Mangiaracina, Song, & Perego, 2015) describes prior research findings that the following antecedents increase the benefits and likelihood of using place postponement: greater product variety, higher demand volatility, seasonality, greater total demand levels of each item, a pull (make to order) system, short order cycle time, and greater product value. In a taxonomy of postponement types, Kim (2014) categorizes time postponement as a type of business-function level postponement (as opposed to product- or process-level) where a downstream supply chain member shifts the risk of ownership to vendors and cites five studies in this area. All of the studies are within the internet retailer context, where customers "arrive" to stores that are online and do not require the shelf presence or availability as traditional brick-and-mortar stores. Regardless of whether it refers to time or place decisions (or both), logistics

postponement deals with distribution decisions between downstream supply chain members (vendors and retailers) of final finished goods.

This study presents *instore logistics postponement* as the purposeful use of both shared and dedicated store space, as an entirely intrafirm approach creating an additional tier of the distribution channel or creating intermediate inventory within the store. A store decides on the degree of inventory investment for each item, which will carry the store's supply needs until the next order cycle or store delivery. This inventory investment can be placed entirely on store shelves so that it is fully available to arriving customers. The same amount of inventory investment can instead be stored in both shared and dedicated space. Store shelf space, where items are given a dedicated capacity, is a finite store resource and part of a store's long-term strategic decision-making process. Normally a store can expand its shelf capacity for a product category by expanding the store footprint (overall store square footage). Physically expanding owned retail buildings or waiting for the renewal date of a lease and switching to a larger retail space or smaller space (Pomerantz, 2014) (Staples CEO Outlines Shift Away from Store Footprint, 2013), is normally the only approach to changing store size. Without expanding the store footprint, it is possible to increase a product category's dedicated space by decreasing the dedicated space of one or more other product categories. Unless there is excess dedicated space for other product categories, reducing alternate category size may create new issues with item availability and revenue for the store overall. With instore logistics postponement, however, the store can delay the forward movement of some of that inventory investment to the customer by storing it in shared store space (like a backroom). Instore replenishment, sometimes referred to as indirect replenishment

(Hubner & Schaal, 2017) or second replenishment (van Donselaar & Broekmeulen, 2008), from shared space to dedicated space would effectively expand the store's space for that postponed item. This concept of shared retail space meets Bucklin's time postponement requirement that the intermediate inventory location is not simply temporarily storing an item while it is being distributed but that it faces the risk of not being sold (Waller, Heintz Tangari, & Williams, 2008). By using shared space as an intermediate inventory location, the store may be able to better manage a variety of products with different item-based attributes which may affect product availability. All hypotheses presented below are illustrated in Figure 3-4 at the end of this section.

3.2.1 The direct effect of systemic drivers on product availability

In this study, item-based attributes having to do with the various aspects of product demand are referred to as "systemic drivers" of retail on-shelf availability (Moussaoui, et al., 2016). Moussaoui et al. (2016) state that systemic drivers are exogenous to the store and consist of: the demand of an item over time (autocorrelation), the rate of sales within a given time period (velocity), how much that velocity deviates from the mean (unpredictability), the number of unique items in a product category (assortment depth) and the distribution channel structure (network design). Autocorrelation, velocity and unpredictability are all item-based attributes within this "systemic drivers" definition, and are the only dimensions used for developing the measure of systemic drivers in this study. Operations literature could argue that these item-based attributes of demand are not exogenous to the store or are insufficient measures of demand. Demand reshape (Eynan & Fouque, 2003) for example, alters a

customer's initial preferences with store decisions including the amount of inventory to display on store shelves (billboard and scarcity effects as in Xue et al. (2017)) and these measures of demand are only realized sales that can be censored by product availability or inflated by stockouts causing customers to switch to available items (the inventory effect (Borin, Farris, & Freeland, 1994)). This study, however, categorizes these dedicated-space related factors as supply drivers and not demand drivers. Similarly, assortment depth and network design, explained in further detail in the methodology section, are supply-related drivers and are not included as systemic (demand) drivers for this study.

Prior studies of systemic drivers' effects on on-shelf availability have conflicting findings (Moussaoui et al., 2016) and either the measures of stockout or that of demand may be at the source of this inconsistency. Moussaoui et al. (2016) outline that scholars have argued that items with higher (lower) velocity are more (less) closely tracked and replenished on store shelves leading to greater (less) availability. On the other hand, higher (lower) velocity items are more quickly (slowly) depleted from dedicated shelf space and may be more (less) likely to SHO. Additionally, higher velocity items can also have a greater level of unpredictability than lower velocity items, with higher unpredictability making it difficult to forecast or manage such items. Using both measures of demand (velocity and unpredictability) as a general systemic (demand) driver may better encompass these opposing forces.

Moreover, the measures of SHO used from one study to another may also have opposing characteristics (Gruen & Corsten, 2008). For example, SHO frequency could

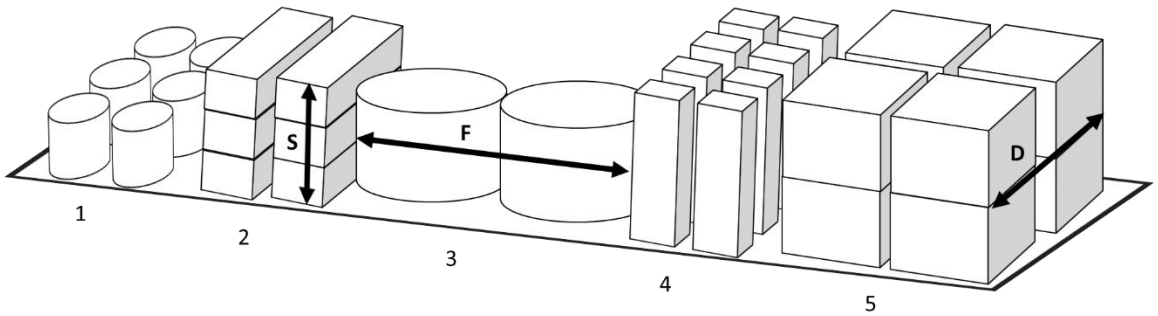
increase if partial replenishments from shared space are made for a highly unpredictable item. This suggests a positive relationship between demand and SHO. A negative relationship would be implied by the decreasing duration of SHO that occurs from partial replenishments which serves to “break up” the longer duration of waiting for a full replenishment from a store delivery. Even setting replenishments aside (as it is a supply driver to SHO and not a demand driver), as the frequency of SHO of an item increases over a period of time, the average SHO duration after a certain frequency may decrease as SHO events happen even more often, because there is a set period of time during which all of the SHO can occur. As that set period of time is marked by even greater frequency of SHO, the average SHO duration should decrease. Such a nonlinear relationship between SHO attributes may also contribute to conflicting findings between demand drivers and SHO. Accounting for the relationship between both attributes and using a multi-dimensional measure of demand may provide a more encompassing view. A more encompassing view of availability would also be attained by including supply drivers into the same model as systemic (demand) drivers. However, given that the supply of an item stays constant, there is a common consensus that increasing demand will reach a point where it exceeds the supply, leading to:

Hypothesis (H3.1): Systemic (demand) drivers are negatively related to availability.

3.2.2 Systemic drivers and instore logistics postponement

Another item-based attribute is a product's dimensions which play a pivotal role in how much shelf space must be dedicated to each item, as in Figure 3-1. The figure presents five items (1, 2, 3, 4, 5) which differ in product dimensions but each have 2 units side by side (facings) on the same shelf. Facings are normally used in retail shelf availability studies with the assumption that items in the study are of equivalent dimensions, even if few product categories meet this requirement in reality. When items are of different dimensions, their stacking (number of units placed on top of each other) and depth (number of units placed behind each other) may vary. Since shelf space is a limited resource for a retailer, the dedicated shelf space for each item should take into consideration all dimensions. If the shelf in Figure 3-1 consisted of items of uniform dimensions equivalent to item 1, then the shelf would be able to hold items 14 facings wide. This could be 14 facings of the same exact item, or 14 separate items, or any combination in between for a total shelf capacity of 42 units of items with product dimensions equivalent to those of item 1. If the items were all of the same dimensions as item 3, the shelf would only be able to hold 10 facings and the shelf capacity overall would also be 10 units. Indeed the total shelf capacity in terms of units could vary from 10 to 60 (the example in the figure has 30 units total) depending on what capacity is dedicated for each item and each item's dimensions. Just as there is a maximum dedicated shelf space, there is a minimum for each item. If an item is offered in a store's assortment at all, there must be at least one facing of that item on dedicated shelf space. Dedicated shelf space units differ even in this minimal single-facing offering: in Figure 3-1, item 1 would only have 3 units on the shelf, whereas item 3 would have only one

unit and would take up the capacity of 6 units of item 1. In this way, the item's dimensions complicate shelf management processes.



Item	Facing	Stacking	Depth	Capacity	Equivalent dimensions: Facings [Total capacity]
1	2	1	3	6	14 [42]
2	2	3	1	6	12 [36]
3	2	1	1	2	10 [10]
4	2	1	4	8	15 [60]
5	2	2	2	8	9 [36]

Figure 3-1 Shelf space utilization and item dimensions

Items with different dimensions may also have different demand attributes adding to the difficulty of managing store shelves. During an inventory cycle in a periodic review system, if the demand velocity (number of units sold in a week) is high, the inventory investment (number of units brought to the store per week) will be higher than for items with lower demand velocity. Similarly, if an item's demand unpredictability is high, the inventory investment will be greater than that of an item with more predictable demand, in order to buffer with safety stock for that increased variability. Higher inventory investment, however, does not necessarily correspond to increased dedicated

space. Especially if the item has larger dimensions, that item's systemic drivers (high demand velocity or unpredictability) may lead to using both dedicated and shared space. If the store's emphasis is on space more so than on labor, then additional labor can handle units of the item multiple times for instore replenishment. Since each item may vary in dimensions and in systemic drivers, and since inventory investment (cycle + safety stock) varies by demand but is constrained by store resources, managing shelves consists of several issues for store managers.

In managing shelves, the traditional trade-off between inventory cost reduction and additional transportation when considering postponement is not the issue for store management. "Store managers' incentives often do not include inventory holding costs…Thus, store managers care less about reducing inventories. They care more about product availability and labor capacity requirements because they are rewarded based on revenues and labor costs" (Donselaar et al., 2010, p. 766). Labor costs include product handling which has been found to exceed inventory holding costs (Curseu et al., 2009) and may be compared to transportation costs between firms, with instore replenishment counting as "additional transportation costs" in the traditional postponement trade-off. The opposing concern is not inventory holding cost, but shelf space, which retail scholars suggest "should be considered as a constraint rather than a cost factor" (van Donselaar et al, 2005) since it has a range of minimum to maximum shelf space. Shelf space is a constrained resource where at least one item holds a dedicated capacity from which customers find their preferred products. If the item is not at its dedicated space upon customer arrival to that space, the product is unavailable. As store management is concerned about product availability and labor, the store emphasis is either on dedicated

space or product handling frequency. The manager's choices are to either speculatively dedicate lots of space for an item on the shelf or to dedicate a minimal amount of space that can be replenished by personnel by bringing items forward to the customer (dedicated space) from shared space. Choosing to use both shared and dedicated space shows a greater emphasis on shelf space over labor. This purposeful use of both types of store space is instore logistics postponement.

This study posits that instore logistics postponement may have a fundamentally

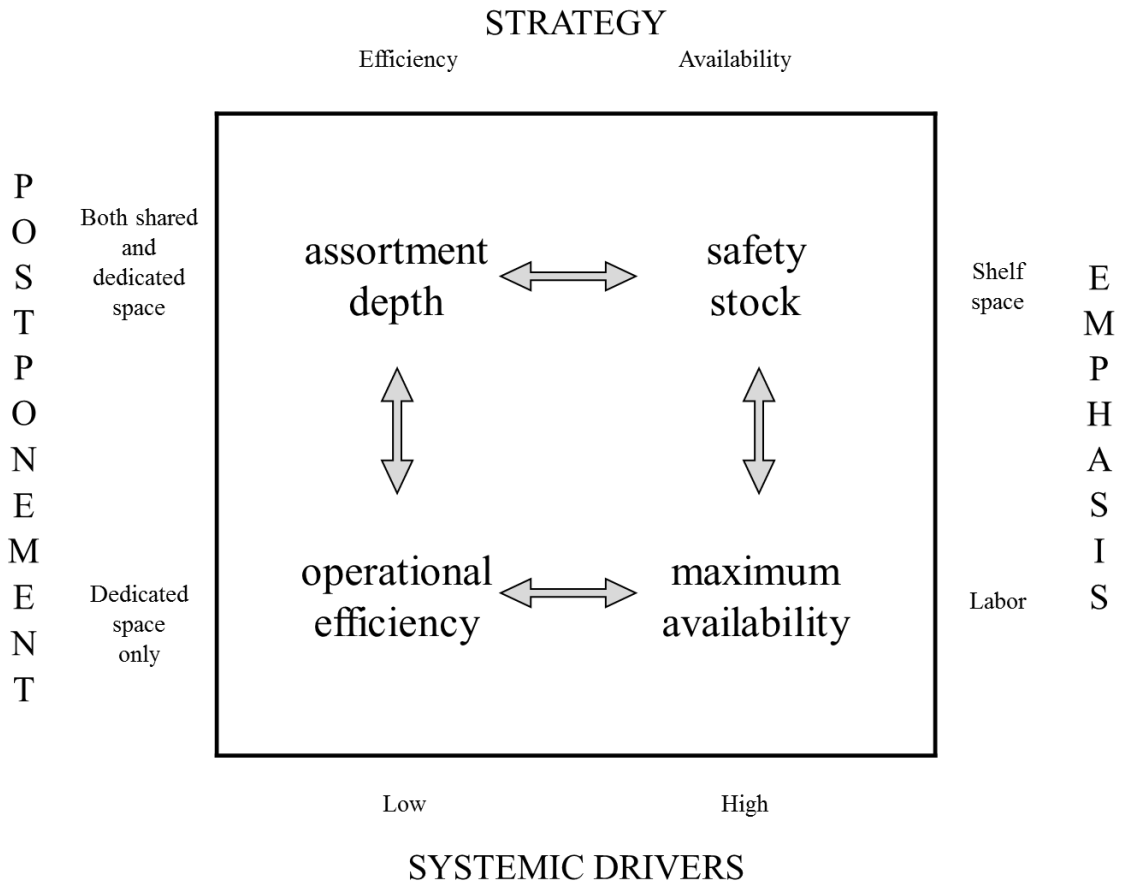


Figure 3-2 Framework for using instore logistics postponement

different mechanism, as outlined in Figure 3-2, than other types of postponement. In the box in the center are four different reasons for implementing or not implementing postponement. Postponement is implemented, as indicated on the left outside of the box,

whenever there is purposeful use of both shared and dedicated space. Dedicated space use on its own may sometimes result in some item overflow from dedicated to shared space, but it is not an example of instore logistics postponement. Intentional use of both types of space takes place whenever the store emphasis is on shelf space, as indicated on the right-hand side of the box in Figure 3-2. If the store emphasis is on labor, then product handling is minimized by using only dedicated space. On the top of the box in Figure 3-2 are the two strategies that the store manager can pursue for each item. The opposing goals (Tabucanon & Farahani, 1985) of customer service (product availability for the retailer) and operational efficiency along with the store's emphasis on different resources constitute the underlying mechanism between the relationship of demand drivers and postponement implementation.

Systemic drivers to product availability are demand drivers and their high or low demand attributes are similar to Fisher's classification of products as either functional or innovative (Fisher, 1997). Fisher finds that "if one classifies products on the basis of their demand patterns, they fall into one of two categories: they are either primarily functional or primarily innovative. And each category requires a distinctly different kind of supply chain" (Fisher, 1997, p. 105). Supply chains with functional or predictable demand items can focus on physical efficiency while those with innovative items need be more market-responsive. This concept parallels the framework presented in Figure 3-2. If an item has high systemic (demand) drivers, the store's strategy should primarily be to respond to high or unpredictable demand with enough supply to ensure product availability. If the item has low systemic drivers, the store's strategy for the item should be efficiency. Efficiency could be in terms of labor or space.

Depending on the item's systemic drivers (demand attributes) and strategy if the store's emphasis is on space instead of labor, then items with lower demand drivers could have minimal dedicated space. Since labor is less of a concern, the dedicated space for a low demand driver item can be even less than what is expected during the inventory cycle, with units of safety and cycle stock stored in the shared space. The purposeful use of both types of space (postponement) would allow for a greater variety (greater number of unique items) in dedicated space, as in the top-left corner of the box in Figure 3-2. The store may prefer this for items with low systemic drivers, so that it can capture a greater portion of the market that may have more predictable segmented preferences. If the store's emphasis is on labor instead of shelf space, operational efficiency (left-bottom corner of box in Figure 3-2) would increase by providing enough dedicated space for a store delivery, decreasing labor costs since personnel would handle the item only one time. Stores are especially incentivized to internally postpone since they cannot usually make use of economies of scale with their smaller-sized order deliveries of items with low systemic drivers. If demand drivers are high and space is an issue, then that increased inventory investment necessary for that item could be split between dedicated and shared space, so that store employees would be expected to replenish from the shared space with safety stock. If instore replenishment is more of an issue to the store because of labor constraints, then high demand-attribute items would be fully available (if it is in the store it is on the shelf, providing maximum availability) on dedicated shelf space only.

Consider that a store's labor resources can be adjusted more quickly and easily than the total shelf space in a store that has a set total square footage. For high systemic

driver items where the store's strategy is focused on item availability, it should be more likely that the store will choose to keep item safety stock in shared space (Figure 3-2, top right) from which to replenish, than to place its entire inventory investment of that item onto constrained shelf space. Minimizing the dedicated space for an item on the shelf is less of an issue for low-demand driver items which can focus its strategy on managing the item efficiently in terms of handling. Following this logic:

Hypothesis (H3.2): Systemic (demand) drivers are positively related to instore logistics postponement activity.

3.2.3 Instore logistics postponement and product availability

While logistics consists of managing the physical movement of goods between two supply chain members, instore logistics focuses on movement within a retail firm's store boundaries. The store initially places an order for a certain number of units of an item from a vendor or retail distribution center. If the order size and resulting on-hand inventory during inventory cycle of a periodic review system is less than the demand which arrives during the same time period, then STO occurs due to insufficient inventory investment (poor forecasting). Questions relating to determining or predicting the effects of insufficient order size is beyond the scope of this study and is similar to decisions on what product assortment or depth to offer (if items should be ordered at all). Hubner and Kuhn (2012) provides further reading on forecasting that takes place in assortment planning, shelf space planning, and instore logistics planning, under the umbrella of category management. Similarly, research questions about product availability under

inventory systems other than periodic review follow a different stream of research (Cachon, 2001).

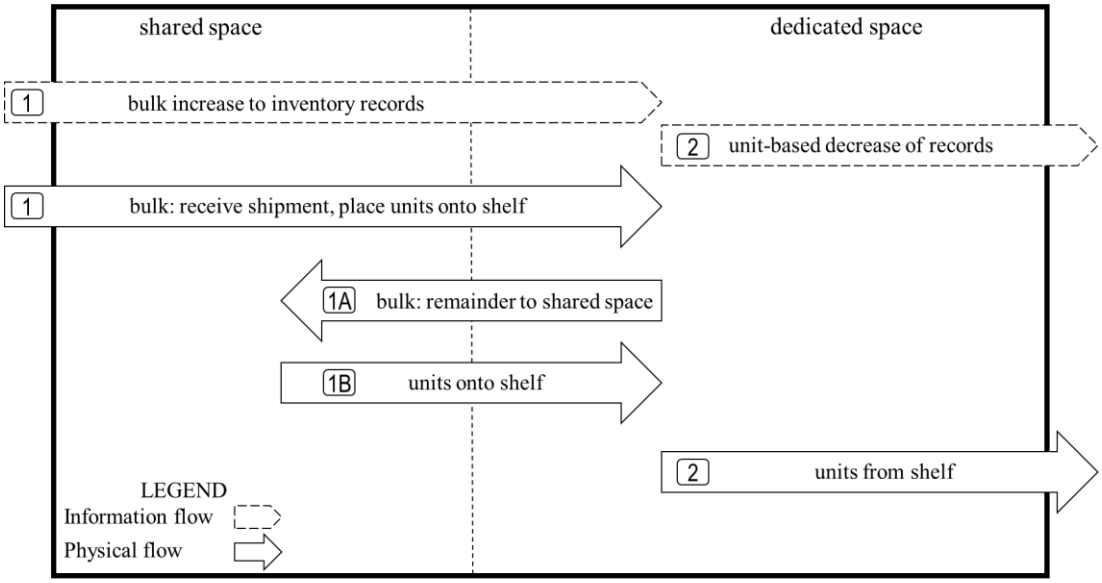


Figure 3-3 Bulk and unit handling in a store, both physical and information

Once out-store logistics (Samli, Pohlen, & Jacobs, 2005) delivers items to the store, information flow and physical flow takes place, as pictured in Figure 3-3. First, the item is initially physically handled in bulk as personnel receive the shipment. This physical flow is mirrored by information flow, as inventory records are also updated in bulk. Both of these flows are illustrated as arrow (1) in Figure 3-3. The next information flow to occur is on a unit-based basis. As customers purchase one or more units of an item that they take from the shelf, the on-hand inventory for that item is updated with an incremental decrease, to mirror the decrease in physical units on hand. What is not reflected in inventory records is the additional physical flow that occurs between step 1

and 2, especially if instore logistics postponement is implemented for an item. Since both shared and dedicated space are purposefully used, after the bulk receipt of items to the store and placement of them in step (1), units remaining in the box for dedicated space are handled in bulk (1A) only to be brought out to the shelves again (1B) after customers take units away from dedicated space.

While the information flow masks the actual physical flow within the store, it is sufficient for the "well known" effect termed the "store-level fill rate effect"(Waller, Heintz Tangari, & Williams, 2008). This is where a greater number of units delivered to the store improves product availability by avoiding more frequent smaller store deliveries which face lead time uncertainty. If the store has more of an item, it can sell more of it. Countering this store fill rate effect is the backroom effect (Waller, Heintz Tangari, & Williams, 2008), which recognizes the physical flow that occurs at arrows (1A) and (1B) of Figure 3-3. The backroom effect states that an inability to carry out step (1B) properly results in stockouts due to poor indirect or instore replenishment (IRO), erasing any possible benefits of higher inventory investment or the store fill rate effect.

This study differs from the theoretical approach of the store fill rate effect versus backroom effect research stream. Backroom effect work essentially study higher inventory investment (larger order size arrival to the store) effects against poor indirect replenishment practice where step (1B) does not take place. Corsten and Gruen (2008) find that 25% of SHO have inventory in the backroom or shared space (a quarter of SHO are IRO in their analysis), so if instore replenishment of dedicated space had taken place *proactively* (before the dedicated space is empty) IRO frequency would be lower. The

reasoning behind the increase of SHO in backroom effect studies are two assumptions. First, since there is no accessible information on the on-hand inventory in shared and dedicated spaces as separate inventories, store personnel are not even aware that there are units on hand in shared space. Secondly, inventory overflow (Eroglu, Williams, & Waller, 2013) is the unintentional phenomenon of some portion of a store delivery not fitting on dedicated space so there is no effort to proactively replenish shelves from shared space.

A purposeful delay of forward movement of goods to the customer, or instore logistics postponement, implies that store personnel are cognizant of shared space inventory even if it is not being formally tracked during the process of information flow. They know to check shared space for an item even before that item is fully stocked out in its dedicated space. Indeed, interviews with some product category managers provided evidence of personnel being required to bring all shared space inventory from that product category out once or twice a day and replenish dedicated space (shelves) with any shared space inventory. This practice of purposefully using both dedicated and shelf space, either for a strategy towards efficiency for items with lower demand drivers, or availability for high demand driver items differs from the inventory overflow of backroom effect studies. In Figure 3-2 the lower portion, where store emphasis is on labor, there may still be shared space use through inventory overflow. Inventory overflow would occur in these cases any time demand is lower during the lead time period than expected or the lead time period is shorter than expected, leading to products delivered to the store not fitting on dedicated shelf space. That type of shared space use is not

intentional and would more likely drive the backroom effect. It would not be considered as instore postponement.

Store managers may want to implement instore logistics postponement to improve product availability along different dimensions. First, the time it takes to replenish from within the store (shared space) does not include the transportation time or uncertainty related to transportation of the items to the store. This is similar to the logic of the store-level fill rate effect. By being able to replenish goods at dedicated space within a shorter time period, stockout (IRO) duration would decrease, and availability would increase. If aggregate availability is the sum of each individual item's availability, then these aggregate store, product category or vendor stockout durations would be lower overall than without postponement. Secondly, since postponement is purposefully being applied and not merely because of unintentional inventory overflow, when replenishing one stocked out item store personnel may also replenish another item within the product category which has yet to IRO. By doing so any aggregated availability (vendor, product category, or store stockout rates) would be improved by lower IRO frequency at these higher-level availability measures.

Postponement may not just be decreasing IRO occurrence, but also STO in two different ways by effectively increasing the capacity of the store. First, items with higher systemic drivers could have increased inventory investment than had the store simply ordered for dedicated space use only. With larger orders to the store, it is more likely that STO will also occur less often. Secondly, items with larger dimensions could take up less dedicated space without having to increase order frequency for that item or any other

item under periodic review. As illustrated earlier, larger items take up more unit space than a smaller item could fit into the same capacity. Since each unit takes up more capacity of dedicated (shelf) space, they are also more likely to stockout because the maximum number of units dedicated to them is likely fewer than for the same capacity of smaller items. Storing such items within the store also would be expected to decrease STO as it would for items with high systemic drivers, because the inventory investment of the item during the inventory cycle is greater. With decreasing IRO and decreasing STO occurrence, the overall effect of the use of both types of store inventory on SHO occurrence is hypothesized to be:

Hypothesis (H3.3): Instore logistics postponement activity is positively related to availability.

3.2.4 The indirect effect of systemic drivers on product availability

Since the earlier theoretical framework argues that store managers may implement instore logistics postponement for different reasons depending on the item's demand attributes and that postponement's effects on availability are largely influenced by an item's dimensions, the indirect effect of these demand attributes on product availability is complicated. On the one hand, if instore logistics postponement is implemented for items with high demand drivers as a way to improve availability by using shared space as a place for safety stock (a positive relationship between demand drivers and postponement occurrence), then the higher the demand driver for an item, the more likely postponement is being implemented and the better dedicated space is being maintained. Postponement

would then weaken or completely remove demand's detrimental effects on product availability.

On the other hand, if postponement is implemented for items with low demand drivers as a way to improve operational efficiency by fitting more unique items (increased assortment depth) within the same constrained shelf space, then the higher the demand driver for an item, the more unlikely postponement is being implemented. Increasing assortment depth captures more sales (Ton & Raman, 2010) but also decreases operational efficiency in two ways. First, since there are fewer units of each item in dedicated space, similar to items with larger dimensions, the space will SHO more quickly with the same demand than for smaller items with many units of dedicated space. Secondly, managing the movement of items from shared to dedicated space may be more difficult for a higher number of items than for a lower number of items, as has been shown at the distribution center level of the SC (Wan, Evers, & Dresner, 2012). This inefficiency may result in postponement not improving demand's detrimental effects on availability or even worsening those effects. Therefore, demand's indirect effect on stockouts is hypothesized as:

Hypothesis (H3.4): Postponement mediates the relationship between systemic drivers and availability.

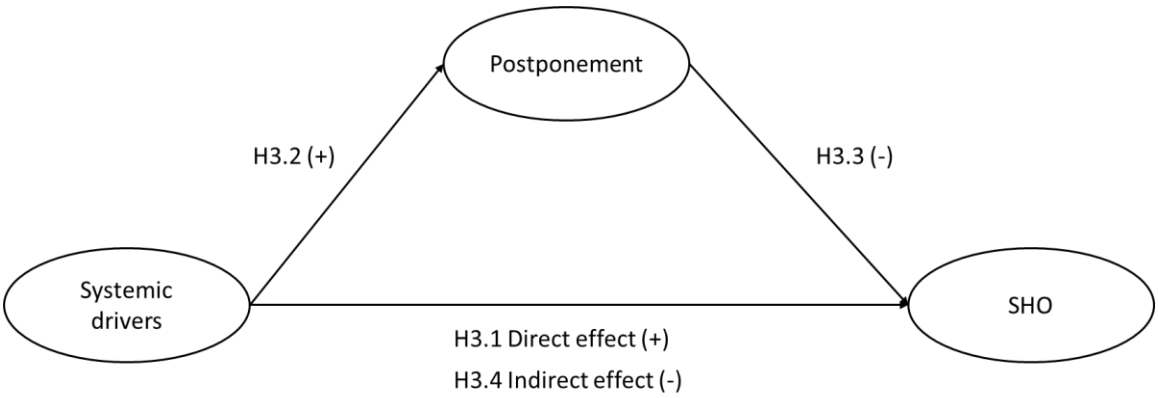


Figure 3-4 Hypothesized model for instore logistics postponement

3.3 Method

3.3.1 Sample and Procedure

Data are from four locations of an American grocery store chain in the Great Lakes area. There are 102 unique items from two product categories (isotonics has 48 items from two vendors while yogurt has 54 different items from three vendors). Data are collected from five sources, as listed in the first column of Table 3-1: shelf liner ping data, prior-year weekly category sales file, pilot study daily SKU sales, store planograms for each product category and store visits and interviews. The shelf liner ping data are obtained from a proprietary shelf technology which gathers time stamps of when a dedicated space for an item is completely empty and when it is restocked. The shelf liner tracks each item in all four locations but varies on when tracking begins from store to store and even within product categories (Appendix IX). While the shelf liner data cover a range of 370 days total, the length of time that each item is tracked is on average 263 days (SD 94.39, min 2, max 370), or 128 days for yogurt and 319 days for isotonics. Each observation from this shelf liner data source is a single stockout occurrence for a store

item. Higher-level (regional) retail managers provided prior-year weekly category sales which cover sales of entire product categories beyond the items whose availability is tracked by the shelf liner technology. A pilot study for the technology lasted 8 weeks. Daily total sales (units sold per day) during the pilot study was separately provided by point of sale reports for each store, generated for each item the shelf liner technology tracked during this pilot period. Tracked items were said to be selected by vendors as high-selling items contributing to most of the product category sales in terms of revenue. Each product category's planogram for each store provided time invariant information on case pack size of items and how they were placed on dedicated shelf space. Store visits confirmed each store's almost complete adherence to planogram product placement in terms of facings and provided information on stacking height and depth. Instore visit information replaced any differences found in stacking height as compared to planograms.

Since items are tracked for stockouts at different times during the period from October 27, 2015 through October 31, 2016 at four stores with different operating hours, the data is an unbalanced panel of real time stockout and replenishment information that is collapsed into a weekly basis. By week, each item's store week (Thursday to Wednesday) total SHO duration and SHO frequency is generated for a total set of 12,109 weekly store item stockout observations. The SHO duration observation (in minutes) for each item is then divided by the total number of minutes each store is open each week, which also varies within the store for weeks with holiday-adjusted hours and for two stores which permanently changed their operating hours during the first half of the pilot study. Bonferroni-corrected (95%) pairwise comparison of the mean number of items

tracked in each store every week shows store A, B, and C do not vary significantly (ranging from a mean of 21.95 to 22.73 units) while store C and D do not vary significantly (21.54 and 21.95) from one another. Even though store D has a significantly lower number of items tracked every week compared to stores A, and B, this number differs by only one unit.

The data's final size (N=12,109) is reached through a series of steps. First, since there is no information about the exact start and end days of tracking each item in each store, of the original 13,271 observations, the very first and very last SHO observations of each item are used to generate known tracking date ranges. The very first "empty" shelf liner time stamp and the very last "replenishment" time stamp for each item in each store is set as the start and end date, respectively, for each item's tracking period. Next, the SHO observations are limited to a maximum duration of 7 days (leaving 13,043 observations), since interviews revealed every store's order size is for this length of time (a periodic review system) and orders are received every week unless the items are managed directly by the vendors themselves. The assumption behind this upper limit is that any SHO duration longer than 7 days occurs from an order forecasting, vendor or distribution delivery issue (out-store logistics issues) and not with instore logistics decisions as is the focus of this study. The remaining SHO events are flagged as unique stockout events to later generate weekly item SHO frequency. Observations are then created for every day that each item is tracked, bringing the total observations to 222,527. These observations are also flagged to be able to count the number of days each item is tracked and the number of items tracked during any given week. The latter values are

used to generate category SHO breadth, or the proportion of items tracked within a category which are SHO.

The data must be changed from observations based on stockout and replenishment pairs to daily observations spanning the date range the item is tracked for 2 reasons. First, it is necessary to be able to properly split SHO that occurred over more than one day to account for times that the store was closed and customers had no access to items anyway. In terms of product availability, it does not matter if the item is not on dedicated space when customers have no access to that space (when the store is closed). Additionally, splitting SHO events into daily values is necessary for those events that span over more than one store week. For example, if a SHO event begins on a Tuesday but is replenished that Friday, since Thursday is the beginning of a store week, the event spans two different store weeks. The duration of open store hours from the moment it stocked out on Tuesday through the store closing on Wednesday must be included in one store week's duration calculations, while the duration on Thursday and the portion of Friday until the item is replenished must count into the following store week's SHO duration measures. Since category systemic driver data (category sales data) were provided as aggregate store week numbers, it is necessary to aggregate all of the data to the largest dimension or highest level of aggregation among all of the data sources. Collapsing the daily unbalanced panel data into weekly values leaves 13,999 observations of data for this study. Items from the data are dropped if they do not appear in all 4 stores. This brings the number of unique items down to 102 (54 in yogurt and 48 in isotonics) from 135 items in combined the files provided, with total observations down to 12,109 store item days. These collapsed observations contain time invariant item attributes, stockout information, and weekly

category sales information. Creating daily observations makes it possible to include weeks into the dataset where an item was definitely tracked (it was within the range of first and last tracking dates) but where no stockouts occurred.

Characteristics of stockout and sales data in this study (Table 3-2) are all tested (see Appendices IV - VII) by pairwise comparison of means with equal variances and Bonferroni corrected for significance at the 5% level. Weekly SHO duration (in minutes and not as a proportion of open selling time) does not significantly differ between stores A, B, and C for yogurt (or store A isotonics), while store D appears to have a significantly higher mean weekly SHO duration. Store B does not have a significant difference in SHO duration between yogurt and isotonics. Store B's SHO duration for isotonics is not significantly different than isotonics in store C. For SHO frequency, only the following groups have no significant difference: store A and store C for yogurt, store B and C for isotonics and store B's yogurt. All other values of SHO duration and SHO frequency between stores significantly differ, requiring control variables for store differences in the model's availability structural equation.

The same process for sales also shows the need to control for store differences in the systemic drivers construct. Multiple comparison testing was preferred over using ANOVA F-statistics to confirm store and category group differences because the Bonferroni grouping allows for seeing how groups are different from one another beyond just not being statistically equivalent. Specifically, stockout duration and frequency is significantly higher for each product category in store D than all other stores, but its item sales is on average less than store B and category sales do not differ from store B in

isotonics and is less than the yogurt sales in store B. This initial look at group differences suggests these duration and frequency variables do not relate to demand behavior at the item or category level in a simple manner (the highest average category or store SHO occurrence is not associated with the largest or smallest realized demand).

The study sample represents the product categories at different proportions and with different item-based attributes. That is, compared to the assortment depth and total category shelf space set aside to each product category, this study's sample differs by category. Yogurt is given 32 shelf feet[6] in 3 of the stores (A, B, and D) and 36 shelf feet in store C. The total number of unique items that are offered in this category vary from 389 to 427 different items. The planograms for isotonics show that this category is given 16 shelf feet in every store and has 88 unique types of items. This study sample tracks 55% of isotonics products that are offered in stores, and about 10% of the yogurt product assortment offered.

[6] Shelf feet are the width of shelves against which facings are measured. Standard retail shelving consists of 4-foot shelving panels that are placed side by side to form the store aisles.

Source	Data fields			Used to generate
	Construct indicators	Model controls	Interim variables	
Daily POS data (56 days)	item sales velocity (mean weekly sales) item sales unpredictability (standard deviation of weekly sales)	item unit price vendors		Systemic drivers (demand) construct
Weekly category sales data (53 weeks)	weekly category sales velocity (prior year weekly category sales) weekly category unpredictability (standard deviation of weekly sales)			Systemic drivers (demand) construct
Shelf liner data (370 days)	item weekly stockout duration (sum of all events over each week) item weekly stockout event rate (total number of unique item stockout events each week)	sum of tracked items in the store each week sum of tracked items within a category each week proportion of week item tracked by liner		Availability (SHO) construct
Planograms		shelf height case pack size	item facings	capacity, case pack size → Postponement construct
Store visits and interviews		pings used	item stacking item depth store hours	capacity variable → Postponement construct stockout rates → Availability construct

Table 3-1 Data sources used for this study

Product category	Store	**Shelf stockout (SHO) weekly duration** (proportion of open store minutes that item is SHO)				**Weekly SHO Frequency** (count of unique SHO events per week)				**Average item weekly sales** (mean item weekly sales units, varies by item only)				**Category weekly sales** (count of total units sold per week of entire product categories)				Obs.
		Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	
Yogurt (weeks 1-26)	A	0	75.67	2.42	6.72	0	6	0.60	0.98	8.23	57.05	29.19	11.60	4,376	8,120	6,205.49	944.13	885
	B	0	71.43	3.23	7.00	0	8	0.92	1.35	13.80	161.09	55.11	28.53	8,777	15,718	12,185.93	1,614.93	943
	C	0	81.26	3.37	10.54	0	8	0.65	1.20	5.13	103.88	33.95	18.42	5,231	9,733	7,642.08	1,158.73	840
	D	0	86.77	6.99	11.28	0	12	2.21	2.17	8.96	122.45	46.84	24.40	5,200	12,387	9,529.38	1,814.15	858
Isotonics (weeks 1-53)	A	0	93.01	2.51	8.33	0	7	0.39	0.91	.89	47.95	12.59	9.72	593	3,075	1,538.73	702.18	2,287
	B	0	97.63	4.30	10.56	0	7	0.86	1.41	.63	50.39	18.62	11.93	1,103	3,115	1,869.01	514.56	2,235
	C	0	89.39	5.08	10.60	0	7	0.90	1.37	1.75	42.16	13.87	8.20	768	2,107	1,257.54	357.17	1,939
	D	0	88.35	5.36	10.99	0	11	1.24	1.69	1.41	76.68	20.09	15.60	1,022	3,109	1,896.96	534.94	2,122

Table 3-2 Descriptive statistics of item sales and shelf stockouts (N=12,109)

Construct	Measurement indicators	Min	Max	Mean	SD	Factor loadings	Measure fit
Systemic drivers (demand)	Weekly category sales (divided by 100)	5.93	157.18	37.77	36.57	0.6276	SRMR 0.020
	Mean weekly sales of item	0.63	161.09	23.66	20.16	0.9437	CD 0.661
	Standard deviation of weekly item sales	0.70	38.59	8.62	6.93	0.9189	
	Cronbach's α : 0.8717 AVE: 0.6937						(errors clustered by store week)
Instore postponement	Units shelf space exceeding multiple case pack size ("excess space")	0	9	2.58	2.94	0.8345	SRMR: 0.074
	Postponement dummy (0=capacity is an exact multiple of case pack size)	0	1	0.60	0.49	0.9717	CD 0.963
	Proportion of shelf space exceeding multiple	0	37.50	14.56	14.01	0.9289	
	Proportion case pack size in shared space	0	7532.54	27.91		0.8879	(errors clustered by store week)
	Cronbach's α : 0.9356 AVE: 0.7840						
Availability (SHO)	Vendor SHO event rate (breadth)	0	100	41.06	26.85	0.8826	SRMR: 0.035
	Store SHO event rate (breadth)	0	100	40.82	20.62	0.6933	CD 0.462
	Category SHO duration rate	0	61.32	4.46	3.82	0.6014	
	Category SHO event rate (breadth)	0	100	43.87	24.83	0.8826	(errors clustered by store week)
	Cronbach's α : 0.8359 AVE: 0.5602						

Figure 3-5 Model measurement indicators

3.3.2 Model specification

Measurement of the theoretical constructs:

The construct indicators are from all of the data sources (Table 3-1) gathered for this study and are listed by construct in Table 3-3.

Systemic drivers (demand) are measured with 2 category- and 2 item-based indicators, as summarized in Table 3-3. One category sales measure, often used as a measure for demand, is the number of units that are sold in each product category in each store per week during the prior year. Year-lagged weekly sales values are used as a measure of demand seasonality or difference in a item's demand from week to week in each product category. The standard deviation of category sales as well as that of individual item sales is often included as a measure of demand unpredictability (Moussaoui et al., 2016). Both the item standard deviation (unpredictability) as well as the mean item sales per week (velocity) are generated from a 56-day daily point of sales file. Since the item sales information is provided on a store-week basis while the availability data is nine calendar weeks, the weekly item sales generate an average daily sales per calendar day which are then combined into calendar week sales.

Moussaoui et al.'s (2016) other measures for systemic drivers are either included in the structural portion of this construct or in another part of the structural equation model. Demand autocorrelation is accounted for with robust standard errors that are clustered around store weeks. Assortment depth is accounted for in the structural portion of the systemic driver equation in terms of the number of items that are tracked every

week by the liner. This is used instead of the number of different items offered in each product category, which would be a static value from week to week. Assortment depth measured by the number of items tracked controls for the absence of SHO events for items during periods when they are not tracked. The last component described for systemic drivers is network design. Network design deals with the physical position of the store in relation to its vendors and retail distribution centers. Since network design is said to involve lead times from retail suppliers to the store (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016), this study views this as a supply driver and not a demand driver. It is included in the structural portion of the postponement equation in terms of vendor dummies.

Item	Shelf capacity	Case pack	Shared space units	(1) Extra dedicated space	(2) Postponement dummy	(3) Extra space per shelf capacity	(4) Shared space units per case pack size
1	6	6	0	0	0	0	0
2	6	5	4	1	1	16.7%	80%
3	2	1	0	0	0	0	0
4	8	7	6	1	1	12.5%	85.7%
5	8	12	4	0	1	0	33%

Table 3-3 Example of instore logistics postponement measures

Instore logistics postponement is measured with 4 indicators relating dedicated space to shared space. Since the literature has limited measures of postponement at the retail level (Li et al., 2005), and this study aims to contribute to postponement literature with the concept of instore logistics postponement, the indicators in this measure have not

been previously tested in prior literature. Since retailers normally refer to their dedicated shelf space in terms of the ratio of case pack size (Eroglu, Williams, & Waller, 2011) this study focuses on the relationship of the two different volumes (case pack size and dedicated shelf capacity) in different ways, as in Table 3-4, columns labelled 1-4.

There are multiple assumptions in developing these indicators. First, it is assumed that instore logistics postponement, or the planned and purposeful use of both shared and dedicated space does not occur if the dedicated space is a multiple of case pack size. For example, if the case pack size is 12 units and the shelf capacity is a multiple of this at 12, 24, or 72 units (etc.), then any time items do not fit in this dedicated space are unintentional inventory overflow. A 0/1 postponement dummy (column 2 in Table 3-4) in this case is valued at 0, because there is no intentional use of both types of retail space. When this dummy variable equals 1 and the dedicated space is not an exact multiple of case pack size, at least one of the other 3 indicators also have non-zero values.

The remaining 3 indicators increase in value as the likelihood that the store is implementing instore logistics postponement increases. The first, column 1 in Table 3-4 called “extra dedicated space,” is a direct measure of the number of units of shelf space that exceeds an exact multiple of item case pack size. It is possible that there is extra dedicated space not to purposefully split a case pack’s contents between shared and dedicated space, but simply as a result of item dimensions. However, it is less likely that a higher number of units of dedicated space is in excess simply because of this item attribute. The third (3) postponement indicator is the percentage that this extra dedicated space takes up in the overall dedicated space for an item. Items 2 and 4 in the example

highlight why this measure is necessary in addition to the second indicator. Both items have 1 unit of extra dedicated space, however it constitutes a higher proportion of the overall dedicated space for item 2 than it does for 4. In terms of using shelf space more efficiently the third indicator suggests that item 2 has a higher likelihood of implementing postponement since the extra dedicated space in proportion to the overall space is higher. On the other hand, if the store manager is more focused on minimizing labor instead of shelf space, the fourth indicator is a measure of how much of the case pack contents are put into the shared space. If the goal is to replenish the shelf one time with store delivery and not have personnel carry out instore replenishment from shelf space (thus minimizing labor), then this fourth indicator would be zero as it would for the third indicator as well. With increased reliance on instore logistics postponement, the proportion of case pack sent to shared space would increase. In instances where the case pack size is bigger than the dedicated space on the shelf for the item, this indicator would have non-zero value even though the excess space-based indicators (columns numbered 1 and 3 in Table 3-4) would both equal zero. The data sample for this study has no items whose case pack size has a greater number of units than the dedicated space for the item.

Shelf stockouts (SHO) are measured with three indicators. Gruen and Corsten (2008) state SHO events have certain attributes (measurable characteristics) and recommend use of multiple SHO attributes at once to better understand the drivers and impact of SHO. In developing the availability measure the types of SHO attributes considered were related to the frequency, duration and breadth of SHO events. The four indicators initially considered are: item SHO frequency and duration, and category SHO breadth and duration. Two item-based SHO measures are chosen since the framework

rests on item attributes (demand and size) and two category-based SHO measures chosen since these are the most often used measures and the model incorporates weekly category sales as part of the demand drivers. Although all four indicators had a Cronbach's α of 0.7655, the AVE was only 0.45. Once the indicator for category SHO breadth was removed, the AVE rose to the minimum recommended threshold of 0.50. Table 3-3 lists the final confirmatory factor analysis results.

All 3 indicators are generated following prior convention. Item stockout frequency and duration are generated as described in the data sample description and summarized in Table 3-2. Category SHO duration is generated in multiple steps. The numerator of this rate is simply the sum of all of the SHO durations (in minutes) during each week, by category. The denominator takes the number of items tracked that week and multiplies it by the number of minutes each store is open that week. The SHO measures that Gruen and Corsten (2008) recommend which are not included in this study are indicators of breadth, intensity and impact. The latter two are financial and unit-based measures of the sales loss incurred due to SHO. They require some sort of assumption or estimation of the demand that would have occurred during the time of day or proportion of the day during which the SHO occurred, and are beyond the research scope of this study. The breadth measure had to be dropped as the fourth indicator in this construct, as the AVE when including it was below 0.5.

All test scales are based on the means of standardized items. Standardization is required because the construct indicators have different scales. For example, in the Postponement construct, there is a 0/1 dummy that only varies by these two values, an

excess space value which is a count of units of space on the shelf that exceed an exact multiple of case pack size, and two variables which are percentages and vary from 1 to 100.

Unidimensionality and validity requirements are achieved through several processes. The factor loading value for the retained indicators for all 3 constructs, as listed in Table 3-3, are greater than 0.60 and are significantly associated with its corresponding construct, supporting unidimensionality of the constructs. The lowest factor loading for the systemic (demand) drivers construct is weekly category sales. It is above the minimum threshold but also the only indicator in the measure which varies by week, so this lower value is not a source of concern. Convergent validity for each measure is achieved since all values (Table 3-3) of AVE are greater than 0.50. Discriminant validity is achieved for each measure and described in detail for the overall model specification (Table 3-8, "Fitted covariances and square root AVE").

Construct validity is achieved when fitness indices meet minimum thresholds. Since longitudinal data is being used and there are clustered errors to account for correlated errors over time, the only two statistics for fit which are provided are in the last column of Table 3-3. The standardized root mean square residual (SRMR) is a measure of how close the model's implied correlation matrix is to the predicted correlation and has a maximum recommended cut-off of 0.08 (Hu & Bentler, 1999). The coefficient of determination (CD) is a measure of what percentage of the variation in the data is explained by the model. Each construct is below the recommended fit statistics for SRMR. For the Availability construct, the measure accounts for less than half of the

variation in the data. Given that the factor loadings of two of the indicators were below 0.70 (but still above the minimum 0.60), it is possible that Availability suffers from some measurement issues.

Reliability of each construct measurement deals with how well each intended latent construct is captured. Internal reliability is achieved with Cronbach's alpha greater than 0.70. Although the availability construct has average variance extracted (AVE) below 0.60, it is still above the conventional threshold of 0.50 and falls into "good measurement" (Hair Jr, Babin, &Krey, 2017, p. 173). The composite reliability is manually calculated to be above the minimum threshold of 0.6 (Demand CR=0.7096, Postponement CR=0.8230 and SHO CR=1.500). Furthermore, no correlated error terms or constraints on covariance between variables are necessary to achieve good overall fit to the final specification of the model, which further supports good measurement (Hair Jr, Babin, & Krey, 2017).

3.3.3 Structural covariates

The structural portion of the model relates the constructs with one another, but each construct may also be affected by factors other than the constructs being tested. The aim of control variables or covariates is not to test new theory but control for factors found from prior work. A full model without any of these controls quickly converges without these additional observed variables, and carries the same sign and significance between constructs as the full model but with inflated coefficient values. Furthermore, a model with measurement and structural models combined reduces standardized root mean squared residuals showing a better overall fit of the model (from 0.082 without the structural controls to 0.071for the full model). Since likelihood-ratio tests are not possible with clustered standard errors (Statacorp, 2015), comparing this measure of residual size between the two models is preferred and shows the benefit of including covariates into the final specification.

There are two controls that are unique to this study and may need further explanation. First, there is a 0/1 dummy controlling for whether store personnel are notified of empty dedicated space. Given this new shelf liner technology, two of the stores in the sample never used the shelf liner alerts and only gathered SHO and replenishment data. The other two stores did begin relaying SHO pings to its personnel after a few weeks of having it functioning in their stores. This dummy is included in case such alerts provide instant notification of a SHO without having the personnel discover it themselves. Such an alert may be linked to lower average SHO duration than stockouts

that are tracked but are not notified to store personnel. Secondly, there is a variable to control for the proportion of the store week that the item is being tracked by the shelf liner. While 97% of the observations have this valued at 1 (the item was being tracked during that entire store week), some 656 observations vary in value from 0.14 (being tracked 1 of the 7 days in a store week) to 0.86 (6 days of tracking). More than half of these 808 observations occur during the first 3 weeks. The remaining observations of partially tracked weeks are dispersed through every week of the data. Every week has at least one item that is being partially tracked. This control is included so that lower SHO values during those weeks when the item is not fully tracked are accounted for.

The data gathered from the different sources serve 3 different purposes as outlined in Table 3-1. The measurement portion of the model has indicators for each construct and uses all data sources to generate these indicators. The structural portion which draws paths between the constructs to show the relationship amongst them, have covariates (controls) for each construct. For example, the shelf height is a count variable from 1 to 7, specifying the relative height of the shelf of each item's dedicated space. This is included to account for marketing efforts (Dreze, Hoch, & Purk, 1994) to capture customer demand by putting them at expected eye-level or near other brands. Interim variables are those that are collected to generate other variables and are not directly used in the path analysis. In other words, the data gathered about each item's facing, stacking and depth units are used to generate the capacity variable. The information about each store's hours are used in two different ways. First, the open and close times help find net stockout duration for each SHO event. If an event's raw duration is 10 hours, the store

hours help to determine what portion of that 10-hour duration occurred while the store was closed. For example, if there are two stockout events beginning on a Tuesday, one at 10 am and the other at 10 pm, even if they both are replenished 10 hours after they stockout, the one which stocked out an hour before the store closes would have a shorter unavailability than the one which was stocked out all day while the store is open. The second way store hours are used is to generate store open selling time in minutes. These values are first generated on a daily basis and then collapsed over the store week as summed total weekly minutes of selling time. The store hours data fields themselves, however, are never used in the model directly as variables. This is why they are interim variables; they are generated solely to aid in generating model variables.

The other variables generated in the study that are used in the path analysis as covariates are listed in Table 3-5. The first step in generating the "item regular unit price" from the 56-day sales file is by taking the total revenue for the day and dividing it by the number of units of each item sold that day. This gives an interim variable for average unit price for that day which may or may not be the item's regular price. Then each item's maximum unit price over all stores is determined and double-checked online to confirm that it is at or near that item's regular retail price at various retail websites. This maximum value, double-checked online, is entered as the item's regular unit price and joined with the master data combining stockout observations and category sales. The dedicated space for an item (its shelf capacity) is generated by: facings x stacking x depth, which are data fields obtained from planograms and store visits.

There are also dummy variables in the model that are generated without having to rely on the data sources. Control variables coded as 0/1 for the store weeks of Christmas and New Years are created to correspond to when those holidays occur. They are included because Christmas is the only day of the year this chain of stores is closed, and New Year's Eve and Day are the only two days where the store hours are the same across all stores, since even the 24-hour location reduces its open selling time on those two days. Since there are 4 stores, 3 dummy variables carrying 0/1 variables all equal zero in the base case of store location A. Similarly, the 5 vendors have four 0/1 dummies with the base case being a yogurt vendor.

Construct	Control variables	Min	Max	Mean	SD
Systemic drivers (demand)	Item regular unit price	0.7	9.99	2.78	2.52
	Shelf number of dedicated space for item (6) (1=bottom shelf (base case), 7=top shelf)				
	Christmas week, New Year's week				
	Store dummies (3)				
	Number of items in category tracked per week	1	51	22.20	13.10
	Dedicated shelf capacity for item	4	72	20.45	12.76
	Weekly minutes each store is open	6180	10080	8101.93	1223.38
Instore postponement	Vendor dummies (4)				
	Store dummies (3)				
	Dedicated shelf capacity for item				
	Case pack size	3	15	9.16	4.27
Availability (SHO)	Weekly item stockout frequency				
	Weekly item stockout rate (proportion of selling time)				
	Whether store personnel are notified of empty dedicated item space	0	1	0.45	0.49
	Proportion of week each item uses liner	0.14	1	0.97	0.13
	Christmas week, New Year's week				
	Store dummies (3)				

Table 3-4 Model controls

3.4 Empirical results

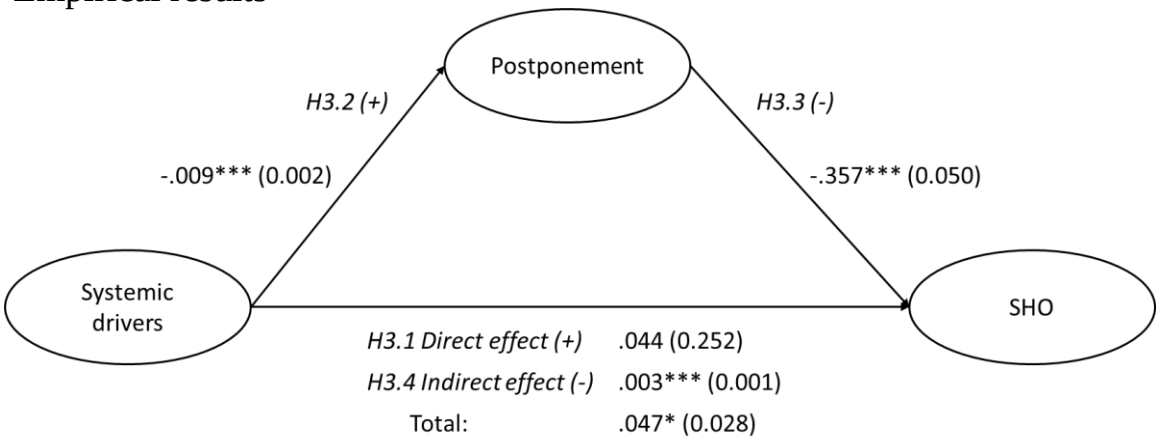


Figure 3-6 Generalized Structural Equation Model: Standardized Main Results (Clustered Standard Errors)

The model estimation's standardized results are provided in Figure 3-5. The model is estimated with Stata/SE 14.1 using the <sem> command. Structural equation modeling (SEM) is preferred over multiple regression equations to study postponement's mediating effects on product availability because the SEM framework provides more advantages (Gunzler, Chen, Wu, & Zhang, 2013) in mediation analysis in various ways including simpler ad hoc analysis. The estimation technique is maximum likelihood instead of ordinary least squares so there is a built-in mechanism (listwise deletion) for missing data (Gunzler, Chen, Wu, & Zhang, 2013). Since the dataset has repeated measures over time and over stores it is necessary to relax the restrictive assumption of independence of errors and the model is estimated with errors correlated within clusters of a store and time variable (Petersen, 2009). These store-week clusters allow for correlation of error terms wi thin these groups, replacing it with an assumption of

independence between store week clusters. This approach is more generalized than using robust standard errors, which also account for heteroscedasticity of the errors but require errors to be independently distributed (Statacorp, 2015). Because of these clustered errors, the fit statistics provided in Table 3-6 are residual statistics: the standardized root mean square residual (SRMR) as a measure for how close the model's implied correlation matrix is to the predicted correlation, and the coefficient of determination (CD) as a measure of what percentage of the variation in the data is explained by the model. The SRMR value of 0.071 is below the recommended cutoff criteria (Hu & Bentler, 1999) for good fit and 83% of the data's variation is explained by the model variables.

Hypothesis 3.1 is not supported; although the coefficient for the relationship between systemic drivers and availability is the correct sign, it is not significant ($p=0.124$). Recalling the conflicting findings (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016) between an item's systemic drivers (demand attributes) and product availability (SHO), these results do not find support that product availability is worse for faster-moving (Corsten & Gruen, 2003) and more unpredictable (Taylor & Fawcett, 2001) items, nor do they support that availability is improved (Fernie & Grant, 2008).

Hypothesis 3.2, which theorized a positive relationship between systemic drivers and instore logistics postponement, is not supported. The data does not support the idea that store managers use both shared and dedicated space for items with high demand attributes, suggesting perhaps they lean towards making fast-moving or unpredictable

items fully available on store shelves. Availability may be the strategy for items at these higher systemic driver levels, where the store's emphasis is on minimizing labor (less product handling). The significant negative coefficient also suggests that items with lower systemic drivers use both dedicated and shared space and internal postponement decreases with higher systemic drivers. The conceptual framework presented in this study reasons that items with lower demand drivers use postponement to be able to increase product variety in a store or assortment depth within a category which emphasizes efficient shelf space use.

Hypothesis 3.3 is supported by the negative and significant path coefficient from instore logistics postponement to SHO. Prior studies which show greater stockouts at higher velocities (Gruen & Corsten, 2008) suggest that slower items simply may not be selling enough for the capacity allotted for them on the shelves. The model in this study controls for average item demand attributes as well as the shelf capacity and item stockout frequency and duration. Therefore, a negative relationship between postponement and SHO suggests it is not that these items are not selling enough to deplete the units in dedicated space, but that they are possibly being proactively replenished within the store using inventory from shared space.

Hypothesis 3.4 is not supported although it has a significant and positive coefficient for the mediating effect of systemic drivers on SHO. Even though the direct effect of demand on stockouts (H3.1) is not significant, and the total effect is strengthened when including this indirect effect, the data shows the opposite of what is

hypothesized. This may be because H3.4 follows the reasoning that increased systemic drivers are associated with increased instore logistics postponement (H3.2). However, following the theoretical framework presented earlier, these results suggest store managers may value efficient use of shelf space for product assortment reasons instead of operational efficiency (less product handling).

Path	Hypothesis	Full Model	Results
Systemic drivers → SHO (direct)	H3.1 +	0.044 (0.252)	Not Supported
Systemic drivers → Postponement	H3.2 +	-0.009*** (0.002)	Not supported
Postponement → SHO	H3.3 -	-0.357*** (0.050)	Supported
Systemic drivers → Postponement → SHO	H3.4 -	0.003*** (0.001)	Not supported
SRMR		0.071	
CD		0.829	
Observations		12,109	

Table 3-5 Standardized path coefficients (standard error in parantheses)

3.4.1 Model respecification

Since 3 out of the 4 hypotheses are not supported, the measure validity tests for the constructs are at borderline test statistics, and the relationship between demand and SHO occurrence is not even significant, model respecification is warranted. Table 3-7 provides coefficient values, model fits, and AIC and BIC values to compare fit across 5 different respecifications of the same model. The only things that change from model to model are structural covariates. The first column is the original model specification whose results are already summarized. The second column is a model which has removed the control variables of store open minutes. Columns 3 through 6 are the different combinations of including or not including week and category dummies onto the model specification in column 2. Whenever the week dummies are included, they are included as Availability and Demand structural covariates, replacing the Christmas and New Year's controls. Whenever the category dummy is included, it replaces all but one vendor dummy (for direct store delivery) and the dedicated space (capacity) control as a Postponement structural covariate.

Specifying the model using different controls does change the fit of the model to the sample data, as in Table 3-7, and it does change outcomes in terms of support of hypotheses. Of the 5 models Model 4 has the best fit (lowest AIC=361,153.8). Model 4 removes the item-based SHO variables and the store hours control and instead adds week dummies. Compared to the original model, including week controls not only decreases the residual size (SRMR 0.071 → 0.030) but improves the explanatory power of the

model (CD 0.829 → 0.842). The week dummies were included into the Availability and Demand equations and improves the fit of each equation by approximately the same amount (R^2 0.668 → 0.686 for Demand and R^2 0.463 → 0.482 for Availability). The demand and availability relationship (H3.1) is supported in model (4) as well as every model except for model (1).

Observations: 12,109	(1) Original	(2) Removes weekly open minutes	(3) Includes week and category controls	(4) Includes week controls	(5) Includes category controls
Systemic drivers → SHO (direct)	0.044 (0.252)	0.076** (0.348)	0.098** (0.032)	0.098*** (0.032)	0.075** (0.035)
Systemic drivers → Postponement	-0.009*** (0.002)	-0.009*** (0.010)	-0.044*** (0.004)	-0.009*** (0.001)	-0.044*** (0.005)
Postponement → SHO	-0.357*** (0.050)	-0.620*** (0.072)	-0.532*** (0.060)	-0.563*** (0.059)	-0.597*** (0.063)
Systemic drivers → Postponement → SHO	0.003*** (0.001)	0.005*** (0.001)	0.024*** (0.003)	0.064*** (0.009)	0.026*** (0.004)
SRMR	0.071	0.077	0.031	0.030	0.078
CD	0.829	0.782	0.854	0.842	0.800
AIC	1,513,930.2	1,217,933.6	362,218.5	361,153.8	1,220,484.2
BIC	1,514,315.1	1,218,318.5	362,603.4	361,546.1	1,220,869.1
Equation-level goodness of fit (R^2)					
Systemic drivers	0.668	0.667	0.701	0.686	0.685
Postponement	0.114	0.115	0.192	0.115	0.196
SHO	0.463	0.327	0.484	0.482	0.328

Table 3-6 Model respecification, estimates and fit

The discriminant validity of the constructs is not an issue in either model (1) or model (4) despite different structural covariates. Table 3-8 provides the fitted covariances

between the latent variables as well as the AVE square root values of each measure on the diagonals of the covariance matrices. For each construct, the square root of the AVE is greater in value than any of the covariances the construct has with other constructs. For example, for Postponement, the square root of AVE is 0.885, which is greater than Postponement's covariance with Stockout (-.083 for model one and -.112 for model 4) and also greater than its covariance with Demand (-0.215 and -0.216). Even though the covariances vary from model to model because of different structural covariates, the discriminant validity of the constructs are sufficient in both models (Awang, 2012, p. 71).

	Model 1			Model 4		
	Demand	Postponement	SHO	Demand	Postponement	SHO
Demand	0.833			0.833		
Postponement	-0.215	0.885		-0.216	0.885	
SHO	0.174	-0.083	0.748	0.200	-.112	0.748

Table 3-7 Fitted covariances and square root AVE (diagonal)

3.4.2 Alternate specifications of instore logistics postponement

The model specification and respecification show an acceptable level of fit in explaining the variation in the observed data and give consistent estimates from model to

model in terms of support of hypotheses. However, there is a growing concern (Streiner, 2003) about the Cronbach's alpha in validating a study's constructs. Very high values of alpha (>0.90) do not necessarily reflect more reliable measures and indeed cause larger errors of estimates because the construct may be underrepresented (narrow range) or may have redundant indicators (Panayides, 2013). The very high Cronbach's alpha for the instore postponement measure (0.936 from Table 3-3) combined with the very low equation-level goodness of fit for Postponement (R^2 0.115 from Table 3-7) suggests that the construct may indeed be underrepresented. Adding in additional control variables increases the equation-level goodness of fit because there are simply more variables in the model.

An additional concern about the Postponement construct specifically is illustrated in Figure 3-6. As currently specified, instore postponement indicators assume the store is more likely to implement postponement as the difference between case pack size and shelf capacity increases. Figure 3-6 illustrates 3 boxes of identical case pack size. The case pack contents can be split, depending on the dedicated space set aside for the item, along a spectrum of possibilities. The possible scenarios in the illustration could be the only case pack for the item or one of a number of case packs for the item's dedicated space (the space being multiples of case pack size). However, these boxes represent the very "last" one that will fill the very last units of shelf space remaining for that item. At the very left of Figure 3-6, dedicated space compared to case pack size allows for all but one unit from the case pack to be shelved. This 1 unit remaining in the case pack is an example of inventory overflow. It does not represent purposeful use of both shared and

dedicated space. Similarly, at the other extreme with the case pack illustrated on the right of the figure, the dedicated space compared to the item's unit dimensions leave excess space that could fit one more unit on the shelf. To fill that single space the store could buy a "last" or "additional" case pack and put all but one unit of the case pack into shared space.

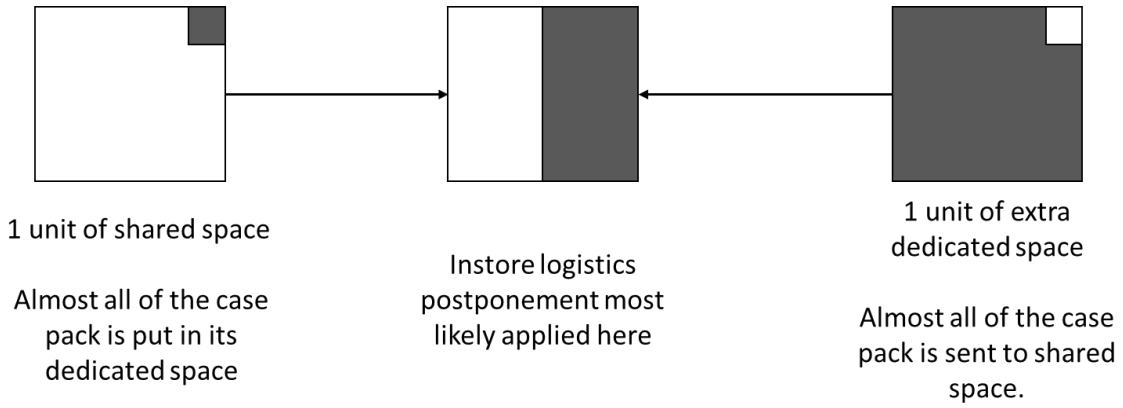


Figure 3-7 Internal postponement and the ratio of shared to dedicated space

In practice, it is highly unlikely that the store would decide to have higher inventory investment just to fill dedicated shelf space that is empty due to the dimensions of the item. With the Postponement construct, the indicator measuring the proportion of dedicated space to case pack size, assumes that this scenario is a more likely instance of internal postponement than the scenario illustrated by the box on the left. Unless completely out of additional shelf space for the item, the store would more likely add a facing or two to increase the dedicated space for the item. In this way, even if there were some units sent to shared space, there would at least be a greater proportion of the case pack on the store shelf. In other words, the figure on the very right of the illustration is

just as unlikely as that on the very left to be an example of instore logistics postponement. The scenario in the center of Figure 3-6, however, where the contents of the case pack are equally split between shared and dedicated space, is a dependable instance of purposeful use of both types of space.

Consider the following decision rules for a new 0/1 indicator, Post, for the Postponement construct, where 1 indicates postponement implementation:

Post=1	If shared units equals excess units Or if excess units is 50-80% of case pack size Or if excess units is greater than or equal to 20% of the dedicated space for the item
Post=0	If shared units or excess units is zero Or if excess units are less than 20% of dedicated space and less than ½ the case pack size

When the shared units for an item is equal to the excess units for that item in dedicated space, that means the case pack is evenly split between shared and dedicated space. If the empty or excess units of space on the shelf is more than half of the case pack size, then that empty space is likely going to be filled by a partial case pack. The 80% upper threshold mirrors the 20% minimum of the next condition which makes Post equal 1. That is, if at least 20% capacity of the shelf space is this “excess” shelf space. In all other cases, the Post dummy is zero. The 20-80% cut-offs may be tested with a more robust dataset having a greater variation of case pack and dedicated space values, but the ½ case pack size needs to remain the most likely scenario of instore logistics postponement. This decision rule results in the following breakdown of observations:

Table 3-8 Postponement dummy grouping of data (number of observations per cell)

Isotonics Yogurt

Post = 0	4,828	2,194	57.99%
Post = 1	3,755	1,332	42.01%
	70.88%	29.12%	

A few observations can be made about how the data is grouped with this new definition of postponement. First, the last row in Table 3-9, shows over 70% of the SHO data is for isotonics items with the remaining for yogurt. This difference is not because of the Post variable definition but simply because isotonics are tracked an average of 319 days while yogurt is tracked for roughly a third of that time (129 days). While the very right column does not show an equal partitioning of the Post dummy by category, it is very close to being equal. Yogurt shows a greater difference in subset size and this may be because of the nearly identical dimensions of yogurt items. One facing of a store shelf can hold a certain number of units of an item depending on that item's dimensions, and a case pack can hold a certain number of units of the same item depending on that item's dimensions and weight. Because of this, any data set will have observations of shared space units and excess dedicated space units occurring in certain combinations more than in others. In other words, the differences in breakdown between the two product categories may rest in the difference in item dimenstions. Similarly, the proportion of case pack size or of shelf capacity will also occur in certain combinations more than others. To illustrate, Appendix X lists two separate tables of the capacity and case pack combinations as well as the shared units and excess unit combinations of this study's data sample.

The Postponement construct is respecified using the Post dummy with similar CFA used to bring the original indicators in the measure together. This time the indicators are: Post, the proportion of case pack size that is in shared space, and proportion of dedicated space that is in excess (columns 3 and 4 of Table 3-4). The structural covariates also change to include: excess units, shared units, store dummies and a vendor dummy for a vendor which provides direct store delivery for its products. The same 5 model variations as in Table 3-7 is retested by running the models with this new Postponement indicator, Post. Table 3-10 shows the results from these model tests. The hypothesis support does not change at all between Table 3-7 and Table 3-10. The fit of each model in Table 3-10 improves as compared to its counterpart in Table 3-7, with lower AIC and BIC values. Model (4) still has the best fit in terms of smallest residuals and increased explanatory power as compared to the original model (1). However, the greatest difference is the explanatory power of the Postponement construct. For model (4) it goes from accounting for 11.5% of the variation of the data to 87.6% when using the Post dummy.

Since the explanatory power of the construct increases so much by changing one indicator, and since the indicator is time invariant, another respecification of the model is warranted. This time, instead of having each indicator come together as an Availability or a Systemic Driver construct the model specification will test for the same hypotheses as before, but with the indicators directly in the model as in Figure 3-7.

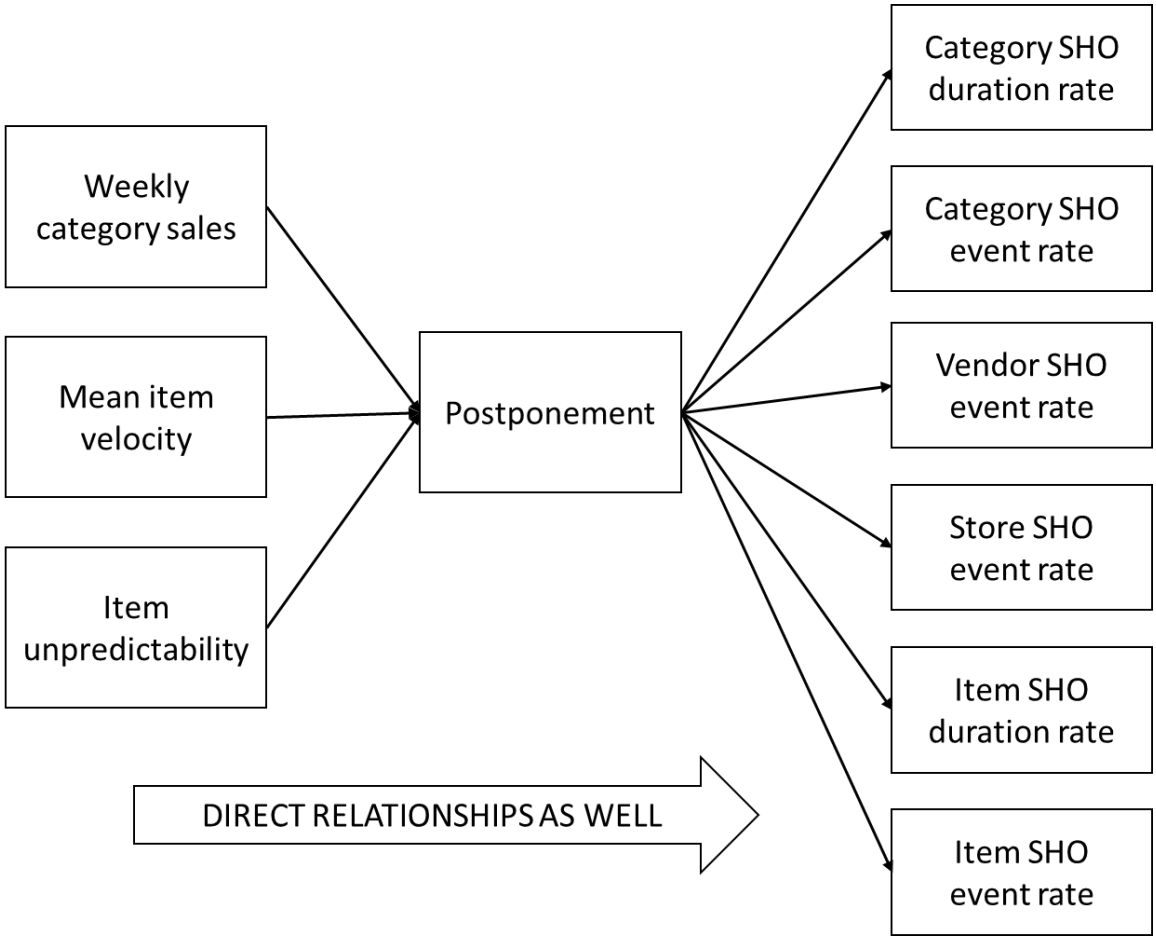


Figure 3-8 Generalized structural equation model

Observations: 12,109	(1) Original	(2) Removes weekly open minutes	(3) Includes week and category controls	(4) Includes week controls	(5) Includes category controls
Systemic drivers	0.002	0.007**	0.008**	0.008***	0.007***
→ SHO (direct)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
Systemic drivers	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
→ Postponement	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Postponement	-0.184***	-0.207***	-0.203***	-0.206***	-0.205***
→ SHO	(0.040)	(0.041)	(0.039)	(0.039)	(0.041)
Systemic drivers → Postponement → SHO	0.001***	0.004***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SRMR	0.056	0.060	0.025	0.024	0.063
CD	0.976	0.969	0.977	0.977	0.970
AIC	1,463,058.3	1,167,201.3	312,169.5	311,155.3	1,169,722.2
BIC	1,463,539.4	1,167,660.2	313,353.8	312,347.0	1,170,188.5
Equation-level goodness of fit (R^2)					
Systemic drivers	0.688	0.687	0.694	0.688	0.694
Postponement	0.876	0.876	0.877	0.876	0.877
SHO	0.483	0.342	0.506	0.506	0.342

Table 3-9 Models retested using Post (0/1 dummy) as a Postponement indicator

The specification of the model in Figure 3-6 is run with Stata's <gsem> command because both the endogenous and all of the exogenous dependent variables have different distributions. The endogenous dependent variable Postponement carries the 0/1 values of Post, so it is a logit model (Bernoulli family with logit link), while almost all of the duration and event rates (the exogenous dependent variables) are Tobit models and belong to the Gaussian family with left- (at 0) and right-censoring (at 100). Item SHO event rate is an exception to this because it is a count variable measuring the attribute of frequency (the other event rates measure breadth). The count variable is modeled with the negative binomial and log link, with constant dispersion assumed. All of the controls used in the model previously are included in each equation of this model as well. For visual simplicity's sake, the arrows from each demand variable (weekly category sales, mean item velocity, item unpredictability) to each stockout variable have not been included in the conceptual illustration although they are included in the model. Only the total effects (any indirect and direct effect together) are provided in the output. Indirect effects of demand variables on availability can be found by multiplying the parameter estimate for the demand to postponement relationship with that of the postponement stockout relationship.

While each individual coefficient estimate is not directly testing the hypotheses in this study, the signs and significance of multiple estimates together certainly shows a pattern towards supporting or not supporting the hypotheses. For the first hypothesis on the effects of systemic drivers on SHO (H3.1) there are three different coefficients for each of the six dependent variables in the availability equations (center of Table 3-11).

These coefficients are the total effect values, which was found to be significant (p=0.028) and positive in the prior specification (Figure 3-5). When using the systemic driver indicators separately as category sales, item velocity and unpredictability, the total effect of each on the SHO measures are almost all positive and significant. Notable exceptions are for the Category Duration and Store Event Rate dependent variables. The average item unpredictability has no significant link on either one of these aggregate measures while item velocity is linked to a significant increase in duration but not store frequency. Additionally, category sales are linked to increased store and category SHO frequency but decreased duration, so that stores on average are able to sell more when they can replenish both more quickly and more frequently. This suggests instore logistics postponement implementation and future research would help in determining how SHO event rates and durations are linked with one another to better explore their relationship with postponement.

Are systemic drivers linked to instore postponement (H3.2)? In the previous specification of the model the relationship was not supported with a negative instead of positive relationship as hypothesized. The theoretical framework for instore logistics postponement presented in Figure 3-2 said increasing systemic drivers should be linked to increased use of both types of store space. The separate indicators in the current model (Table 3-11) are all significant. However, negative coefficients for category sales and item velocity and a positive coefficient for item unpredictability suggest that systemic drivers differ among one another. The results show that more unpredictable items are more likely stocked on dedicated shelf space and in shared space as safety stock (top-

right corner of Figure 3-2), whereas higher velocity items and product categories are more likely stocked only on the dedicated shelf for maximum availability (bottom-right corner of Figure 3-2). In this way the current model supports the hypothesis of the positive link between systemic drivers and postponement while it also supports the previous model's results based on the Systemic Drivers construct. This more nuanced view in the current model shows competing demand drivers to postponement.

The postponement implementation and product availability relationship (H3.3) finds partial support in this current model (Table 3-11). The path coefficient in the prior specification (Figure 3-5) supported the theorized negative relationship between Postponement and SHO. In Table 3-11, the only significant negative parameter coefficients for postponement is at the category level, with a decreased SHO duration and frequency. Both of these lend support to the hypothesis. There is no support of H3.3 for the SHO attributes by store and by item. Further research categorizing items by high- or low-systemic driver classification could compare expected likelihood of postponement implementation with survey data from stores on whether both inventory space types are indeed used depending on item attributes.

On the other hand, there is a positive relationship, or increased SHO frequency for vendors when there is instore logistics postponement. The vendor SHO rate also had a significant positive link to all three systemic driver variables. This suggests postponement has a different mediating relationship with availability when considering the vendor as opposed to the category, store or item SHO frequencies. Although substitutability is not

within the scope of this study, the store receives planograms to implement based on prior sales and retailer negotiation with vendors. The total dedicated space per vendor is smaller than the total dedicated space for the product category and each vendor prefers to offer an assortment of items to capture a greater share of customers. Customers are likely to switch to items from the same vendor when faced with SHO of their initially preferred item. Therefore, the average item by vendor may appear to SHO more often when postponement is implemented with an emphasis on shelf space and a strategy leaning towards efficiency instead of availability (top-left corner of Figure 3-2).

Postponement equation (H3.2: higher demand drivers linked to postponing items)						
Category sales	-0.008*** (0.001)					
Item velocity	-0.044*** (0.003)					
Item unpredictability	0.079*** (0.008)					
Availability equations						
(H3.1 and H3.4: Total effects of demand on stockouts)						
	Category Duration	Category Event rate	Vendor Event rate	Store Event rate	Item Duration	Item Event rate
Category sales	-0.003*	0.097***	0.070***	0.014**	0.004	0.002***
	(0.001)	(0.008)	(0.010)	(0.006)	(0.008)	(0.001)
Item velocity	0.024***	0.072**	0.113***	0.0153	0.122***	0.008***
	(0.004)	(0.025)	(0.317)	(0.019)	(0.023)	(0.001)
Item unpredictability	-0.003	0.279***	0.251**	0.070	0.127*	0.021***
	(0.012)	(0.073)	(0.091)	(0.055)	(0.068)	(0.004)
(H3.3: postponement linked to lower stockout occurrence)						
	Category Duration	Category Event rate	Vendor Event rate	Store Event rate	Item Duration	Item Event rate
Postponement	-0.182**	-1.645***	0.996*	0.508	-0.457	-0.027
	(0.078)	(0.471)	(0.590)	(0.357)	(0.471)	(0.033)
Degrees of freedom	500					
AIC	448,726.5					
BIC	452,427.4					

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table 3-10 GSEM results

3.5 Discussion

Previous product availability research suggests that an item's case pack size and dedicated (shelf) space are important considerations in shelf and store stockout levels (Eroglu, Williams, & Waller, 2011). These item-based attributes are said to affect aggregate level availability measures because an item's dimensional characteristics make way for inventory overflow (Eroglu, Williams, & Waller, 2013), which then sets the store up for poor instore replenishment because of the backroom effect (Waller, Heintz Tangari, & Williams, 2008).

Previous postponement research suggests postponement and speculation can improve retail inventory management for online retailers. The general consensus, however, is that "…in traditional retail environments, […] inventory replenishment decisions are centered on either inventory location speculation *or* postponement" (Bailey & Rabinovich, 2005, p. 161). This study suggests that traditional brick-and-mortar stores do implement both speculation and postponement within their store walls. Instore logistics postponement is the purposeful use of both dedicated and shared space for an item and combines time and place postponement by speculatively holding more than the dedicated space allotted for an item, and delaying its forward movement to the customer by keeping units in shared space.

This study combines postponement and product availability streams to show that logistics postponement decisions do occur within a firm. The suggested mechanism

behind its implementation in a retail store, or instore logistics postponement, differs from the traditional trade-offs considered in interfirm postponement. A store has a strategy in maintaining an item, either towards efficiency or availability depending on that item's demand attributes (velocity and unpredictability). If the store's emphasis is on shelf space, it will opt for instore logistics postponement, instead of an emphasis on labor which would result in the item being stored in dedicated space only.

Does instore logistics postponement improve availability? The answer is complicated. The postponement measure in the study is at the item-level, and provides a measure for the relationship between case pack size and dedicated shelf capacity. Availability measures are at aggregated levels, either by vendor, category or store, as is convention in retail availability research. While the models presented here show a negative relationship between instore logistics postponement and SHO levels, further exploration with different structural covariates illustrates a more complicated picture. The decreasing explanatory power of the Postponement construct when a product category dummy is added suggests that the store does not decide to use both shared and dedicated space at aggregate levels---over an entire category or throughout the entire store at once. Additionally, due to the negative relationship between demand drivers and postponement implementation, the indirect effect of demand on stockouts through postponement is still significantly positive.

This study presents a conceptual framework which may partially explain such unexpected results. One of the two hypotheses that are not supported is that

postponement is increasingly implemented with an item's increasing systemic (demand) drivers. This in turn plays a part for the lack of support of the hypothesis of postponement's mediating the demand and stockout relationship. Stores appear to be following an efficiency strategy with items having lower demand drivers, with an emphasis on shelf space. While it does make sense to try to capture a greater portion of more predictable demand using a greater assortment of items, it can be more difficult to manage a larger number of items (Ton & Raman, 2010) within the same amount of dedicated space.

This study aims to spur discussion and further research into refining both the measure of instore logistics postponement and its antecedents as well as performance effects. The data sample used here does not group items as having high- or low-demand drivers. An empirical test using such groupings may be able to pinpoint which items also used shared space and their corresponding stockout levels. Other measures than the postponement indicators presented here may be developed. Using the shipment size of an item instead of the case pack size may provide a clearer picture of when an item is being internally postponed, instead of assumptions about when it is more likely that postponement is taking place.

Future studies should focus on item characteristics to categorize items into one of the four areas in the box of Figure 3-2. A study (Dubelaar et al. 2001) of demand uncertainty and inventory levels found no support for a relationship between the two, and noted that their measure of inventory levels did not differentiate between safety stock and

cycle stock. Although they did not consider sales velocity as well, their suggestion to separate inventory levels by type of stock was based on the reasoning that unpredictability increases safety stock. Velocity increases cycle stock, and the differences between these two stock types formed the basis of the conceptual framework in Figure 3-2. In this suggested framework, high-demand driver items where the store emphasizes shelf space has its safety stock in shared space. For items with low demand drivers, the shared space is used for both cycle and safety stock, to be able to increase assortment depth among these more predictable items. Categorizing items by both their velocity and unpredictability into a high- or low-demand driver grouping would allow for empirical testing on whether space or labor is emphasized by the store for each. It would also allow for empirically testing if stores implement different strategies (availability for high demand driver items and efficiency for low demand driver items) as Fisher (1997) observes for supply chains of innovative and functional items. Studying ways of categorizing items could be the first step in finding empirical support for the demand driver and postponement implementation framework presented in this study.

Another avenue of research is to explore what factors may affect the relationship between instore logistics postponement and product availability. For example, the differences between the physical and information flows that take place in the store has been presented in this paper. Simulations of generalized inventory record inaccuracy and its impact on fill rates, which is essentially a measure of product availability, have found a significant relationship between increasing record inaccuracy and decreasing fill rates (Nachtmann, Waller, & Rieske, 2010). However, the literature still lacks study of

inventory error specifically caused by the out-of-synch physical and information flows presented here. With shelf liner technology development it may be possible to get a clearer idea of how many units of an item has entered the store and how many units of it is in dedicated space. Using such information may aid in the study of any moderating effects record inaccuracy may have on the postponement and availability relationship.

Some of the limitations of this study can be improved upon in future work. For example, the first step of instore logistics activity for a store is placing an order (Curseu et al., 2009). Order size, even at an aggregate product category, brand or store-level was not available. Interviews revealed the stores receive orders every week and that the order size is meant to cover 7 days of sales. Because of this the model included average weekly item sales, generated from a 56-day sales data file. However, this value was not used as a proxy for order size. That may have been possible if inventory-level information (starting on-hand inventory or ending inventory, etc.) was provided, but even then the model would have to assume that item order size does not significantly vary from week to week. Although weekly category sales were provided, they were for the year prior to the data sample, and were included in the model to account for demand differences between stores and from week to week. Similarly, no information on how much safety stock or target service levels of items or product categories were provided. Although prior studies using case pack size and shelf capacity also do not include information on order size, future studies would benefit with some sort of measure of what actually enters the store.

Other limitations have to do with the data fields that were included. First of all, since this is a new shelf liner technology and has not been implemented throughout an entire product category or entire store, the data sample consists of only two partial product categories. One of the product categories, yogurt, has a single case pack size regardless of vendor or item. While the dedicated space for this product category does vary from 18 to 72 units, 90% of the 86 items have less than 54 units of dedicated shelf space. One quarter of the items (22 items) have 18 units allotted to dedicated space. Due to the limited range of both of these values, it is not possible to test the structural equation model's group invariance by product category. This suggests that the case pack size and shelf capacity relationship may only play a part in those product categories where both of these fields sufficiently vary. Secondly, the capacity data is in terms of units, as it is for case pack size. However, the overall shelf capacity itself is constrained and carries a different number of units depending on the item size. In other words, 10 units of a small item does not take up the same amount of shelf space as 10 units of a large item. It may be possible to include some measure of the *proportion* of shelf space an item takes up based on its dimensions, instead of on its units. The store is more likely to internally postpone items that are larger in size because they take up a greater proportion of their constrained shelf space. Whether the store internally postpones items more because of item dimensions or demand attributes may also be related to availability outcomes.

Chapter 4 Are sales greater for a product with frequent short stockouts or less frequent long stockouts?

This chapter focuses on SHO events themselves, taking advantage of a technology source which captures the actual SHO event in its entirety. Two different attributes of SHO are pulled from the event data, frequency and duration, and are linked to both SHO antecedents and its effects on sales. While antecedents and effects of SHO are normally two different streams of research, the SHO attributes used in this study are not estimated through sales or simulation methods, so that it is possible to link the attributes to one another, to sales, and to item-based (supply and demand drivers to SHO) differences. The demand driver is sales from the previous store week and the supply driver is the backroom effect, measured with the relationship between case pack size and item shelf capacity. This study contributes to item the availability literature by showing a nonlinear relationship between an item's SHO frequency and duration as well as a nonlinear relationship between these SHO attributes and sales, once SHO demand and supply drivers are controlled for.

4.1 Introduction

Since 1721, a "store" has meant a "place where goods are kept for sale" (Online Etymology Dictionary, 2016) so that having items available for customers is the core function of retailers. In 1916 the first self-service store Piggly-Wiggly opened its doors with today's "impersonal selling technique" (Regan, 1960) allowing customers to walk

around the store inventory and select items from shelves themselves without relying on a store clerk to bring a product upon request. This shift to self-service made it possible for a product to be stocked out or demanded by customers without the knowledge of store personnel, making the store a “retail black box” (Decker, Kubach, & Beigl, 2003) for management. Since the switch to self-service stores, there have been efforts to measure and control shelf availability and replenishment, to improve instore logistics processes and to manage product categories so that customers searching for items can find them on store shelves.

Despite four decades of research (Aastrup & Kotzab, 2010) on the retail shelf stockout (shelf-out or SHO) phenomenon, the global average stockout level of a store has remained fairly constant at 8% (Gruen, Corsten, & Bharadwaj, 2002). Depending on the product category, 4.9% to 12.3% of the items a store offers is not on the shelf at any given time (Gruen, Corsten, & Bharadwaj, 2002). Some studies suggest higher SHO rates, finding that most of the SHO are due to poor instore replenishment (Stuttgen, Boatwright, & Kadane, 2018). When an item is not at its dedicated shelf space, some customers may ask about the item’s whereabouts and store personnel may spend time looking for the SHO item. One study found customers lose 20% of their shopping trip time while stores lose $200 to $800 per week on labor (“lost time cost”) looking for the SHO items (Gruen & Corsten, 2008). Despite online spending reaching 9% of all purchases, since more than half of customers still prefer going to a store to buy an item even if it is also offered online (Stanley, 2016) new retail stores and chains continue to open (Buzek, 2017).

Consequently, the SHO phenomenon remains a relevant issue for brick-and-mortar retailers.

The effects of SHO are generally worsening retail performance. Industry studies find SHO events cause an annual worldwide loss of $600 billion dollars (Buzek, 2015) in retail sales. Scholars also find that market share decreases both in the short (Wu, Zhai, & Liu, 2015) and in the long term (Campo, Gijsbrechts, & Nisol, 2003). This may be because customers who face SHO switch stores at rates of about 50% (Zinn & Liu, 2008) to 81% (Corsten & Gruen, 2003). With repeated SHO experiences, customers lose store and brand loyalty (Turk, 2012). In some cases, however, retailers can take advantage of SHO of certain brands so customers switch to the store's own brand which has higher profit margins (Shah, Kumar, & Zhao, 2015). For this reason, research using SHO as the backdrop (antecedent) to customer loyalty and stockout-based substitution phenomena view SHO as something to optimize instead of merely minimize (Trautrims, Grant, Fernie, & Harrison, 2009).

Measuring SHO events has been the greatest stumbling block to research, requiring scholars to test relationships of observable attributes of SHO instead. Four general attribute types are: frequency, breadth, duration, and intensity (Gruen & Corsten, 2008). Frequency measures how often a SHO occurs for an item during a unit of time and is referred to as an item SHO event rate. The event rate for entire product categories, vendors, or stores measures the breadth of shelf-outs by counting the total number of items that are SHO at a moment in time and finding its proportion to the number of items

offered. Another proportional rate is the duration rate which measures the amount of time within the store's selling time where an item is SHO. Finally, stockout intensity attributes measure the amount of money or number of units of an item or entire baskets (Anderson, Fitzsimons, & Simester, 2006) of items which are lost due to SHOs.

Each attribute type is measured using different methods. Breadth measures are by far the most prevalent attribute used in studies (Corsten & Gruen, 2003) since it simply requires looking at a store shelf and counting the number of stocked out items. Multiple visits auditing the same shelf allows for use of frequency measures (Taylor & Fawcett, 2001), as do computer software simulating stockouts under various scenarios (Eroglu, Williams, & Waller, 2011). Simulation programs also allow for measuring stockout duration, which is the least often used measure in stockout research (Wu, Huang, Blackhurst, Zhang, & Wang, 2013). SHO intensity requires estimation methods using an experimental setup (Anderson, Fitzsimons, & Simester, 2006) or with point-of-sale data (Grubor & Milicevic, 2015), where it is assumed SHO occurs based on when item sales drop uncharacteristically. Point-of-sale data can also be used to estimate both the stockout duration and frequency within the same study (Gruen & Corsten, 2008), however such studies are limited to studying the antecedents of SHO, since the same sales data cannot then be used to study SHO effects.

Although Gruen and Corsten (2008) recommend using multiple attributes in a single study, such work is extremely scarce. Of the four attributes, intensity can be considered a consequence (a financial or unit-based loss) of SHO. As such it is most

often found with the attribute of SHO breadth (Anderson, Fitzsimons, & Simester, 2006) where customers who face SHO in a greater number of items in one period make fewer purchases in a later period (intensity or customer impact). Of the other combinations of attributes studied together, Campo et al. (2004) are often cited as having studied the effects of SHO frequency and duration. Their work surveyed customers on temporary SHO versus a retailer's decision to decrease product variety. Reduced assortment is similar to customers experiencing repeated (frequent) SHO (Kim & Lennon, 2011) or SHO which last a long time (Hendricks & Singhal, 2003), and they found that more customers would switch stores in cases of an assortment reduction than a temporary SHO. SHO duration and frequency are not measured, estimated, operationalized, or tested in their study. Indeed, they state that the effects of SHO frequency and duration attributes are still unknown (Campo & Gijsbreachts, 2005).

This paper brings together both SHO frequency and duration using pilot study data from one of a few different types of shelf technologies currently in development. Developing technology for "real-time notification" (Gruen & Corsten, 2008) of SHO reflect retailer efforts, such as Amazon Go (Liptak, 2017), to continually provide self-service shopping while also measuring and controlling shelf availability. Availability can be tracked using video imaging (Rosado, Goncalves, Costa, Ribeiro, & Soares, 2016) (Moorthy, Bhargave, Behera, Ramanathan, & Verma, 2015) to detect when items have been moved. Notification of completely empty shelves can use light emitting technology (Frontoni, Mancini, & Zingaretti, 2014) or weight-based sensors, as in this study. RFID technology can only track case packs and not individual units since it is cost prohibitive

(Gruen & Corsten, 2008) and mobile apps which require customers to scan in a product code when faced with a SHO is deemed too consumer-centric, as it relies on the shopper to alert the store (Moorthy, Behera, & Verma, 2015).

This work contributes to SHO literature by not only studying both frequency and duration at once, but by also studying the relationship between the two attributes themselves, their antecedents, and their impact on sales. First, after a review of duration and of frequency in extant work (Section 4.2) the paper presents a conceptual framework on how duration and frequency are affected by demand (4.3.1) or supply drivers (4.3.2), how the attributes relate to one another (4.3.3) and how they in turn are linked to sales (4.3.5). The method used to test the hypotheses is described in Section 4.4. Results (4.5) are discussed (4.6) in terms of SHO frequency and duration attributes followed by implications of where future SHO research may lead with actual SHO events being captured as in this study, and as scholars recommend (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016).

4.2 The duration and frequency of retail shelf stockouts

Since the bulk of SHO research is around the breadth of SHO occurrence in product categories or stores, the effects of item SHO duration and of SHO frequency on an item's sales has not been firmly established in empirical literature. A quick overview of two empirical studies of interest for these two attributes are summarized in Table 4-1.

There are other studies that mention SHO frequency (Waller et al., 2008) or duration (Rosales et al., 2018) in developing the theory behind their research questions or experimental design, but the summary table includes only those works that actually focus on frequency or duration as outcomes or antecedents to other performance measures.

Paper	Wu et al. (2013)	Eroglu et al. (2011)
Study focus	Effects of SHO duration	Antecedents to SHO frequency
Attribute measurement	Experimental factor Levels: 4, 16, 28, or 42 hrs.	Number of shelf stockouts per year
Dependent variable	Difference in market share for product or for store	SHO frequency
Independent variable	-SHO duration -proportion of customers with various responses to SHO depending on product category -original market share of store	-mean units sold per day -units of shelf capacity for item -case pack size in units
Analysis method	NetLogo discrete agent simulation (provides parameter estimates with significance levels)	ANOVA tests discrete event simulation outcomes
Findings	-SHO duration has a greater effect on market share difference than the initial market share -product and store market share outcomes worsen at different rates depending on product category	-demand strengthens the negative relationship between shelf space and SHO -case pack size strengthens the negative relationship between shelf space and SHO -there is a 3-way interaction among shelf space, case pack quantity and consumer demand

Table 4-1 Selected research on SHO duration and frequency

There are three main studies that refer to SHO duration. The first study is funded by Proctor and Gamble and reports the stockout duration pattern for 13 stores following an unknown number of product categories (Gruen, Corsten, & Bharadwaj, 2002). In that sample, one fifth of items are replenished less than 8 hours after becoming SHO, while another fifth are replenished after 3 days. Most SHO last between 1 and 3 days, although it is not clear if these SHO durations include closed store hours. Closed hours are included in SHO duration in a second study where SHO events are estimated from point-of-sale data (Gruen & Corsten, 2008). Point-of-sale (POS) data for 200 of the fastest moving items over two weeks in one US store shows 25% of SHO events last 3 days or more while 40% last 1 day or less. While this seems like a similar pattern as the first study, a second sample from a different retailer over 12 months and 100 items shows more than 90% of SHO last at least 3 days and very few items appear to be replenished on the same day (Gruen & Corsten, 2008, p. 18). Both of these often-cited works revolve around descriptive statistics of SHO attributes and methods of estimating them.

Performance in terms of product and of store market share is linked to SHO duration in the third study (Wu, Huang, Blackhurst, Zhang, & Wang, 2013), listed in Table 4-1. Wu et al. use SHO duration as an experimental factor and assign different levels of SHO duration in a series of simulation experiments. Experiments show that if a retailer reduces an item's SHO duration from 14 hours to 2 hours, the product market share of an item from the paper towel, salted snacks or coffee categories will improve even more than the improvements in store market share. Store market share will improve more for cosmetics and shampoo items, which researchers suggest points to another one

of their experimental factors: consumer response to SHO. For cosmetics and shampoo, the proportion of customers who would switch stores is set to be 1.5 to 2 times greater than the other product categories. Product categories with brand- or item-loyal customers, it is implied, are more important for a store in terms of focusing on the duration attribute of shelf-outs.

Two of the three main studies on SHO frequency do not directly test this attribute in terms of antecedents or effects. Corsten and Gruen mention SHO frequency in both phases of their research (2002, 2008). In the first phase of their study, they report that a greater proportion of surveyed customers would not buy an item at all if the customers face three SHO as opposed to the proportion of customers who would not buy the item after facing a single or two SHO (Gruen, Corsten, & Bharadwaj, 2002, p. 30). In this case frequency refers to the number of times a customer finds his preferred item stocked out, and it is possible that multiple customer SHO experiences are linked to one long SHO event. In the second phase of their study, SHO frequency is only directly mentioned as a SHO attribute which gives shelf-out rates at the item level. However, a look at POS-estimated shelf-outs over a store week shows different patterns of SHO events with implications about frequency. For example, a weekly chart of 21 weeks shows shelf-outs occurring twice a week (on Saturdays and Wednesdays). The authors state that the store has two deliveries per week but should have four. Another example shows an item stocking out every single day. Gruen and Corsten (2008) state higher weekly SHO frequency indicates insufficient dedicated shelf space for this item. While not studying

SHO frequency directly, their observations imply different antecedents to SHO are linked to different levels of SHO frequency.

The remaining work, listed in Table 4-1, on SHO frequency uses simulation to study the antecedents of SHO events (Eroglu, Williams, & Waller, 2011). SHO events occur more often for an item at different shelf capacities as demand rates increase or as case pack size increases. Case pack size and demand's moderating effects on the negative relationship between shelf space and SHO frequency is found through a simulation experiment. The experiment uses values from one store's cereal product category. The product category is audited by a third-party once a month for 24 months for 3 different items. The 3 items have different levels of daily demand (differing from each other by 2 units per day) and also differ in terms of the item case pack size and shelf space in units. Units of shelf space is considered to be a factor at 13 different levels (12 to 24 units) for a total of 117 different scenarios in the simulation experiment. Depending on the combination of demand, case pack size, and shelf capacity, the item may have additional units available in the store that do not fit on the shelf. Shelves are automatically replenished with these units and whenever the store's total inventory of the item reaches a reorder point, an order is placed. The experiment's scenario outcomes are the number of times an item SHO over ten years. ANOVA analyzes these outcomes to test the hypotheses of the moderating effects of case pack size and demand on the relationship between shelf capacity and SHO frequency. A three-way interaction between demand, case pack size, and shelf capacity is also found.

The simulation experiment does not entirely reflect a realistic environment in practice but serves to explore the relationship between case pack size, shelf capacity and item demand. In practice, store shelves and shipping boxes (case packs) are of standard dimensions and the number of CU that fit on the shelf or case pack depends on the item's dimensions. Single unit incremental changes for the item's dedicated shelf space or case pack size is not likely. For example, a cereal box may be placed on the shelf facing the customer with additional boxes, perhaps 5, behind it for a total of 6 boxes. Any additional facings would change the dedicated shelf space by increments of 6 instead of a single CU. Similarly, if one shipping box (case pack) can fit 6 boxes of cereal, the next incrementally larger shipping box will not usually fit exactly 7 boxes of cereal. Unless the manufacture purposefully makes shipping boxes incrementally larger for this cereal box example or for items of every dimension, the number of CU from one case pack size to another will also not vary by single increments. In this way the simulation study (Eroglu, Williams and Waller, 2011) appears to test a nonrealistic combination of case pack sizes and shelf capacities, but does so in order to show how the relationship between the two time-invariant item-based characteristics has an effect on an item's SHO frequency over time. Increasing shelf capacity for an item of a set case pack size decreases SHO frequency, while increasing case pack size for an item with a set shelf capacity increases SHO frequency (Eroglu, Williams and Waller, 2011, Fig. 2, p. 427).

4.3 Theory and Hypotheses

Consider a store using periodic inventory review, where orders are placed for items at regular time intervals and varying order size. The order size for each item may depend on a number of item- and store-specific factors, but that order size is chosen to ensure sufficient on-hand inventory at the beginning of an inventory cycle (in terms of meeting all or an acceptable proportion of estimated incoming demand) until the next order is received. Until the next order is received, whether a SHO occurs during the inventory period depends on customers taking units away from the shelves (demand) and store personnel putting units on the shelf (supply). The "black box" of the store consists of these supply and demand drivers to SHO. If the demand for the inventory cycle is greater than forecasted then this demand driver may result in a store stockout (STO type of SHO). Depending on the case pack size and shelf capacity relationship, the replenishment of the shelf may not properly take place (Waller et al., 2008) and this supply driver may result in an instore replenishment stockout (IRO type of SHO). This study considers both demand and supply drivers as antecedents to SHO to see its impact on store sales, as illustrated in Figure 4-1. A more detailed model is illustrated in Figure 4-3 at the end of this section.

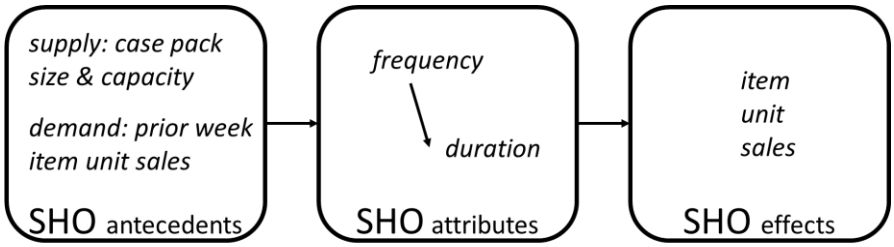


Figure 4-1 Study overview

4.3.1 Demand driver to SHO

If the demand during a store's inventory cycle is greater than forecasted for that cycle, SHO occurs for one of two reasons. The first is poor forecasting in terms of not being able to properly gauge cycle demand. While the impact of poor forecasting is beyond the scope of this study, a review of product availability has found that poor forecasting includes both under- and over-forecasting and that both increase SHO (Moussaoui et al., 2016). Their review includes an umbrella categorization of higher than anticipated demand as "systemic drivers": demand autocorrelation, unpredictability and velocity. Demand autocorrelation is the phenomenon observed in sales patterns over time where there is a similarity or connection between sales quantities in consecutive time periods (days, inventory cycles, or other time unit of analysis). Seasonal purchases or those during a store week with a promotion or discount are examples of auto-correlated demand. Moussaoui et al. (2016) state that not accounting for autocorrelation leads to poor forecasting. Zinn, Marmorstein and Charnes simulate different levels of demand autocorrelation and find "when (positive) autocorrelation is present but goes undetected or is ignored by management, customers will experience larger and more frequent stockouts" (1992, p. 182). Just like demand autocorrelation, increased unpredictability has been linked to increased SHO through poor forecasting (Moussaoui et al., 2016). Poor forecasting, for whatever reason (demand-based characteristics, computational limitations, software issues, personnel issues, etc.) is actually an out-of-store supply driver to STO and outside the scope of this study.

The second way demand during a store's inventory cycle can be greater than the available supply is the last demand-related systemic driver identified by Moussaoui et al. (2016), demand velocity. Demand velocity is the number of units an item is purchased over a specific period of time. In their review, Moussaoui et al. (2016) state that there is no consensus of the effect of sales velocity on SHO since some studies link increased velocity with higher SHO occurrence while others suggest high-velocity items are better monitored and replenished by store personnel and thus linked to lower SHO. The idea behind how velocity drives SHO is that the units of an item move so quickly off of the shelves that store personnel do not have enough time to replenish the items either from the backroom or from a store delivery. The difference between high- and low-velocity items gets exacerbated if the low-velocity item has "too much" dedicated shelf space and the high-velocity item has "too little" shelf space (Corsten and Gruen, 2003). If the items are somewhere in the store but are not available to customers because they are not in the space dedicated to them on customer-accessible shelves the sales velocity drives an IRO type of SHO. A STO occurs if the demand velocity is sufficiently large enough to exceed the store's entire stock of the item. In this case, it does not matter how quickly personnel are able to replenish the item, as there are no units of the item left in the store with which to replenish until the next store delivery.

The contradictory findings on the relationship between demand velocity and SHO may stem from which hypotheses were tested and the use of different attributes of SHO. Taylor and Fawcett (2001) count the number of different items that are SHO in a store's inventory and compare the proportion of inventory that is SHO between items that the

store is promoting and those that are not. This measure of SHO is a breadth attribute which Corsten and Gruen (2003) also use when finding that the proportion of faster-moving items that are SHO is greater than the overall SHO rates of the items in their study. They state that promotional items and fast-selling items largely overlap, as the promoted item attracts more customers than a regular item. DeHoratius and Raman (2008) measure inventory record inaccuracy as the difference between recorded and actual inventory levels of items. They take the total annual units sold of their data and find that the top 10% of unit sales are items that have the greatest inventory record inaccuracy. No SHO attribute is measured, but it is assumed the larger the record inaccuracy the more likely personnel will be unaware of items running low or being completely out of stock on store shelves. The Grant and Fernie (2008) study which is said to suggest lower SHOs with higher sales is a survey of four different types of retailers. This explorative qualitative paper reports that one of the retailers interviewed (a mobile phone retailer) stated that product availability has been an issue for promotional items and additional staff training is taking place for better promotions management. None of these studies actually test for the relationship between an item's speed of sales and its SHO occurrence.

The two arguments for whether demand velocity and SHO occurrence are positively or negatively related both consider instore replenishment behavior within a single period of time. For the argument that they are positively related, the speed at which units of an item are taken off the shelves (customer purchases) is faster than the speed at which store personnel restock the shelves during the same period of time. The instore

replenishment is done insufficiently, so that a SHO occurs when it need not have. For the negative relationship, store personnel know that an item sells at high velocity and pay special attention to it, dedicating more time and effort to replenishing the fast-moving item compared to the slow-moving item. Instore replenishment tasks are sufficiently carried out, and the store knows which item will have a higher velocity than another similar item from either store marketing efforts (promotions, discounts, etc.) or prior sales. The demand reshape (Eynan & Fouque, 2003) that results from store efforts is beyond the scope of this study, leaving prior period sales as an explanation for improved instore replenishment for faster-moving items.

Theoretically, there are two ways in which prior period sales can be linked to SHO. The first is the argument of improved instore replenishment behavior for fast-selling items. High velocity one inventory period could signal personnel to focus on the item's shelf availability the subsequent period. The second way prior period sales can be linked to SHO in a subsequent period is without store personnel even being aware of prior period sales. The combination of an item's dedicated shelf space size, case pack size, and units sold (Eroglu et al., 2011) may lead to inventory overflow and the backroom effect (Waller et al., 2008). The backroom effect is an increased occurrence of IRO because personnel are not aware or able to move the backroom inventory of an item onto customer-accessible shelves.

This poor instore replenishment behavior may be more likely in one period if the store orders (in multiples of case pack size) and the prior period's item sales are low. It is

more likely that a greater number of units will be put in the backroom for an item with low prior weekly sales as compared to high prior sales. Given the same shelf capacity for two items, the item with lower sales will have more units of stock left over from one inventory period to the next. Using the same logic, an item with higher prior period sales would have lower inventory overflow. This in turn would mean fewer units in the backroom than for lower prior period sales items. Even if store personnel were focused on product availability and instore replenishment there may be more stock in the backroom for lower sales the prior period than higher sales. In such a case fast-selling items may not have improved instore replenishment, leading to the hypothesis:

Hypothesis (H4.1): An item's sales one period is positively linked to its SHO frequency in the subsequent inventory period.

4.3.2 Supply driver to SHO

Customer supply of an item delivered to a store is governed by when and how store personnel stack the item onto customer-accessible shelves in the store's selling area. Selling-floor shelves have assigned product category areas and dedicated space for each item. The capacity of an item's dedicated space is most often (91% of time according to Gruen and Corsten, 2008) determined by the item's case pack size. A single or multiple case packs of an item may be ordered by the store based on the item's shelf capacity and depending on the store's replenishment strategy and resource emphasis (Ch. 3, Figure 3-2). A resource emphasis on shelf space (instead of labor) means the store uses both

shared (backroom) and dedicated shelf space for an item. On the other hand, a resource emphasis on labor means the store avoids instore replenishment and increases shelf capacity to match order size.

Not being able to observe store personnel replenishment behavior, researchers have used the item attributes of case pack size and shelf capacity as measures of whether ordered items delivered to the store completely fit in dedicated shelf space or require instore replenishment (Ch. 3). Waller et al. (2008) show that a retailer's market share decreases in relation to case pack size because of inefficient instore replenishment (the backroom effect). If the retailer receives a large number of units of a product that does not all fit in its shelf capacity, then the backroom effect negates any gains from a store-level fill rate effect, which is increased product availability from instore replenishment. Worsening product availability is found with decreased shelf capacity and increased case pack size (Eroglu et al., 2011). Decreasing shelf capacity whether or not there is instore replenishment means that fewer units of an item are available to the same volume of customers, increasing the likelihood of SHO occurrence. If there is instore replenishment, then decreasing the shelf capacity suggests multiple points during the inventory cycle for SHO events unless personnel proactively replenish shelves. Regardless of personnel behavior, the relationship between SHO events (frequency) and capacity can be hypothesized as:

Hypothesis (H4.2a): Item shelf capacity is negatively linked to that item's stockout frequency.

While increasing shelf capacity may lead to fewer SHO events, its impact on SHO duration is unclear. Even if the dedicated capacity for an item were increased to take up an entire product category space, if personnel do not have the item on-hand, they cannot place it in that dedicated space. Any benefit from increasing dedicated space to avoid the possibility of SHO occurrence from inefficient replenishment is constrained to order size and on-hand inventory. Increased case pack size, on the other hand, means that even if the order size is of a single case pack quantity there will be more units of the item in the store. Given a constant shelf capacity, increased case pack size means increased opportunity to replenish the store shelf within the inventory period without waiting for the next store delivery. In cases where the manufacturer increases case pack size and the retailer does not increase shelf capacity, the store is more likely to be replenishing dedicated shelf space with instore stock. Efficient instore replenishment leads to:

Hypothesis (H4.2b): Item case pack size is negatively linked to that item's stockout duration.

Considering the interaction effects of changes to case pack size and shelf capacity brings inventory review types and instore replenishment into question. The question is whether instore replenishment purposefully and efficiently takes place in between store replenishments. Eroglu et al. find the interaction effect decreases SHO frequency because the decreasing capacity has a larger impact for smaller case pack sizes than larger case pack sizes. They reason: "as case pack quantity decreases, supplier replenishments become more frequent, which decreases shelf stockouts" (2011, p. 428). Their study has a

reorder-point (ROP) inventory system, where the store places an order whenever the inventory position has decreased to a certain level instead of at set order intervals. They also assume instore replenishment is not planned so that the entirety of deliveries to the store are placed onto dedicated shelf space. Any inventory overflow results in mismanaged shared space inventory which is not efficiently used to restock shelves. Avoiding inventory overflow with more frequent smaller shipments and larger shelf capacity thus is linked to fewer SHO.

Hypothesis (H4.2c): Case pack size strengthens (negatively moderates) the negative relationship between capacity and stockout frequency.

Similar logic may be applied for case pack size and shelf capacity interactions for stores with periodic inventory review. With set inventory order points in time, order frequency would remain constant despite decreasing case pack size. Since order size may vary from week to week, any decrease in case pack size could mean more case packs per order than at prior (larger) case pack size. An increase in shelf capacity may not be needed to be able place the entire store delivery onto the shelf. Any increase in case pack size could mean an increase in total units delivered for multiple case pack orders of an item, depending on the static capacity level. Increasing case pack size is hypothesized to be linked to decreased SHO duration (H4.2b) since a larger case pack size means more units of the item in a single-case pack store order. Having more units at the store with increased case pack size and being able to stack them on customer shelves would also minimize inventory overflow as capacity increases. This leads to:

Hypothesis (H4.2d): Capacity negatively moderates the relationship between case pack size and stockout duration.

4.3.3 Relationship between SHO attributes

This study not only contributes to product availability literature by empirically testing supply and demand drivers to SHO, but also tests the relationship between two different SHO attributes: item SHO duration and frequency. There has been a call (Gruen and Corsten, 2008) for using multiple SHO attributes at once as a best practice in the retail industry. Use of multiple SHO attributes allow retailers to "determine the overall impact" and root causes of SHO events within product categories (Gruen and Corsten, 2008, p. 13), but the relationship between the SHO attributes themselves have scarcely been discussed in practice or in the literature. This is surprising because every time an item stocks out during a store week the increasing SHO frequency requires that total SHO duration also increases. However, the total weekly SHO duration can increase without increased SHO frequency. The SHO attributes are clearly not interchangeable with one another.

Consider a SHO event, illustrated as shaded areas in Figure 4-2, occurring within a period of time of interest, such as the working (open) hours of a store day or week. The left-most boxes in Figure 4-2 show an item having a SHO with a short (top box) or a long (bottom box) duration within that time period. The next set of boxes show an increase in SHO frequency with the same SHO event duration as in the left-most boxes. There are

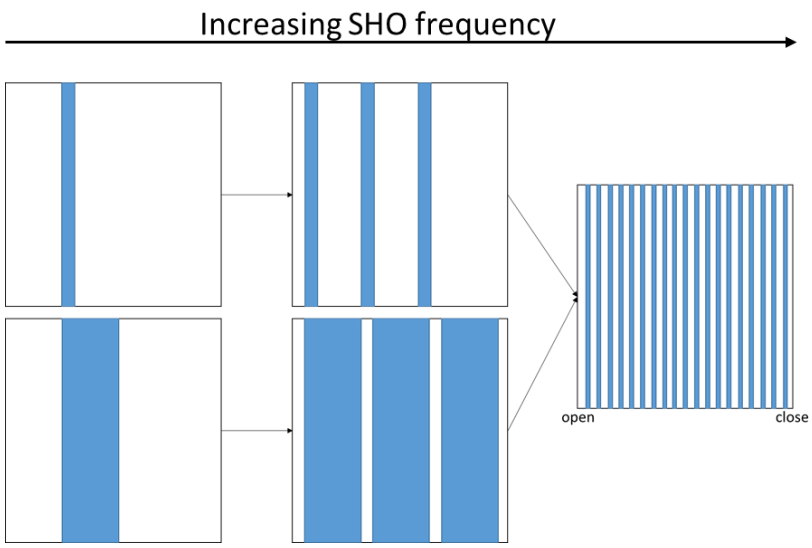


Figure 4-2 The relationship between SHO duration and frequency

now three SHOs of short or long duration. Increasing SHO frequency even further in a limited period of time results in shorter durations per SHO event. Indeed, if time is a constrained quantity like shelf space, then increasing total SHO frequency requires decreasing total SHO duration, just like increasing product variety on store shelves, requires decreasing shelf capacity for each item. This suggests a nonlinear relationship between the two.

Additionally, the same store practice may be linked to SHO frequency and duration attributes moving in opposite directions. For example, instore logistics postponement (Chapter 3), or the purposeful use of both the shared backroom space and dedicated shelf space, enables replenishing store shelves without having to wait for a new shipment from the supplier. Instore replenishment would thus reduce SHO duration. On the other hand, instore replenishment may increase SHO frequency, especially if shelf capacity is only partially replenished from the backroom stock. If a store practice such as instore logistics postponement may lead to an increase in one SHO attribute and a

decrease in another, then the link between each SHO attribute and store performance measures may also be in opposing directions. Since studies normally do not include more than one SHO attribute at once, it is possible that when there is a positive link between SHO frequency and sales, there may be a negative link between sales and SHO duration that has not been uncovered. This is why establishing the relationship between attributes is important, leading to:

Hypothesis (H4.3): There is an inverted U-shaped relationship between an item's weekly stockout frequency and duration, such that weekly total duration increases at low stockout frequency level and declines past a certain point.

4.3.4 SHO effects on sales

Using more than one SHO attribute in a single study not only allows for studying how they are linked with one another but also testing any direct and interactive effects on item sales as illustrated in Figure 4-3. Research on the effects of stockouts on item sales have generally found increased stockout occurrence linked to lower sales (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016).

Increased SHO attribute of…	May be a symptom of…
IRO duration	(1) mismanaged inventory in shared space, or (2) insufficient labor to replenish shelves, or (3) not realizing SHO has occurred
STO duration	(1) late deliveries to the store, or (2) earlier SHO than expected (spike in demand)
IRO frequency	(1) earlier SHO than expected (spike in demand) and (2) insufficient dedicated space capacity
STO frequency	(1) increased store deliveries and (2) insufficient order size

Table 4-2 Different possible causes among SHO types and attributes in this study

In theory, SHO duration and frequency may be longer or occur more frequently for different reasons depending on whether the SHO is a STO or IRO, as listed in Table 4-2. For IRO events, there are still units of the item in the store, somewhere other than in its dedicated shelf space. IRO duration could be a symptom of any one or a combination of: not knowing where additional units of the item are located within the store (to be able to use them for replenishing dedicated space), not having enough (or quick enough) labor to replenish the shelves, and not being aware that the shelves need replenishing. STO duration, on the other hand, could be driven from increased lead time and higher than expected sales during lead time (stocking out earlier than expected so that SHO duration

is longer even if items are delivered to the store when expected). IRO frequency could be driven by higher than expected sales between instore replenishments but also due to a shelf capacity that is too small, requiring more frequent instore replenishments. For STO frequency to increase the frequency of shipments to the store must increase and the quantity of each shipment must be below inventory period demand. Overall, IRO duration and STO frequency appear to depend on supply issues, STO duration on supply or demand drivers, and IRO frequency on both supply and demand.

Adding to the complexity of different drivers for different SHO types and attributes is that the inventory management and deployment of items within and across product categories are not identically implemented. Within a product category, there could be one or more items that are managed by the store's vendor. The on-shelf availability of those items may be checked more or less often than other items. Stores may have their own brand items to which more space is allocated than to vendors who must pay for the store's dedicated shelf capacity. Some items may have a higher sales velocity than other items despite having the same shelf capacity. Store managers may direct personnel to pay special attention to specific brands or promoted items more so than the rest of the product category. Lead time could vary by item, depending on the supplier, unless the store receives frequent shipments from a retail distribution center. Although the volume of space on the shelf for two items may be the same, a larger item may have a smaller shelf capacity, or number of units that fit into the dedicated shelf space. The larger items may be replenished (either by store delivery or instore replenishment) more often because there are fewer units of it compared to a smaller item

with the same dedicated shelf area. All of these differences between the management of items may amount to a change in a SHO attribute for one item and not the other, so that the overall effect on an average item may not be clear.

For SHO duration the link to sales may appear clear on the surface. If an item is not available on the shelf then customers cannot buy it. The longer the period of time that the item is unavailable, the greater the number of the customers will face a SHO. However, given that STO duration could be driven from greater than expected sales, sales may be higher for items with a longer STO duration. Presumably the store holds safety stock, and in cases of STO (instead of IRO) the safety stock has been purchased along with the cycle stock. In that case a longer SHO duration would point to more unit sales than with no SHO duration at all. On the other hand, IRO duration for a promotional item may be shorter due to a focus on replenishing that item more so than regular items or may be longer if there is insufficient labor resources for replenishment. If a store temporarily increases the shelf capacity of a promotional item (perhaps by also having items at the end of aisles or in temporary endcaps or sidekick displays) in anticipation of increased sales, then it may be less likely to SHO so that the SHO duration is shorter than for items with lower sales volume. This leads to the hypothesis:

Hypothesis (H4.4a): There is a curvilinear relationship between an item's weekly stockout duration and its sales, such that sales increase at low stockout duration levels and then increase at a decreasing rate past a certain point.

In contrast, SHO frequency may differ depending on whether the SHO is an IRO or a STO and may moderate the duration and sales relationship. First of all, an increase of SHO frequency from zero to one necessitates a nonzero SHO duration so that customers may arrive during a SHO event and leave without item purchase. Such lost sales would mean a negative relationship between SHO frequency and sales because of the increased (nonzero) SHO duration, regardless of whether it is IRO or STO. Short, rare SHO occurrences throughout the inventory period signal instore replenishment, suggesting higher sales with increased SHO frequency. Indeed, increasing SHO frequency in a periodic review system points to IRO events, as store deliveries do not increase during the order cycle. In an order-up to system (ROP), store deliveries may occur more frequently so that increased SHO frequency may be of STO type. Increased STO frequency with long duration would be expected to have the greatest loss in sales as both SHO attributes are high.

But what of short, frequent SHOs and long, seldom occurring SHOs? Given the hypothesized negative relationship between SHO duration and sales, increasing SHO frequency could decrease sales at a higher rate than at lower SHO frequencies. For example, consider a weekly SHO duration of 10 hours. In one scenario this SHO duration may occur over two separate events or 5 separate events, or any combination of lengths totaling 10 hours. In the 5-event case, each individual SHO event is preceded and followed by product availability so that the 10-hour weekly duration has fewer consecutive SHO hours than the less frequent SHO case. In this more frequent scenario, with SHO hours more dispersed throughout the store week, it is more likely for a

customer facing a SHO earlier in the week to return to the store later in the week and face yet another SHO. Even without returning customers, since the customer traffic of a store is not constantly distributed (some hours, like after work, are busier than other store hours) having fewer longer periods of SHO may coincide with both high- and low-traffic hours, whereas many shorter SHO periods would more likely coincide with high-traffic hours. The increase in SHO frequency could thus:

Hypothesis (H4.4b): Increased stockout frequency negatively moderates (weakens) the positive relationship between stockout duration and sales.

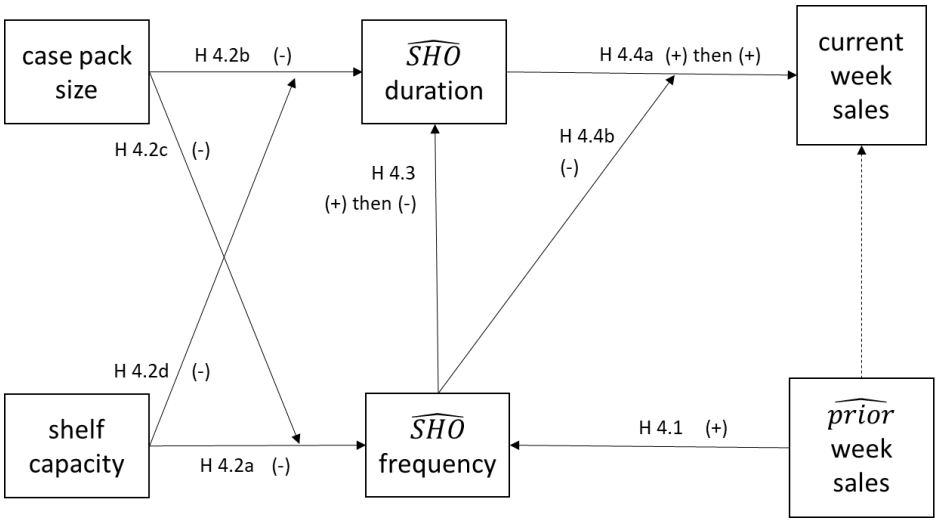


Figure 4-3 Hypothesized model of SHO attributes, antecedents and effects

A summary of all hypotheses in this chapter is illustrated in Figure 4-4.

4.4 Method

4.4.1 Sample and Procedure

Data for the stockout events are from a 56-day (November 8, 2015 through January 2, 2016) pilot study at 4 stores of a new shelf liner technology tested on 148 unique items in two product categories (yogurt and isotonics). During this time period yogurt has 86 unique items being tracked which accounts for 20% of that entire product category's assortment and about 37% of its item movement. 70% of the isotonics assortment is tracked with 62 unique items constituting about 94% of its item movement during the time period of the data sample. While the data spans a period of 56 days, all four stores are closed on Christmas day so there are only 55 daily observations for each store's items. Although this study's data sample comes from the same shelf liner technology's pilot study, the number of items in each category differ than those in Chapter 3. In Chapter 3, the data sample spans 11 months of time instead of 56 days. During the 56 days 148 items are tracked in all four stores. All four stores do not track all of these items beyond the 56-day pilot period, which explains why this chapter has more items for yogurt and isotonics than the prior chapter (54 and 48 items, respectively). While the number of items involved differ, data is from the same four stores as in Chapter 3.

The data is gathered from four different sources, as outlined in Table 4-3. Planograms obtained from regional retail managers provide the layout of each item

within each product category at each store. Each store offers the same items but may carry them at different amounts and place them at different shelf heights within the

	Description	Information gathered
Manual gathering	Store visits, phone calls, e-mails	-Store hours, including changes to store hours, holiday and weekend hours -Product placement on the shelves: facings, stacking and depth of shelves
Planograms	Yogurt and isotonic drinks shelving plans for each store	-Product shelf height (base case = bottom shelf) -Case pack size -Facings and stacking
Point of sale	Item movement totals for every day for SKUs that are being tracked in the shelf liner, in every store	-Daily total units sold of each tracked SKU
Shelf liner	Shelf stockout and replenishment pings of SKUs with shelf liner technology	Stockout (0 units on shelf) and replenishment (non-zero units on shelf) date and time for every occurrence Dates: November 8, 2015 – January 2, 2016

Table 4-3 Data gathering methods

product category. Manual gathering includes store visits, calls and e-mails to gather information on store operating hours and product placement on the shelves. The store operating hours are needed for two purposes: (1) reduce recorded SHO durations to when the store is open, since customers have no access to the item when the store is closed, even if the item is in stock; (2) calculate the selling time for each store per week as this

value is used as the denominator in the SHO duration measure. Product placement on the shelves refers to not only checking how well planograms are implemented but also to get information on how many units deep into the shelf each item can be placed in order to generate total dedicated shelf space for each item. If the planogram indicates an item stacked 2 units high but the manual audit shows 3 units stacked on top of each other, then the manually gathered data is used. There were no differences between planograms and on-shelf observations of the number of facings of a product or its shelf height. Item sales are provided from store managers and show the daily sales for each tracked item, whether or not it has stocked out at all during the time period of the study. Last of all, the stockout attributes are generated from stockout events of tracked items which have observations of paired time stamps recording when the dedicated shelf space for the item becomes completely empty and when the store personnel have replenished that shelf space. The company developing this proprietary shelf liner technology provided these paired stockout and replenishment observations.

The following process brings these data sources together to form the study sample. First, the item sales file serves as the base of the data since it includes all of the items being tracked. It has 33,376 daily observations as a balanced panel, with 8,065 observations (24%) having zero item sales. This file is merged with a cleaned shelf liner data of stockout event observations. The stockout event observations, prior to merging, have been coded to become daily observations of stockout durations during each store's open hours. Each store's open hours, or selling time, is important because recoding the original stockout event file to count the stockout duration during store open hours

lowered stockout duration by 30% on average. One item is dropped at this time because it was not tracked on the shelf and was incorrectly included in the sales file. Of the remaining 33,152 daily observations, 2,754 observations or 8% of them stockouts. These are not individual stockout events, as an event can last up to 7 days, but of the 33,152 store-item days, 8% of those items have some period of stockout, which coincides with a meta-analysis on average stockout rates in terms of SHO breadth (Gruen, Corsten, & Bharadwaj, 2002, p. 13).

	SHO duration (percentage)	SHO frequency (count)	Case pack size (units)	Shelf capacity (units)
Mean	2.68	0.62	11.00	26.42
Std. Dev.	7.24	1.24	3.78	14.08
Min.	0	0	3	4
Max.	80.94	9	15	72
Observations	4,736			
SHO duration	1.00			
SHO frequency	0.7479	1.00		
Case pack size	-0.1336	-0.11	1.00	
Shelf capacity	-0.0692	-0.03	0.51	1.00

Table 4-4 Descriptive statistics and variable correlations for sales equation

Once the two files are merged, the data is collapsed from daily to weekly observations. The main reason for collapsing data into weekly observations is because of the stockout frequency variable. On a daily basis, it shows items have a mean frequency of 0.089 stockouts (SD 0.305) with a minimum of 0 stockouts occurring 90% of the time, and a maximum of 3 stockouts per day. In contrast, the weekly descriptive stats and correlations are posted in Table 4-4. Collapsing the data over weeks brings down total observations to 4,736, with 1,369 weekly observations (28.91%) having some duration of

stockout occurrence. This process also illustrates why Gruen and Corsten (2008) emphasize that retail stockout studies must clearly address what attribute of stockout events are being measured and over what unit of time.

The two SHO measures generated for this study are the item SHO event rate and the item SHO duration rate per store week. The SHO event rate is a measure of the frequency attribute of a SHO event and is simply the sum of daily SHO occurrence by item and week for each store. If a SHO event starts one week and continues into the next week, it is counted for both weeks. The duration rate is the proportion of weekly store selling time where the item is stocked out. For the stockout event observation spanning two store weeks, the actual portion of selling time that the item is SHO in each week would be included into that week's duration only. In this study's sample, as seen in Table 4-4, items are never stocked out for the entire store week and most of the observations have 3 or fewer stockouts. Appendices IV-VIII provide further descriptive statistics for the entire shelf liner data used in both Chapters 3 and 4.

4.4.2 Model Specification

Since both the antecedents and effects of stockout are included in this model, and since the variables are all either censored or count variables and due to their endogenous and multicollinear nature, the equations in this model are run simultaneously as

generalized structural equations (GSEM). GSEM allows for identifying dependent variables that are endogenous and have varying distributions. It is possible to identify sales and frequency with the negative binomial family (where the "family" is the distribution) and log link (where "link" is the way the expected mean relates to independent variables) while having the stockout event rate with a Gaussian censored distribution (a Tobit). Additionally, GSEM produces parameter estimates for case pack size and shelf capacity even though these variables do not change over time and are included to represent any instore replenishment or inventory overflow effects over the 8 weeks of the data (see Appendix XI for panel data testing). GSEM allows for studying a complicated system of paths between variables in a multilevel simultaneous manner where the residuals from each level (each equation) are independent and normality of the dependent variables need not be assumed.

The first step in the GSEM analysis for this study involves looking at SHO frequency, duration, and weekly sales variables individually. The frequency is predicted using a negative binomial regression (see Appendix XIII) with a lagged weekly sales, case pack size, the interaction of pack size and shelf capacity. The sets of dummy variables control for differences between stores, between vendors, and during what proportion of the store week the personnel were alerted to SHO. SHO alerts require that the shelf liner is active and sends pings to store personnel's cell phones. The liners are not activated on the same day for each store and the data has either: 0 days (62.50% of the observations), 1 day (3.13%), 5 days (3.13%), or 7 days (31.25%) active liners. There are 3 binary dummy variables for these alerts with 0 as the base case where the shelf liner

does not alert personnel at all (0 days active). The frequency equation is exposed over open store week hours and has clustered standard errors over each SKU. Clustering over each item requires independent observations from item to item, while relaxing the independence requirement within item data subsets. Clustered standard errors accommodate for heteroscedasticity and autocorrelation. The exposure option is necessary for the negative binomial because the dependent variable of SHO frequency is otherwise assumed to count the number of SHO events during the same period of store hours over every observation. Store hours vary by store and by week within store so the exposure option essentially weights the SHO frequency counts based on the period of total time SHO events could have occurred. Not accounting for different store week hours would bias results since some items have higher and some would have lower non-normalized counts.

Since there is a squared term of SHO frequency in sales and SHO duration equations, the predicted SHO frequency is squared to generate the squared frequency term. This process is repeated for the Yogurt and Isotonics subsets. The lagged weekly sales variable is regressed with negative binomial, with the same exposure and clustered errors. The observations are not lagged, but the predicted weekly sales variable is lagged in the main model. To obtain the predicted weekly sales, the right-hand variables are all fitted from the prior frequency and duration equations: the interaction between SHO duration and frequency, and the linear and quadratic terms for SHO duration. There are also 6 control variables: regular price of item, percent discount, shelf height, store, week and vendor dummies. The sales equation is also predicted overall and as category subsets.

The second step in the GSEM analysis for this study involves creating the paths of equations (Table 4-5) that are considered simultaneously in the model. Each path or level of equation consists of one dependent endogenous or exogenous variable and more than one right-hand side variable and controls. The endogenous variables of duration and frequency from the first step are predicted and fitted as right-hand side variables in this second step. The squared frequency term is first predicted then squared, for the overall sample and each subset. Capacity and case pack size are included into the SHO attribute equations as supply drivers to SHO and the demand driver to SHO, lagged sales, are also predicted values from the first step that are then lagged in the second step. While multiple case packs can be included for shipments arriving to the store from a vendor, the model considers the number of units in a single case pack for each item. The model drops case pack size for the yogurt category subset as it does not vary by item. The interaction term can thus be interpreted as how the SHO attribute changes with increasing shelf capacity and a case pack size of 12.

There are also a number of control variables, listed in Table 4-5, for the different levels of the GSEM. Percent discount is calculated by dividing the difference between an item's regular retail price (obtained from multiple websites) and daily unit price (the total daily revenue divided by daily units sold for the item) by the regular retail price. Items can be purchased at a discounted price in multiple ways: by offering store "loyalty" membership discounts, promotions ("buy 10 for $10" or "buy 2 get the 3rd free", etc.), and manufacturer's coupons. Because of this, the calculated daily discount amount of the daily observations show different percentages within a store week, whereas normally

there should be a single promotional discount from the beginning of the store week on Thursday until the end of the week on Wednesday. The lowest discount percentage (corresponding to the highest unit price below the regular retail price) for each store's item during each week is then recorded as that item's weekly store discount. Item discounts vary by item store weeks and are included as a control in the sales level (equation) since they are linked to increased demand for an item (Borin, Farris and Freeland, 1994).

Path to	Type of variable	Independent variables	Controls	Model, comments
Weekly sales	Exogenous, unit sales, sum of all units for an item sold during store week	SHO duration, Squared SHO duration, interaction between frequency and duration	Regular item price, Percent discount, Shelf height dummies, Store dummies, Week and vendor dummies	Negative binomial
SHO duration	Dependent endogenous, proportion of store hours where the item is SHO	SHO frequency and its squared term, Item's case pack size and its interaction with capacity	Proportion of store week staff alerted to SHO, store and vendor dummies	Gaussian, left censored 0, right censored 100 (Tobit)
SHO frequency	Dependent endogenous, count of SHO events per store week	Lagged fitted item sales by store week, shelf capacity, and its interaction with case pack size	Week and vendor dummies	Negative binomial

Table 4-5 GSEM levels of equations

The time-invariant controls are included in the model for different reasons. The shelf height is a set of four binary (0/1) dummies with the bottom shelf being the base case where all dummies are zero. Shelf height dummies are in the sales equation since there is a link between what item the customer is more likely to buy and its vertical shelf position (Wongkitrungrueng et al., 2018). Vendor dummies (4 of them) are included to capture any differences by manufacturer of items. The isotonics category has only two vendors, both of which participate in direct store delivery (DSD), and the remaining 3 vendors in the yogurt category do not participate in DSD. The item regular price variable is determined by taking the largest price found for an item through online search and varies from 70 cents to $9.99. It is included as a control for customers choosing between two similar products based on price difference. For example, a store-brand strawberry yogurt may be preferred over a well-known yogurt brand of strawberry yogurt for price-conscious customers who prefer the lower-priced store brand item. Item characteristics are often included in retail category management studies, even though they are time invariant (Frontoni, Marinelli, & Rosetti, 2017) (Dubelaar et al. 2001) (Corsten & Gruen, 2003).

The parameter estimates in this study are also estimated with a three stage least square (3SLS) technique after dependent variable transformation. The dependent variables of SHO frequency and duration are normalized into independent units of measurement with mean 1 by dividing each observation by the mean store week value of that variable. For example, each SHO frequency count observation is divided by the mean SHO frequency per store week (all items within each store are divided by the same

mean value for each week) so that the distribution of SHO frequency changes from a count (min=0, max=9, mean=0.62, std. dev.=1.24) to a continuous distribution (min=0, max=23.68, mean=1, std. dev.= 2.10). The simultaneous equations still remain in the same format, with frequency being fitted into duration and both duration and frequency using fitted values in the main sales equation.

There are two main differences between the two simultaneous equation techniques. Since 3SLS uses least square estimators instead of maximum likelihood estimation, there is no longer a need to weight the dependent variable count values over open store time periods. The dimensionless normalized dependent variables are equally weighted across all observations. The second difference is with regard to the path diagram from variables to one another. Instead of using lagged predicted values from a sales equation in a previous step, the frequency equation simply has a lagged sales variable, and the sales equation has no lagged sales as a control for temporal effects. Similarly, there are no robust errors clustered by item in the 3SLS structure, as the least squares estimator is not consistent with heteroscedasticity. Even with the clustered standard errors in the GSEM structure, since it is not a linear panel model the standard errors may also be biased. A visual comparison of the predicted sales values against items suggests similar range of values (similar variance) across items (Appendix XIIV). An issue with the transformed 3SLS dependent variables would be interpreting the results beyond sign and significance.

4.5 Empirical results

The empirical results of both specifications are listed in Table 4-6 and include the parameter estimates, their standard errors, and the fit statistics for the entire model and product category subsets. Even-numbered columns list the results for 3SLS and the sign and significance of the coefficients are almost all the same as odd-numbered columns with GSEM estimates. There are two sets of differences. The first involves the yogurt subset where case pack size does not vary and is omitted from the model. The second involves the isotonics subset where capacity's lack of significant effect on frequency affects the duration and sales equations. The AIC and BIC values cannot be compared across GSEM and 3SLS models although they have exactly the same set of data and variables since 3SLS has a different set of residuals than the original model.

For the demand driver to SHO, the hypothesis that an item's prior week sales are linked to an increase in SHO frequency finds support ($p<0.001$) in all 6 columns. Stores focusing on decreasing SHO occurrence as part of their performance metrics would benefit from directing replenishment personnel one week to the most sold items the prior week. Even if new shipments arrive every week, it is likely that the person who places orders to the store is not the same one who receives orders at the backroom. Upon shipment receipt, personnel could be directed to unpack and replenish selling area shelves in descending order of item sales volume.

	Entire Sample		Yogurt		Isotonics	
	(1) GSEM	(2) 3SLS	(3) GSEM	(4) 3SLS	(5) GSEM	(6) 3SLS
Sales equation (H4.4)						
Duration	0.090***	0.149***	0.135***	0.137***	0.063***	0.029***
	(0.007)	(0.006)	(0.022)	(0.006)	(0.006)	(0.002)
$Duration^2$	0.001**	0.003***	0.003**	0.008***	0.0002	0.0008***
	(0.0004)	(0.0002)	(0.001)	(0.001)	(0.0001)	(0.0001)
Duration * Frequency	-0.024***	-0.054***	-0.041*	-0.038***	-0.012***	-0.007***
	(0.004)	(0.005)	(0.016)	(0.005)	(0.003)	(0.0008)
Duration equation (H42b, H4.3)						
Frequency	14.205***	1.846**	16.181***	1.019***	31.029***	1.271***
	(1.809)	(0.589)	(2.454)	(0.189)	(5.820)	(0.265)
$Frequency^2$	-1.644***	0.219	-2.434***	0.079	-7.113**	0.091*
	(0.304)	(0.315)	(0.486)	(0.045)	(2.216)	(0.045)
Case pack and the interaction between case pack size and capacity provided in Appendix IV						
Frequency equation (H4.1, H4.2a, H4.2c)						
Capacity	0.200***	0.097***	0.0004	-0.031***	0.134	0.103**
	(0.045)	(0.025)	(0.007)	(0.004)	(0.080)	(0.036)
Prior week sales	0.020***	0.013***	0.013***	0.013***	0.037***	0.024***
	(0.004)	(0.002)	(0.003)	(0.002)	(0.007)	(0.003)
Case pack Size: 4	-0.013	0.061**			-0.027	0.028
	(0.028)	(0.021)			(0.027)	(0.019)
8	-0.216***	-0.137***			-0.199***	-0.138***
	(0.041)	(0.031)			(0.044)	(0.029)
12	-0.202***	-0.127***			-0.175**	-0.135***
	(0.043)	(0.024)			(0.055)	(0.026)
15 *Capacity	-0.173***	-0.100v***			-0.147**	-0.112***
	(0.034)	(0.020)			(0.053)	(0.026)
Observations	4,144	4,144	2,408	2,408	1,736	1,736
AIC	50,619.96	48,991.36	30,555.06	28,400.84	19,334.62	19,017.96
BIC	51,024.04	49,390.11	30,815.46	28,649.67	19,634.88	19,312.77

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 4-6 Empirical results

For the supply driver to SHO, there are four different hypotheses throughout the different levels of the model. Shelf capacity's negative link to SHO frequency (H4.2a) is significant ($p<0.001$) but not supported. It is only significant and negative as hypothesized in column 4. The negative relationship between case pack size and SHO duration (H4.2b) was included in the models in terms of case pack dummies, with the base case as 3 units. The parameter coefficients are all not significant ($p>0.05$) except in 3 places. Column 6, the isotonics subset under 3SLS, is positive and significant ($p<0.05$) at case pack size 15 units. Column 2, the yogurt subset under GSEM is significant and negative effect ($p<0.05$) at case pack sizes 8 and 12. These conflicting results may be due to insufficient variation in case pack size, leaving estimates extremely sensitive to model specification. Finally, there are two interaction-related hypotheses (H4.2c and H4.2d) where the negative moderating effect on SHO frequency is supported ($p<0.01$) for all case pack sizes except for 4 units. Pack size 4 is significant ($p<0.01$) but positive in column 2 also conflicting with other case pack coefficients. The negative moderation of shelf capacity on the case pack size SHO duration relationship is supported at different rates of significance in different columns. Looking only at the product category subsets (as the overall models are weighted averages of the subsets) yogurt finds no significant moderating effect in the GSEM structure and a negative ($p<0.05$) at case pack size 12 units. Both of the 3SLS coefficients for the moderating effect are significant and negative (GSEM $p<0.001$, 3SLS $p<0.05$). Since the case pack size variable does not take on many values, it is included in the model as a factor variable.

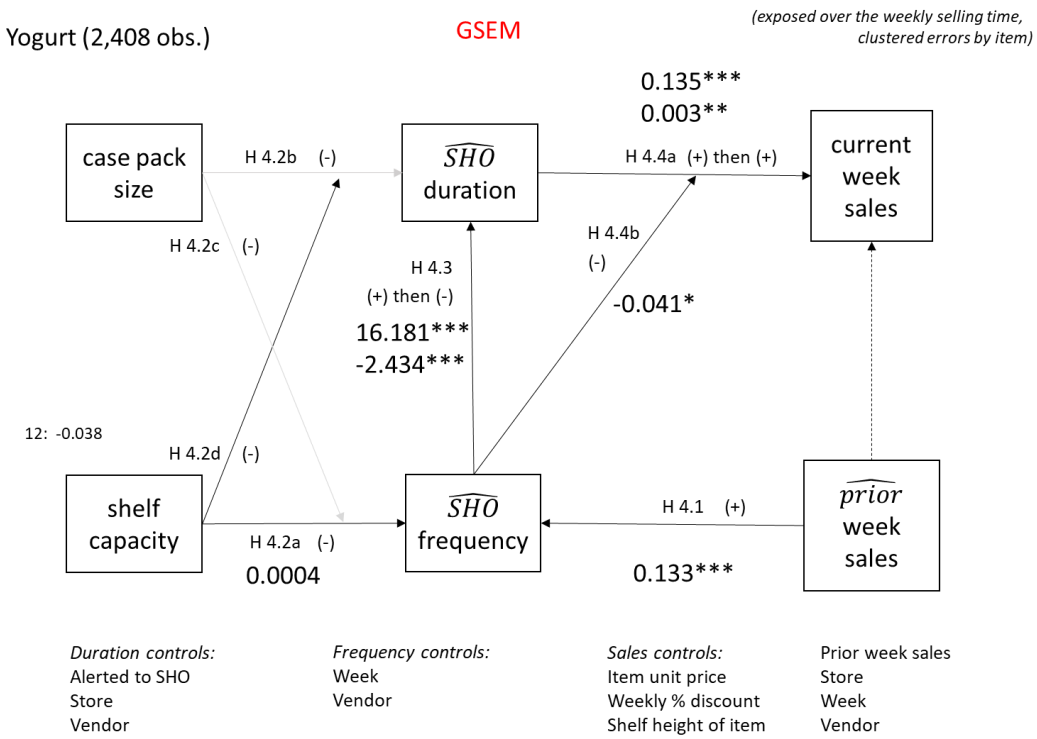
The different case pack sizes vary in coefficient estimates such that some support the study hypotheses while other sizes are not significant. This suggests that there may be a relevant range of values in which changes in case pack size can explain changes in SHO attributes. While the coefficients in the duration equation generally find support with negative and significant signs, the positive and significant sign in the isotonics product category for frequency are opposite of the hypothesis. Increasing case pack size from 3 to 15 units is linked to increased SHO frequency. This category has DSD and items in case packs of 15 are all 32-ounce single bottles with up to three times more sales per week than the other isotonics items. In contrast items with case packs of 3 units are 8-packs of 20-ounce bottles where 99% of weekly sales is less than the mean weekly sales of the 32-ounce items. This suggests the finding (Eroglu et al. 2011) of a 3-way interaction between case pack size, capacity and sales on SHO frequency may need to be incorporated into future functional forms.

The relationship between SHO attributes also has mixed findings. Looking at the duration equation, the direct linear effect of frequency in all 6 columns is positive and significant, supporting the first part of H4.3. For the quadratic term, H4.3 is supported only in GSEM ($p<0.001$), whereas it is insignificant in the overall and yogurt categories of 3SLS and significant with a positive sign ($p<0.05$) instead of negative for the isotonics category. Total weekly SHO duration increases on average with increased SHO frequency. However, the 3SLS model does not find support for a decreasing relationship past a certain amount of SHO frequency.

SHO attributes effects on sales have two different hypotheses on the sales level of equations. The curvilinear relationship between SHO duration and sales is supported (H4.4a, $p<0.001$) in all columns for the linear term and at least the 95% confidence level for the quadratic term. All coefficients are positive and significant except for the isotonics category in column 5. Here there is only a positive linear effect of duration on sales. This result leans towards the logic that sales are lower when there is little to no SHO duration because the items are on-hand at greater levels than needed for the inventory period. SHO duration is longer when sales are so high that even the safety stock is sold earlier than expected. Last of all, frequency does indeed moderate ($p<0.001$ for overall model) the SHO duration and sales relationship, but since that relationship is not negative as hypothesized but positive, the negative impact of frequency doesn't strengthen the relationship but instead weakens it.

In other words, sales are higher as SHO duration is longer, but as that SHO duration is split into a greater number of SHO events, sales suffer. The possible explanation for the positive link between SHO duration and sales is that the item has a STO with all safety stock sold much earlier than expected for the inventory period. If that same duration is split among multiple SHO events, then the source of the SHO is an IRO instead of an STO (since this is as periodic review system). Instead of all of the cycle and safety stock selling out and having a long STO duration (linked to higher sales), the cycle stock on the dedicated space has SHO with stock remaining somewhere in shared space (an IRO type of SHO). Future research could build on the implications of this study by identifying SHO events as either IRO or STO and testing their multiple attributes at once.

Finally, the results in Table 4-6 warrant discussion of model specification. Since there are squared frequency terms in the model, there may be a question as to whether the model needs to be respecified with centered variable values. Appendix XII shows that models with centered variables have larger AIC and BIC values than the same models without centering. That is why non-centered variables are presented in Table 4-6. All of the hypotheses and results of this study are illustrated by subset in Figure 4-5 by product category. Since the case pack size in the yogurt category does not vary, the coefficient estimates for H4.2b and H4.2c are omitted.



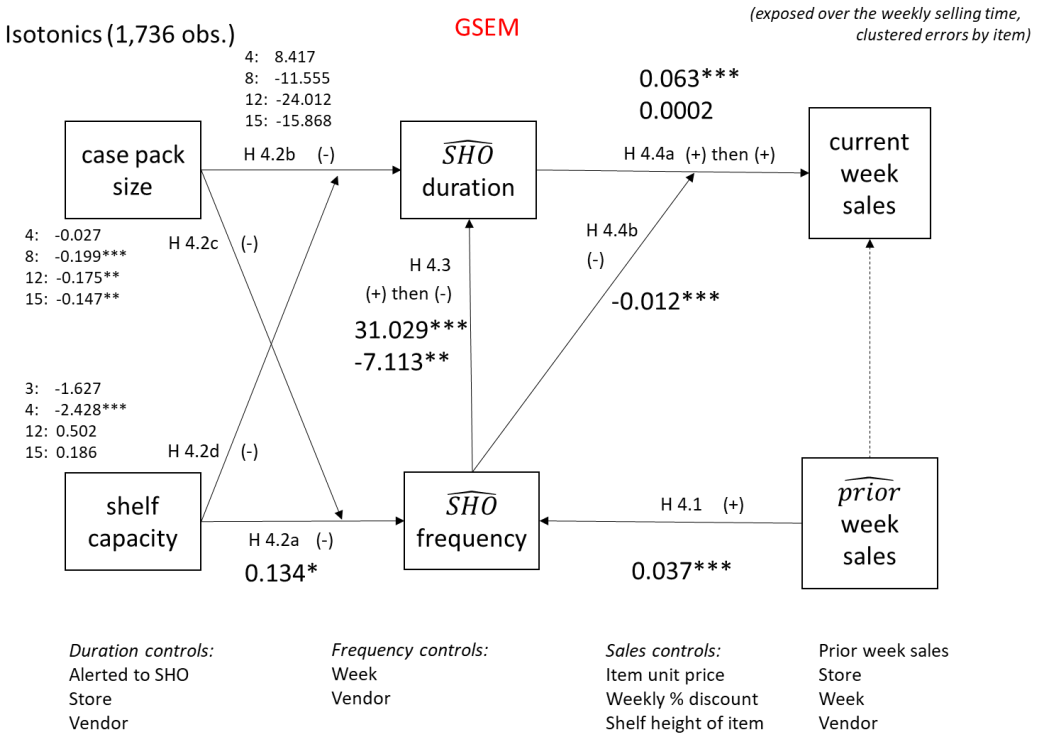


Figure 4-4 Hypotheses and results by product category

4.5.1 When are weekly item sales the highest?

When are item weekly sales at their highest? Intuitively sales are highest when an item is always available on dedicated shelf space because all demand can be met. However, if there are a million units of an item on store shelves and only a dozen people wanting to buy it, the fact that it is always available does not necessarily mean it has higher sales than another item with 12 units of availability and a dozen customers. Similarly, without knowing inventory review policy, replenishment strategy, resource emphasis, and whether there is instore logistics postponement, it is difficult to ascertain whether frequent short SHO events are linked to higher sales than infrequent long SHO events.

Comparing non-SHO weekly sales with 2x2 high-low groups of frequency and duration can answer the question of which combination of SHO attribute levels are linked to highest sales. First, high-low cut-off values for duration and frequency are generated by taking the mean overall SHO rate and mean overall SHO duration of non-zero SHO weeks. Cut-offs are also generated for each product category. Each observation is assigned an overall group and a product category grouping per week: no SHO events, low frequency short duration, low frequency long duration, high frequency short duration and high frequency long duration groups. The predicted sales (both overall and product category model-based predictions) observations are separated by group and are then compared by Bonferroni-corrected pairwise comparison of means. The results of the

comparisons are qualitatively presented in Figure 4-5 and quantitatively provided in Appendix XV.

High-Low frequency and duration's effects on sales?

All data	**Low frequency**	**High frequency**	**No stockout**
Short duration	middle	highest	(A) lowest
Long duration	(A) lowest	2nd highest	

Yogurt	**Low frequency**	**High frequency**	**No stockout**
Short duration	(B) 2nd lowest	(C) highest	(A) lowest
Long duration	(AB) lowest	(C) highest	

Isotonics	**Low frequency**	**High frequency**	**No stockout**
Short duration	n.s.	n.s.	n.s.
Long duration	n.s.	n.s.	

Bonferroni groups of average number of units sold GSEM

High-Low frequency and duration's effects on sales?

All data	**Low frequency**	**High frequency**	**No stockout**
Short duration	(B) middle	highest	(AB) lowest, middle
Long duration	(A) lowest	(B) middle	

Yogurt	**Low frequency**	**High frequency**	**No stockout**
Short duration	(A) lowest	highest	(A) lowest
Long duration	(A) lowest	middle	

Isotonics	**Low frequency**	**High frequency**	**No stockout**
Short duration	n.s.	n.s.	n.s.
Long duration	n.s.	n.s.	

Bonferroni groups of average number of units sold 3SLS

Figure 4-5 When weekly item sales are the highest

The postestimation testing yields a number of interesting findings about average weekly sales. First, there is no significant difference in sales of DSD-managed isotonics items in either the GSEM- or 3SLS-generated sales predictions. Secondly, the lowest sales group in yogurt is in the no SHO situation. Based on the GSEM model, high-frequency SHO weeks (more than 2 SHO events) are associated with the highest average weekly sales of yogurt, regardless of the length of SHO duration. The 3SLS results for yogurt show high-frequency short-duration SHO weeks have higher sales than high-frequency long-duration SHO weeks. In contrast, there is no significant difference among 3SLS low-frequency groups whereas in post-GSEM tests low-frequency long-duration weeks have lower sales than high-SHO long-duration weeks.

4.6 Discussion

This study provides item-level analysis of both the antecedents and effects of two different SHO attributes, frequency and duration. Duration and frequency attributes are obtained from records of actual SHO events instead of simulation, estimation from sales data, or as an experimental condition created by the researcher. Actual SHO event data enables studying the drivers and effects of SHO together, whereas they normally constitute two separate streams of research (Aastrup and Kotzab, 2010). It also allows for linking the antecedent-SHO relationship with sales outcomes to capture an overall picture of product availability.

Some of the findings of this study support or are in contrast to prior SHO research. The analysis of archival data in this study builds on the findings of the two main

papers (Wu et al., 2013; Eroglu et al., 2011) cited in the literature review and which use simulation to study item SHO. First, regarding the antecedents to SHO frequency, increasing the dedicated space for an item (its shelf capacity) is linked to decreasing SHO frequency (Eroglu et al., 2011). While Eroglu et al. (2011) find case pack size strengthens this negative relationship, this study shows a path from case pack size to SHO duration instead. Secondly, regarding the effect of SHO duration, unlike Wu et al. (2013) who find worsening store performance with increased SHO duration, this study finds improved performance linked to higher SHO duration. The different findings may be due to different measures of performance and of scope of research. Wu et al. (2013) focus on the phenomenon of stockout-based customer switching and measure performance in terms of product and store market shares. This study completely overlooks this substitution phenomenon and the measure of units sold is not capable of gauging performance relative to other items or products. It is possible that stocking substitutable items pools the risk of SHO so that a greater proportion of incoming customers buy one item or the other resulting in both items having greater sales as well as an earlier SHO duration (Chapter 2).

Other study findings contribute to product availability literature and have implications for future research. First of all, sales of an item in one period are directly and positively linked to the item's SHO frequency in the subsequent selling period. While scholars have looked at autoregressive models of STO's impact on sales (Calderon-Cortes & Morales-Arroyo, 2015) it has not been on the unit level of analysis or in more than one product category. Secondly, the SHO attributes of frequency and duration do not

relate to item sales in the same manner. While SHO duration is positively related to sales, frequency's relationship with sales is nonlinear and weakens the positive relationship of duration with sales. Future product availability research at the item level of analysis should include both measures of SHO in antecedent or effects studies. Since SHO breadth is the most-often used attribute (Chapter 1) of product availability literature, future studies at this higher level of analysis (product category or store) should consider manually auditing store shelves multiple times (like Taylor and Fawcett, 2001) to get a better gauge of the SHO duration of each item. If enough observations are obtained from store visits over time, it may be possible to see if the breadth and duration attributes also have a similar nonlinear relationship to sales. Being able to differentiate between IRO and STO would also help to confirm that items with high SHO frequency and high sales are being managed with instore logistics postponement.

The managerial implications of this research push the store to view SHO events differently. Just as prior sales help store managers place orders for an upcoming inventory period, they can also guide managers on how to prioritize store order receipt and shelf replenishment tasks for the upcoming store week. A set of postestimation tests show that the highest sales occur in high frequency and short duration situations, followed closely by a high frequency and long duration environment. The lowest sales are associated with zero SHO occurrence. Counterintuitively, in low frequency SHO situations, long SHO duration scenarios are not significantly different than low duration scenarios. Instead of focusing on eliminating SHO occurrence altogether, tracking SHO

in terms of both frequency and duration would be a useful tool in deciding whether to change order size or request shelf-planning personnel to adjust dedicated space for items.

Research limitations include aspects of the data used in this study and the method of analysis. With only two partial product categories in four stores, this study faces generalization issues. Not all items within a product category or a store are managed in an identical fashion; some items may be directly managed by vendors, other items may have manufacturer- or store-based promotions which shift focus away from regular items, and the availability of substitutable items may not carry as much importance to store managers wishing to hold as little inventory as possible. Differences in inventory management by item attributes is not considered in this study, nor are different types of inventory review practices. Additionally, there is no differentiation between SHO durations during high-traffic periods from those at low-traffic periods. Since a SHO is more likely to occur during high-traffic periods, the SHO duration during low-traffic periods carries more weight than it probably should, possibly resulting in an inflated coefficient for SHO duration parameter on item sales. The GSEM model overall benefits from being able to fit multiple nonlinear and discretely distributed endogenous variables at once but is difficult to interpret beyond parameter coefficients. On the other hand, the 3SLS model overall benefits from accounting for covariance of errors of dependent variables across equations but cannot easily account for heteroscedasticity or serial correlation.

Chapter 5 Conclusion: Take-aways from the retail shelf

This dissertation focuses on product availability which spans the research streams of retail category management, fast moving consumer goods, packaging logistics, shelf management and instore logistics. Chapter 1 provides a closer look at the shelf replenishment process and within retail-specific research. The chapter also looks at broader product availability research and risk pooling through substitution as well as through postponement. The research deficits in this area stem from the inability of fully capturing retail shelf stockouts (SHO) which results in use of various SHO attributes and separate streams of research into the causes and effects of SHO. Chapter 2 considers customers and products in studying inventory management decisions and its performance impact in terms of customer service. Chapter 3 considers product and employees in shelf display and in-store logistics decisions and its performance impact also in terms of customer service. Chapter 4 focuses on products and the financial (unit sales) impact of in-store logistics decisions. The sections below highlight dissertation context, contributions, and avenues for future research.

5.1 Dissertation context

The first chapter sets the research theme of the dissertation, of uncertainty and on-shelf availability with decisions on in-store logistics, inventory management, and product assortment and display. It applies the three opposing objectives of managing inventory according to Tabucanon and Farahani (1985) within the setting of instore replenishment, as illustrated in Figure 5-1. The illustration shows the number of units (y axis) of an item

at the store overall (dotted inventory investment curve) as well as the number of units on customer-accessible shelves (bold solid black customer service curve) over a time span of 8 days (x axis). Operational efficiency is illustrated by shaded regions in the graph which are the difference between what is in the store and what is on selling floor shelves. Shelves should have been replenished during these shaded regions which reflect poor instore replenishment. Replenishment is necessary whenever the customer service curve decreases, in other words whenever a customer purchases the item. One item is illustrated in this figure without considering parallel and sometimes competing demands and replenishment of other items. The 3 studies in this dissertation (chapters 2-4) reflect different aspects of Figure 5-1.

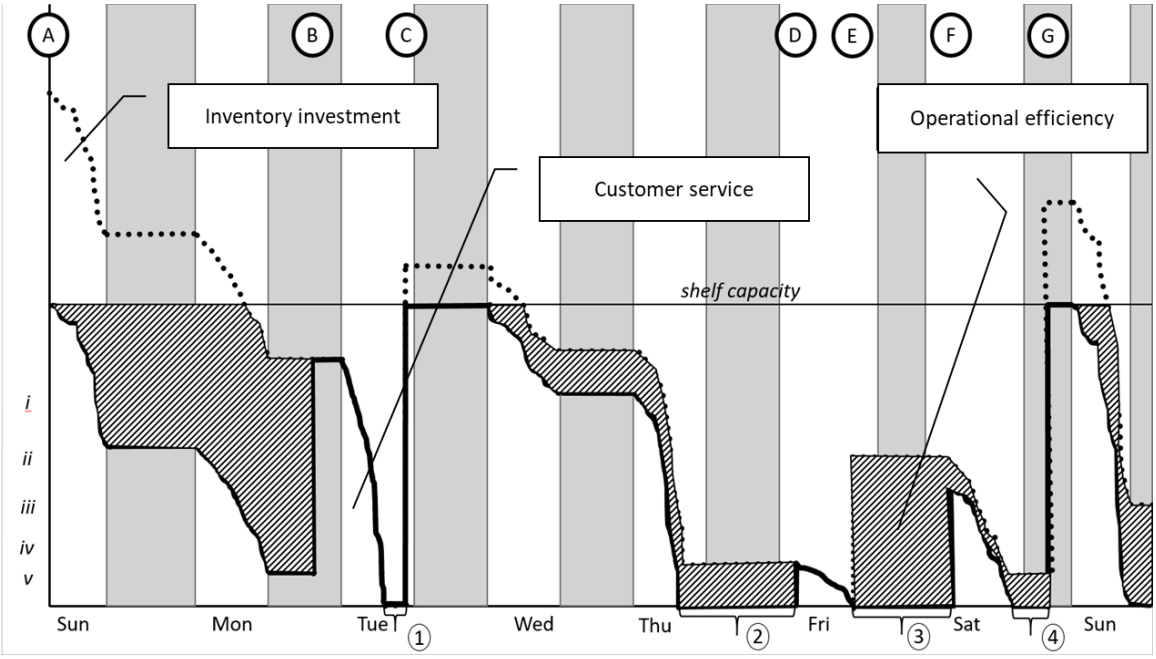


Figure 5-1 Shelf replenishment with two-tier inventory

5.2 Contributions

The three main areas of contribution to theory are in risk pooling through substitution, through postponement, and in product availability research. Each research question area is summarized below.

5.2.1 RQ1 What is the relationship between stockout-based customer switching and fill rates?

The illustration in Figure 5-1 is considered for multiple items in Chapter 2 where the inventory investment and customer service curves are identical since there is no secondary inventory space and no instore replenishment. With no instore replenishment, there is no backroom effect, or supply issues affecting the customer service curve over time. The customer service curve is thus only affected by demand. The demand of two separate items intersect in Chapter 2, when one item stocks out and customers switch to the alternate item as a substitute purchase. The alternate item faces an inventory effect (Borin, Farris, & Freeland, 1994) so its decreasing inventory investment curve now reflects an “effective demand” (Rajaram & Tang, 2001), or the sum of both initial and alternate customer demand. For example, during the stockout period 3 in Figure 5-1, customers who arrive to the store for this item and find it SHO may be willing to switch to another item. If they switch, although they are not accounted for at all in the graph of this item, they are now in another graph (of the alternate item) as part of that item’s decreasing inventory investment curve.

Substitute purchases drive inventory investment levels downward, and yet traditional substitution theory posits that customer service (and thus firm performance) increases with substitution since customers facing SHO are able to leave the store with a substitute item purchase. These substitute purchases effectively pool the inventory investment of two separate items into a larger supply source so the risk of customers leaving without a purchase is mitigated. With 100% two-way substitutability (every customer is willing to buy the alternate item when faced with SHO) the inventory investment curve (Figure 5-1) for an item switches to the inventory investment level of the other item whenever its own inventory level falls to zero. The alternate item's customer service curve captures these substitute purchases. When the item's own shelf inventory is replenished, the inventory investment curve returns to its own levels. The inventory investment curve fluctuates and is thus separate from the customer service curve even without instore replenishment. With instore replenishment, the difference between both curves holds a region of operational inefficiency. However, with customer switching the temporarily higher inventory investment level marks the period of time when product availability is higher without actually holding more stock. Not holding more stock but having greater product availability is at the core of risk pooling through substitution. This view is in contrast to more recent views on substitution which find: higher inventory levels (Yang & Schrage, 2009), decreased retailer market share (Wu et al., 2013), decreased profits (Hsieh, 2011) as well as greater difficulty in forecasting demand (Cooper et al., 2006; Anupindi et al, 1998) which drives SHO—all of which oppose risk pooling theory.

The contribution of Chapter 2 is reconciling opposing views on whether the customer switching phenomenon helps or hurts firm performance. Performance is studied by bringing together various fill rate measures into one study and by relaxing assumptions of prior substitution studies. The first of two assumptions that are relaxed is symmetric substitutability; in Chapter 2 the proportion of focal customers willing to buy the alternate item may be greater (or less) than the proportion of customers willing to buy the focal item as a substitute. The second is the introduction of random customer arrival sequence (Gilland and Heese, 2013) in an agent-based model perspective which is compared to fill-rate heuristic (FRH) outcomes. The FRH essentially represents outcomes from a customer service curve in Figure 5-1 where customers are assumed to buy an alternate item only after *all expected initial* customers have bought it. This FRH assumption represents the traditional perspective on the substitution phenomenon in terms of risk pooling in empirical and analytical papers. In contrast, the agent-based model does not assume how the demand or inventory investment curve in Figure 5-1 will decrease over time. Over time the agent-based simulation program tallies which simulated customer buys which item and the emerging outcomes are compared to FRH outcomes using ANOVA followed by Tukey's Studentized Range test to find Tukey groupings.

Comparison of outcomes from the two different perspectives on customer switching helps to find a middle ground in the dispute of whether this type of substitution increases or decreases sales and SHOs. Both sides of the debate find support in the findings. Customer service (as measured by various unit fill rate outcomes) has a

nonlinear relationship with the customer switching phenomenon. At very low inventory investment (as measured by target service levels, TSL) there is not enough in stock for customers to switch to so that increasing substitutability is linked to worse sales outcomes contrary to risk pooling theory. The main reason for this may be because replenishment cycles which originally start out in sync are pushed out of sync with substitution. At very high inventory investment, customer service does not improve through substitution. Substitutability makes little to no difference in terms of increasing either initial or alternate sales, perhaps because SHO does not occur often enough to require customer switching.

At the medium level of inventory investment not only the degree of substitutability but the direction of substitutability matters in terms of what proportion of focal customers get which item and which walk away with no purchase. With symmetric two-way substitutability, overall the number of focal customers who face SHO are as expected by the FRH. Fewer customers walk away without a purchase as symmetric substitutability rates increase (supporting risk pooling theory). It is interesting to note that fewer *focal* customers are able to buy the initial item with increasing substitutability, because alternate customers, whose item is now being purchased by focal customers, buy the focal item as a substitute in greater numbers. For one-way substitutability, risk pooling theory does not find support. If one-way substitutability towards the alternate item goes up (increasing focal customer willingness to switch), then focal customers willing to substitute face SHO more, contrary to risk pooling. If one-way substitution

towards the focal item increases, then focal customers not willing to switch face SHO more, also contrary to risk pooling.

Another contribution to the theory of risk pooling through substitution is that its impact on customer service can be masked or accentuated depending on which type of fill rate is chosen as a performance measure. Performance measures are compared with MANOVA tests followed by ANOVA to obtain partial eta square effect size values of substitutability and inventory investment's impact on fill rates. The most aggregated fill rate measure, category fill rate, is significantly linked to the level of inventory investment of each item as well as the proportion of customers willing to substitute. While both inventory investment levels have equal impact size ($\eta_P^2 = 0.30$) on category fill rate, the other three fill rates (item, customer, initial) accentuate focal item inventory investment ($\eta_P^2 = 0.74$-0.76) while masking the effects of alternate item investment level changes (no significant effect).

This means that holding more inventory of an alternate item appears to have no effect on focal item fill rate or focal customer fill rate, contrary to risk pooling theory. Even though there is an impact, it is only captured by the category fill rate or the alternate item or alternate customer fill rate measures. This parallels the discussion of Figure 5-1, where customers facing SHO and switching to the alternate item do not appear anywhere in a customer service and inventory investment graph of the focal item. Similarly, risk-pooling effects of increasing the proportion of focal customers willing to switch is only captured through category fill rate measures or fill rates of the alternate item or alternate

customer. This suggests that perhaps reaching a consensus in the literature on the effectiveness of risk pooling through substitution would be assisted by analysis with multiple performance measures within a single study and by comparing across studies with the same fill rate types.

Finally, comparing the two views on substitution leads to practical contributions for retailers. For a retailer managing a group of items that are a mixture of non-substitutable, fully-substitutable, and two-way substitutability (symmetric and asymmetric), the FRH serves as a reasonable rule of thumb for inventory investment (TSL) decisions. TSL decisions should consider what proportion of customers are willing to switch between items. Items where less than half of customers would switch are good candidates for risk pooling through substitution. Substitution however, will hurt sales when half or more of customers would switch between items. Such items should have higher inventory investment than store software recommends or store managers estimate. If the store would like to focus on a specific group of customers (those who buy a certain brand or set of items as initial or alternate purchaes) then they should consider which other customers would switch to the focal customer item(s) and have higher target service levels for those items. It may seem counterintuitive to have higher service goals for items that the focal group may not be interested in. The reasoning behind this is trying to avert other customers from switching to the focal customer's items. These items may be the focal customer's initial or alternate choice.

5.1.2 RQ2 How does instore logistics postponement affect product availability?

This chapter contributes to the theory of postponement by introducing the term "instore logistics postponement" (ILP) which reflects the two-tier inventory already existing in most brick-and-mortar stores (Tiwari, Jaggi, Gupta, & Cardenas-Barron, 2018) and is illustrated by Figure 5-1. Figure 5-1's inventory investment is when the store implements ILP and the inventory investment curve disappears (merges with customer service) when there is no item ILP. When there is ILP, the operational efficiency areas are the shaded regions of Figure 5-1 and are measured with six different SHO attributes. This chapter's methodology is two-fold, resting largely on how to measure ILP. First, a factor analysis attempts to capture when a store takes part in ILP with four different indicators. Then ILP is measured with three indicators including a binary dummy for taking part in ILP. This latter measure of the ILP construct has greater explanatory power. Secondly, a GSEM structure includes ILP with only the binary definition instead of as a construct. Both approaches lead to similar results. Since the GSEM structure allows for each observed variable to be included directly into the model instead of as construct indicators, a more nuanced view of the relationship between an item's demand characteristics and product availability (SHO attributes) is reached. Overall, both the CFA and GSEM show decreasing product category SHO occurrence with ILP.

Since items within a product category may have different characteristics (demand unpredictability, velocity, and other systemic drivers) another study contribution is

adapting Fisher's (1997) classification of items as either functional (low systemic drivers) or innovative (high systemic drivers) to suggest a framework to decide when to use ILP. ILP is appropriate whenever the store's resource focus is efficient use of shelf space instead of labor. If the item's systemic drivers are high, then ILP allows the backroom to be an area for item safety stock. This fits in with a store strategy that of increasing availability, as the safety stock can be used to replenish the store shelf whenever an item's next shipment to the store is on its way and shelf stock levels are SHO or nearly SHO. This is possible whenever high-systemic driver items have had an unusually high inventory period sales. On the other hand for items with low systemic drivers, ILP is still a suitable inventory investment approach if the store has limited shelf space and high product variety. Both the safety stock and a portion of the cycle stock can be stored in backroom space so that the store can offer a wider assortment of items than possible if all or most of the item inventory were kept on customer-accessible shelves. The two main reasons when ILP would not be preferred is when the store emphasizes labor over shelf space. Then an item strategy towards efficiency would push the store towards avoiding using labor resources for instore replenishment, having personnel put these low-systemic driver items entirely onto the dedicated shelf space upon store delivery. Similarly, an item strategy towards availability would push the store toward avoiding using labor resources and instead increasing item shelf capacity so a store delivery could fit dedicated shelf space entirely.

Empirical results show that the components of systemic drivers are linked to ILP in different ways. More unpredictable items (an increased systemic driver) are more

likely to take part in ILP where safety stock is kept in the backroom to use during inventory periods where the shelf stock may run out before the next store delivery. However, higher velocity items (an increased systemic driver) are less likely to take part in ILP so that the items are more likely stocked only on the dedicated shelf for maximum availability. Furthermore, average item unpredictability has no significant link on either SHO duration for the overall category or SHO frequency for the store. On the other hand increased item velocity is linked to increased SHO duration for the category but not to SHO frequency for the store. Larger category sales (an increased systemic driver) are more likely to take part in ILP as evidenced by empirical results showing increased store and category SHO frequency but decreased SHO duration.

5.1.3 RQ3 How are SHO measures related to one another and to item sales?

Last of all, Chapter 4 contributes to product availability research at the item level of analysis by combining two different SHO attributes, SHO antecedents and their effects on sales. Sales are measured in units and reflect the customer service curve of the opposing objectives in Figure 5-1, with a steeper slope as a measure of higher sales velocity. The methodology in this chapter is GSEM and 3SLS as the two measures of SHO, frequency and duration, are endogenous regressors of a main unit sales equation. With a control for whether store personnel are alerted to an item's SHO (alerts should decrease the region of operational efficiency in Figure 5-1) this study finds higher sales linked to longer SHO durations, with the lowest sales for items which never SHO. This result initially sounds counterintuitive, since SHOs are usually linked to lower sales.

However, a look at the sawtooth model (Figure 5-2) of a periodic review system over time shows that when demand is higher during an inventory period (such as Q_2 sales in period 1) the average demand slope is steeper than period 2 sales which have not stocked out during inventory period 2. Since the stores used in this sample have a periodic inventory review period of a week, it makes sense that items with the lowest of weekly sales (mean 15.19 units for isotonics, 32.53 units for yogurt) will have lower SHO duration, on average, than items with the highest weekly sales (isotonics 20.78, yogurt 61.46 units of mean weekly sales).

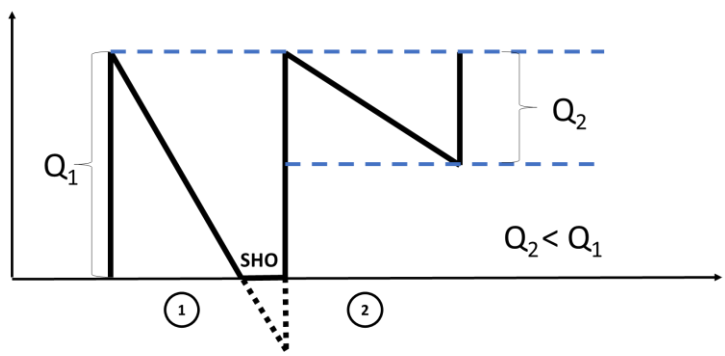


Figure 5-2 How zero SHOs can be associated with the lowest sales

The finding that higher SHO frequency is linked to higher unit sales also sounds counterintuitive at first glance. However, it supports the concept of instore logistics postponement in Chapter 3. Stores are quickly and perhaps partially restocking items from backroom space during the inventory period and between store deliveries.

Since this data is from a pilot study of a new technology and SHO frequency in the data can be as high as 9 SHO per week it may be necessary to consider the accuracy of data. Although assured by the pilot study company that this is not the case, a less exciting and still possible explanation may be that the shelf liner technology, of a standard surface area, has multiple sheets of sensors for those items with large shelf capacity. For example, yogurt with 100 units on the shelf may use 4 shelf liner sheets, while another yogurt item with shelf capacity 25 may use only one sheet. It is possible that the SHO event records for the 100-unit item captures when the stock on the shelf diminishes by one quarter multiples, so that only one out of 4 SHO events recorded is truly when there are 0 units of the item left on the shelf. So what if this is the case despite contact assurances that it is not so? Then both the SHO duration and the frequency will be higher, on average, for items whose number of facings exceed the width of the shelf liner. That will affect outcomes in exactly the way discussed here, with higher SHO frequency and longer SHO duration linked to greater sales than compared to lower frequency and duration conditions. Looking at the data by category, there are potentially 6 items (out of 86) where such a problem could occur for yogurt, and 24 (out of 62) for isotonics. The isotonics proportion may affect overall outcomes more so than yogurt, which may explain why the Bonferroni-corrected pairwise comparison of means found no statistically significant difference between any of the high-low frequency and duration groupings, including the no SHO case. One of the problems with historical data is having to assume the data is valid or accurate without being able to directly confirm that as in observational (researcher-collected, not archival) data. Though time limitations prohibit checking the isotonics items, a check was made for the 6 large-capacity (by looking at

number of facings) yogurt items which were all in only one of the 4 stores. Of the 48 possible observations in the data for these items, 44 observations had zero SHO. Only one observation for one item fell into the high-frequency group. With 2,752 observations total, the single yogurt observation could not affect the data enough to produce the counter-intuitive results of higher SHO frequency being linked to higher sales.

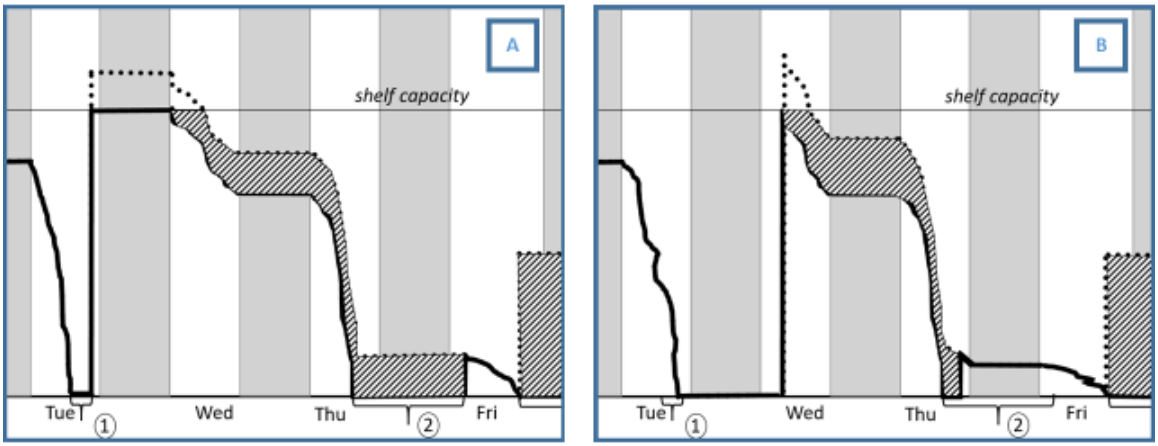
SHO frequency's relationship to SHO duration has mixed findings, only partially supporting the hypothesis of an inverse-U shaped relationship between the two. In the GSEM specification, the inverse-U relationship is significant for the overall model and product category subsets. However, for the 3SLS model, an inverse-U shaped relationship is not supported at all. Indeed, for isotonics, there is a curvilinear relationship where both the linear and quadratic terms are significant and positive. Both models have the same right-hand side variables (identical paths) but the GSEM structure controls for heteroscedasticity and autocorrelation whereas the 3SLS model automatically controls for covariance between the errors of dependent variables at the different levels but there is no user-written panel data correction. These results warrant a sensitivity analysis that is currently time-prohibited.

5.3 Future research

This section lists future research questions that have not already been discussed in the future research sections of prior chapters.

Numerous extensions can be made to other research areas. An immediate extension is comparison of the dissertation findings with online and omnichannel

retailers. Given their technology they are more likely to have full records of when (and for how long) items are SHO, and even how many customers "clicked" for that item, and what they chose next when being faced with SHO. The substitution chapter does not include any backroom inventory; considering instore replenishment (ILP) with substitutable items could look into if ILP allows for risk pooling through substitution to improve performance at every level of inventory investment. This dissertation did not consider retail distribution centers (RDC) specifically. Is ILP better than RDCs or even direct store delivery? A comparison of these 3 types of shelf replenishment approaches would be interesting both in terms of substitution and product availability. In direct store delivery the vendor incurs the time and money costs of replenishing shelves, whereas ILP is completely on a store location and RDCs move costs up to the regional or corporate retailer. There is also the concept of shelf replenishment when a SHO has occurred (reactively) and proactively, before the item has SHO. Reactive replenishment may be similar to emergency transshipments and a comparison of the two types of instore replenishment could parallel studies on transshipments and what types of risk (demand or lead time, or both) they pool against. Finally, companies like Amazon have been looking into different packaging options for case pack size to stores as well as shipping boxes to online customers (Stevens & Phillips, 2017). Vendors may wish to alter their case pack size to retailers—either retailer or store-specific packaging—because even though there is increased cost with customization, there may be fewer lost sales from the backroom effect or the costs could be offset by increased sales through ILP.



Scenario A replenishes the shelf at an earlier time than scenario B, because the delivery arrives earlier than at B.
The next delivery is still at Friday night, so that doesn't change between the two scenarios.
The customer purchases start happening earlier in A than in B, but the number of units sold in total is the same.
Scenario B replenishes from the backroom earlier than in A, so SHO 2 is shortened.

What happens to the operational efficiency? Which one has a greater inefficient area from Wednesday to Friday??? Or over the week overall? Don't sales also depend on whether the different SHO periods have different demand rates?

Figure 5-3 Example for future research

The opposing objectives of Tabucanon and Farahani (1985) were only referred to in this chapter and Chapter 1 as a guide for how on-shelf availability has supply (inventory investment) and demand (customer purchases) drivers that interact with one another, and are not directly tested anywhere in the hypotheses. Research in the areas of operational efficiency, of inventory investment decisions and customer service, can test for these concepts as constructs with some of the indicators being the variables used in this dissertation. While analytical work can try to model the 2 curves and the regions of operational efficiency between them, a combination of modeling and empirical work (perhaps through simulation) could study what it means to have certain amounts of operational inefficiency and whether efforts toward greater efficiency are warranted given changes in performance. An example set of questions is presented in Figure 5-2.

This example brings into question the concept of efficiency in retail. If the upper limit for sales is on-hand inventory levels, does it matter how long or how often SHO occurs? If the same number of units are sold in Scenario A as in Scenario B (Figure 5-2) does the operational efficiency differ? Do the average SHO duration and frequency attributes differ? Is operational inefficiency (poor instore replenishment) a greater issue when stores are open or closed? Since store personnel are no longer busy with customers during closed hours, all regions of operational inefficiency should disappear. Do they?

Future research opportunities also exist by combining the streams of literature in Chapters 2-4. For the empirical studies (Chapters 3 and 4) the demand driver is actually simply realized sales, which overlooks how many customers faced SHO and either switched to another product or left without purchase. There is also no order size for these two chapters as there is for the simulation in Chapter 2. Order size information would make it easier to pinpoint which items are being managed through ILP. Another major difference is that Chapter 2 has a reorder point model of inventory review while the empirical data for the other two chapters were obtained from stores with periodic review. Periodic review is purposefully not used in the simulation study (Chapter 2) as it would have inflated SHO duration by forcing the store to wait until the order point of the periodic review system instead of placing a store order at a reorder point at an earlier period of time. Thus, periodic review would be a confounding effect upon the substitution effects studied. However, if the review type is switched to periodic instead of continuous review, and the simulation model validated by the empirical data when there is no substitution (for example, by taking one item from one product category and one

item from another category), then the combination of ILP and substitution effects could be studied for substitutable items within both product categories. Substitution, just like ILP, causes a distinction between the inventory investment and customer service curves. Substitution may decrease areas of operational inefficiency for an item but also magnify the distance between the inventory investment of an item and its customer service curve of the alternate item. The research area of inventory review policies could study what types of review policies are most beneficial for risk pooling through substitution, through ILP, and when both phenomena exist.

Appendix

I. How packaging and logistics interact in the retail supply chain (Hellstrom & Saghir, 2006, p. 212)

Table 4. The interacting packaging aspects in the retail supply chain processes

Supply chain members	Manufacturer		Distribution centre						Retail outlet		
Logistics Processes	Filling process	Warehousing process	Transport	Receiving process	Storing process	Picking process	Shipping process	Transport	Receiving and shipping	Replenishing process	Re-use and recycle
Packaging system											
Primary	Packing line efficiency Filling speed Label application Closing and scaling technology Flexibility									Mandling efficiency Promoting sale Shelf adaptation Product identification	Handling efficiency Material
Secondary	Handling efficiency Packing line efficiency					Handling efficiency Identification Ergonomics Protection Stability			Handling efficiency Identification	Handling efficiency Identification Protection shelf adaptation Ergonomics	Handling efficiency Material
Tertiary	Handling efficiency Stackability	Handling efficiency Btackability Protection Stability	Cube utilization Stackability Weight and height Stability	Stability Identification	Cube utilization Weight and height	Handling efficiency Material	Handling efficiency Weight and height Stability	Cube utilization Weight and height Stability Stackability	Handling efficiency material Stability	Handling efficiency Store concept adaptation Product identification Promoting sale	Handling efficiency Material

II. Findings of systemic review on OSA drivers (Moussaoui, Williams, Hofer, Aloysius, & Waller, 2016)

Table 2 – List of OSA drivers by category and associated key references from the literature

Category of drivers	Drivers	Key references
Operational drivers	Poor forecasting	Corsten and Gruen (2003); Nachtmann et al. (2010); Taylor and Fawcett (2001)
	Poor ordering	Corsten and Gruen (2003); Ehrenthal and Stölzle (2013); Grant and Fernie (2008)
	Backroom operations	Nachtmann et al. (2010); Raman et al. (2001); Ton et al. (2010) ; Waller et al. (2008)
	Inadequate shelf space allocation	Eroglu et al. (2013); Nachtmann et al. (2010)
	Inadequate case pack size	Eroglu et al. (2013); Waller et al. (2008)
	Low replenishment frequency	Eroglu et al. (2013)
	Poor shelf maintenance	Gruen et al. (2002); Raman et al. (2001)
	Inventory record inaccuracy	Fernie and Grant (2008); Raman et al. (2001)
	Upstream execution failures	Corsten and Gruen (2003); Ettouzani et al. (2012); Raman et al. (2001); Willams and Waller (2012);
	Incorrect master data	Corsten and Gruen (2003); Nachtmann et al. (2010)
Behavioral drivers	System override	Corsten and Gruen (2003); Kremer et al. (2011); Ren and Crosson (2013); Van Donselaar et al., (2010)
	Poor shelf maintenance	Aastrup and Kotzab (2009); Corsten and Gruen (2003)
	Checkout errors	DeHoratius and Raman (2008); Nachtmann et al. (2010); Raman et al. (2001)
Managerial drivers	Lack of managerial emphasis on the importance of OSA	Raman et al. (2001)
	Management turnover	Raman et al. (2001)
Coordination drivers	Poor coordination and lack of synchrony	Corsten and Gruen (2003); Ettouzani et al. (2012)
	Poor communication and lack of visibility	Corsten and Gruen (2003); Lee et al. (2000)
Systemic drivers (contingency factors)	Demand autocorrelation	Zinn et al. (1994)
	Demand unpredictability	Taylor and Fawcett (2001)
	Demand velocity	Corsten and Gruen (2003); DeHoratius and Raman (2008); Fernie and Grant (2008); Taylor and Fawcett (2001);
	Product variety	Raman et al. (2001); Stassen and Waller (2002); Wan et al. (2014)
	Network design	Fisher et al. (1997); Steckel et al. (2004); Wan and Evers (2011)

The proposed framework for studying retail shelf availability drivers (antecedents to SHO). Included with permission from authors.

III. Simulation outcomes for item Y

The next set of tables for item Y are included in the Appendix and mirror the tables in Figures 2-8 thru 2-11 for focal item X.

$P_Y^y + P_Y^{yx}$	Target Service Level								
	50%			75%			95%		
Buy Y if X is out	N	N	N	D	D	D	A	A	A
None	50 51.35	50 51.34	50 51.19	75 77.85	75 77.85	75 77.85	95 97.77	95 97.77	95 97.77
	P	K	L	F	F	E	AB	B	AB
Half	50 49.24	50 54.25	50 52.07	75 74.59	75 74.61	75 75.05	95 97.47	95 97.46	95 97.50
	O	M	J	G	I	H	C	C	C
All	50 49.63	50 51.75	50 57.62	75 72.52	75 69.07	75 70.23	95 96.98	95 97.05	95 97.16
	None	Half	All	None	Half	All	None	Half	All

Buy X if Y is out

$P_Y^y + P_Y^{yx}$ **Difference (t stats)**	Target Service Level								
	50%			75%			95%		
Buy Y if X is out									
None	1.34 (53.46)	1.34 (49.85)	1.19 (50.47)	2.85 (94.40)	2.85 (94.40)	2.85 (94.40)	2.77 (88.62)	2.77 (88.62)	2.77 (88.62)
Half	-0.76 (-19.49)	4.25 (123.50)	2.07 (36.97)	-0.41 (-10.20)	-0.39 (-11.07)	0.05 (1.44)	2.47 (87.66)	2.46 (75.63)	2.50 (102.6)
All	-0.37 (-4.03)	1.75 (10.29)	7.62 (75.16)	-2.48 (-44.08)	-5.93 (-100.0)	-4.77 (-110.0)	1.98 (64.55)	2.05 (102.6)	2.16 (50.19)
	None	Half	All	None	Half	All	None	Half	All

Buy X if Y is out

P_0^y	Target Service Level								
	50%			75%			95%		
Buy Y if X is out: None	C	F	R	I	L	R	O	Q	R
	50	25	0	25	12.5	0	5	2.5	0
	48.66	24.35	0	22.15	11.08	0	2.23	1.11	0
Buy Y if X is out: Half	A	H	R	E	K	R	N	Q	R
	50	25	0	25	12.5	0	5	2.5	0
	50.76	22.87	0	25.41	12.74	0	2.53	1.26	0
Buy Y if X is out: All	B	G	R	D	J	R	M	P	R
	50	25	0	25	12.5	0	5	2.5	0
	50.37	24.12	0	27.48	15.46	0	3.02	1.47	0
	None	Half	All	None	Half	All	None	Half	All

Buy X if Y is out

P_0^y **Difference (t stats)**	Target Service Level								
	50%			75%			95%		
Buy Y if X is out: None	-1.35	-0.65	-	-2.85	-1.42	-	-2.77	-1.39	-
	(-53.46)	(-16.41)	-	(-94.40)	(-40.05)	-	(-88.62)	(-82.18)	-
Buy Y if X is out: Half	0.76	-2.13	-	0.41	0.24	-	-2.47	-1.24	-
	(19.49)	(-47.30)	-	(10.20)	(6.36)	-	(-87.66)	(-63.71)	-
Buy Y if X is out: All	0.37	-0.88	-	2.48	2.96	-	-1.98	-1.03	-
	(4.03)	(-8.51)	-	(44.08)	(103.80)	-	(-64.55)	(-47.62)	-
	None	Half	All	None	Half	All	None	Half	All

Buy X if Y is out

P_X^{yx}		Target Service Level								
		50%			75%			95%		
Buy Y if X is out	Buy X if Y is out	None	Half	All	None	Half	All	None	Half	All
None		M	E	B	M	H	F	M	L	J
		0	12.5	25	0	9.38	18.75	0	2.38	4.75
		0	10.16	19.05	0	6.27	9.41	0	1.08	2.12
Half		M	I	C	M	D	B	M	KL	J
		0	12.5	25	0	9.38	18.75	0	2.38	4.75
		0	2.78	14.46	0	12.11	18.75	0	1.23	2.42
All		M	G	H	M	D	A	M	K	I
		0	12.5	25	0	9.38	18.75	0	2.38	4.75
		0	7.26	6.16	0	12.39	29.04	0	1.44	2.76

P_X^{yx} Difference (t stats)		Target Service Level								
		50%			75%			95%		
Buy Y if X is out	Buy X if Y is out	None	Half	All	None	Half	All	None	Half	All
None		-	-2.34	-5.95	-	-3.11	-9.34	-	-1.30	-2.62
		-	(-46.78)	(-120.0)	-	(-54.55)	(-110.0)	-	(-73.61)	(-84.53)
Half		-	-9.72	-10.54	-	2.73	-0.00	-	-1.14	-2.33
		-	(-200.0)	(-74.92)	-	(51.82)	(-0.024)	-	(-74.47)	(-100.0)
All		-	-5.24	-18.84	-	3.01	10.29	-	-0.94	-1.99
		-	(-44.73)	(-170.0)	-	(29.67)	(288.1)	-	(-42.84)	(-43.45)

P_0^{yx}		Target Service Level								
		50%			75%			95%		
Buy Y if X is out		L	F	C	L	I	G	L	L	L
	None	0	12.5	25	0	3.13	6.25	0	0.13	0.25
		0	14.16	29.76	0	4.80	12.74	0	0.05	0.11
		L	D	B	L	K	H	L	L	L
	Half	0	12.5	25	0	3.13	6.25	0	0.13	0.25
		0	20.10	33.46	0	0.55	6.20	0	0.04	0.08
		L	E	A	L	J	K	L	L	L
	All	0	12.5	25	0	3.13	6.25	0	0.13	0.25
		0	16.87	36.22	0	3.08	0.73	0	0.04	0.08
		None	Half	All	None	Half	All	None	Half	All

Buy X if Y is out

P_0^{yx} Difference (t stats)		Target Service Level								
		50%			75%			95%		
Buy Y if X is out	None	-	1.66	4.76	-	1.68	6.49	-	-0.08	-0.14
		-	(33.83)	(92.68)	-	(25.95)	(74.32)	-	(-16.71)	(-17.69)
	Half	-	7.60	8.46	-	-2.57	-0.05	-	-0.08	-0.17
		-	(165.9)	(91.87)	-	(-73.78)	(-0.37)	-	(-25.29)	(-25.17)
	All	-	4.37	11.22	-	-0.04	-5.51	-	-0.08	-0.17
		-	(71.84)	(162.9)	-	(-0.64)	(190.0)	-	(-24.86)	(-23.23)
		None	Half	All	None	Half	All	None	Half	All

Buy X if Y is out

IV. Full GSEM and 3SLS output (Ch. 4)

GSEM of overall data below.

```
Log pseudolikelihood = -25245.771

                                                (Std. Err. adjusted for 592 clusters in storeupc)
------------------------------------------------------------------------------------------------
                             |               Robust
                             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------------------+------------------------------------------------------------------
soldwk                       |
c.durratehat#c.eventratehat  |  -.0236813   .0044214    -5.36   0.000    -.0323472   -.0150155
                             |
                durratehat   |   .0898371   .0074626    12.04   0.000     .0752107    .1044636
               durratehat2   |   .0010671   .0003935     2.71   0.007     .0002959    .0018383
                  regprice   |  -.3430823   .0205297   -16.71   0.000    -.3833197   -.3028448
                  percdisc   |  -.0204105   .0060338    -3.38   0.001    -.0322365   -.0085844
                             |
                   shelfno   |
                         2   |  -.0031359   .1546408    -0.02   0.984    -.3062262    .2999545
                         3   |  -.0710701   .1292707    -0.55   0.582    -.3244359    .1822958
                         4   |  -.0146525   .1351972    -0.11   0.914    -.2796342    .2503291
                         5   |  -.1576235    .131418    -1.20   0.230    -.4151981    .0999511
                         6   |  -.1788821   .1373337    -1.30   0.193    -.4480512    .0902869
                         7   |  -.1266018   .1475192    -0.86   0.391     -.415734    .1625304
                             |
                     store   |
                      6515   |  -.0532447   .0518624    -1.03   0.305    -.1548931    .0484037
                      6520   |  -.1540976   .0474586    -3.25   0.001    -.2471148   -.0610804
                      6539   |  -.5753876   .0621234    -9.26   0.000    -.6971471    -.453628
                             |
                      week   |
                         1   |          0  (empty)
                         2   |  -.3511602   .0275326   -12.75   0.000     -.405123   -.2971974
                         3   |  -.2513359   .0219528   -11.45   0.000    -.2943625   -.2083092
                         4   |  -.0061798    .023534    -0.26   0.793    -.0523057     .039946
                         5   |  -.0690468   .0243895    -2.83   0.005    -.1168493   -.0212444
                         6   |  -.1472804   .0221271    -6.66   0.000    -.1906487    -.103912
                         7   |  -.0642516   .0313051    -2.05   0.040    -.1256084   -.0028948
                         8   |          0  (omitted)
                             |
                    vendor   |
                         2   |  -.2413811   .1112931    -2.17   0.030    -.4595115   -.0232507
                         3   |  -.8350309    .053822   -15.51   0.000    -.9405201   -.7295416
                         4   |  -.6295834   .0418883   -15.03   0.000     -.711683   -.5474839
                         5   |  -.8186177    .059968   -13.65   0.000    -.9361527   -.7010826
                             |
                       _cons |   -3.46142   .1868212   -18.53   0.000    -3.827583   -3.095257
                ln(openwk)   |          1  (exposure)
-----------------------------+------------------------------------------------------------------
durrate                      |
              eventratehat   |   14.20502    1.80937     7.85   0.000     10.65872    17.75132
             eventratehat2   |  -1.644395   .3044691    -5.40   0.000    -2.241143   -1.047646
                             |
                   casepack  |
                         4   |    20.1577   8.509766     2.37   0.018     3.478868    36.83654
                         8   |  -17.24898   8.639632    -2.00   0.046    -34.18235   -.3156123
                        12   |  -16.99621   8.000418    -2.12   0.034    -32.67674    -1.31568
                        15   |  -16.60315   11.35101    -1.46   0.144    -38.85073    5.644428
                             |
       casepack#c.capacityr  |
                         3   |  -1.627422   1.044781    -1.56   0.119    -3.675156     .420311
                         4   |  -3.837301   .3289252   -11.67   0.000    -4.481982   -3.192619
                         8   |          0  (omitted)
                        12   |  -.0523975   .0646855    -0.81   0.418    -.1791788    .0743838
                        15   |    .096489   .3679361     0.26   0.793    -.6246525    .8176306
                             |
                   pingint   |
                         1   |  -1.648932    1.50916    -1.09   0.275    -4.606832    1.308968
                         2   |    .991843   1.870381     0.53   0.596    -2.674037    4.657723
                         3   |  -.6540911   1.218006    -0.54   0.591     -3.04134    1.733157
                             |
                    vendor   |
                         2   |  -8.635879   4.086856    -2.11   0.035    -16.64597   -.6257887
                         3   |   1.850535   2.056727     0.90   0.368    -2.180575    5.881645
                         4   |   1.117814   2.018143     0.55   0.580    -2.837674    5.073301
                         5   |  -1.329979   2.758302    -0.48   0.630    -6.736151    4.076193
                             |
                     store   |
                      6515   |   1.224667   1.429229     0.86   0.392    -1.576571    4.025905
                      6520   |  -1.029785   1.845497    -0.56   0.577    -4.646893    2.587322
                      6539   |   1.193236   1.755136     0.68   0.497    -2.246767     4.63324
                             |
                       _cons |  -9.875748   8.305193    -1.19   0.234    -26.15363    6.402132
                ln(openwk)   |          1  (exposure)
-----------------------------+------------------------------------------------------------------
eventrate                    |
       casepack#c.capacityr  |
                         4   |  -.0134475   .0276799    -0.49   0.627    -.0676991    .0408041
                         8   |  -.2160552   .0412684    -5.24   0.000    -.2969397   -.1351707
                        12   |  -.2015776   .0426558    -4.73   0.000    -.2851815   -.1179738
                        15   |  -.1731639   .0342966    -5.05   0.000     -.240384   -.1059437
                             |
                 capacityr   |   .2002424   .0451363     4.44   0.000     .1117768     .288708
                             |
                   soldhat   |
                       L1.   |   .0202267   .0035843     5.64   0.000     .0132016    .0272518
                             |
                    vendor   |
                         2   |  -1.114027   .4794021    -2.32   0.020    -2.053638    -.174416
                         3   |   .6422445   .2348551     2.73   0.006     .1819369    1.102552
                         4   |   .2785198   .2162087     1.29   0.198    -.1452414     .702281
                         5   |  -.2658868   .3743439    -0.71   0.478    -.9995875    .4678138
                             |
```

week						
1	0	(empty)				
2	0	(omitted)				
3	-.4093416	.121561	-3.37	0.001	-.6475968	-.1710863
4	.0696827	.096254	0.72	0.469	-.1189716	.258337
5	.0098683	.1076246	0.09	0.927	-.2010721	.2208087
6	-.2318675	.1145406	-2.02	0.043	-.4563629	-.0073721
7	-.6575025	.1137955	-5.78	0.000	-.8805376	-.4344673
8	0	(omitted)				
_cons	-10.35927	.3607862	-28.71	0.000	-11.06639	-9.652138
ln(openwk)	1	(exposure)				
/soldwk						
lnalpha	-1.446489	.0466918			-1.538003	-1.354975
/eventrate						
lnalpha	.8431637	.0875523			.6715643	1.014763
var(e.durrate)	247.9128	28.63497			197.6886	310.8968

GSEM of yogurt category below.

(Std. Err. adjusted for 344 clusters in storeupc)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
soldwk						
c.durratehaty#c.eventratehaty	-.0409285	.0160664	-2.55	0.011	-.072418	-.009439
durratehaty	.1354977	.0225631	6.01	0.000	.0912748	.1797205
durratehat2y	.0032287	.0012385	2.61	0.009	.0008012	.0056562
regprice	-.5090604	.1285852	-3.96	0.000	-.7610827	-.2570382
percdisc	.0000637	.0073307	0.01	0.993	-.0143043	.0144317
shelfno						
4	.0403655	.0466112	0.87	0.386	-.0509907	.1317217
5	-.1543446	.0415224	-3.72	0.000	-.2357269	-.0729623
6	-.2719564	.05738	-4.74	0.000	-.3844192	-.1594937
store						
6515	.184114	.0613424	3.00	0.003	.0638851	.3043429
6520	.0323712	.0783676	0.41	0.680	-.1212265	.1859689
6539	-.5511514	.06301	-8.75	0.000	-.6746486	-.4276541
week						
1	0	(empty)				
2	-.3965897	.0361131	-10.98	0.000	-.4673701	-.3258094
3	-.2821866	.0254559	-11.09	0.000	-.3320792	-.232294
4	-.1027701	.0355672	-2.89	0.004	-.1724805	-.0330597
5	-.1266155	.0326392	-3.88	0.000	-.1905871	-.0626439
6	-.1211434	.0289853	-4.18	0.000	-.1779535	-.0643334
7	.024517	.0478643	0.51	0.608	-.0692953	.1183293
8	0	(omitted)				
vendor						
3	-.8501474	.0570897	-14.89	0.000	-.9620412	-.7382536
4	-.6884022	.069725	-9.87	0.000	-.8250607	-.5517437
_cons	-3.357125	.1595901	-21.04	0.000	-3.669916	-3.044334
ln(openwk)	1	(exposure)				
durrate						
eventratehaty	16.18051	2.454376	6.59	0.000	11.37002	20.991
eventratehat2y	-2.433799	.4858767	-5.01	0.000	-3.3861	-1.481498
casepack#c.capacityr						
12	-.0377243	.0589333	-0.64	0.522	-.1532314	.0777828
pingint						
1	-1.961593	1.676656	-1.17	0.242	-5.247778	1.324591
2	-1.999583	1.911807	-1.05	0.296	-5.746657	1.74749
3	-1.186295	1.219804	-0.97	0.331	-3.577066	1.204477
vendor						
3	1.650018	1.876104	0.88	0.379	-2.027078	5.327115
4	1.323544	1.825226	0.73	0.468	-2.253833	4.90092
store						
6515	-3.648518	1.66722	-2.19	0.029	-6.91621	-.3808261
6520	-6.3298	2.073778	-3.05	0.002	-10.39433	-2.265271
6539	-3.324514	1.864477	-1.78	0.075	-6.978821	.3297934
_cons	-22.573	3.439246	-6.56	0.000	-29.3138	-15.8322
ln(openwk)	1	(exposure)				
eventrate						
capacityr	.0003902	.006907	0.06	0.955	-.0131473	.0139276
soldhaty						
L1.	.0133218	.0031741	4.20	0.000	.0071007	.019543
vendor						
3	.4164862	.2307966	1.80	0.071	-.0358669	.8688393
4	.1075211	.2144796	0.50	0.616	-.3128511	.5278934
week						
1	0	(empty)				
2	0	(omitted)				
3	-.3056056	.1535513	-1.99	0.047	-.6065606	-.0046505
4	.2197407	.1185729	1.85	0.064	-.0126579	.4521394
5	.0598243	.1319394	0.45	0.650	-.1987722	.3184208
6	-.1388169	.1402351	-0.99	0.322	-.4136727	.1360389
7	-.8308943	.1548542	-5.37	0.000	-1.134403	-.5273857
8	0	(omitted)				

_cons	-10.05826	.3540676	-28.41	0.000	-10.75222	-9.364296
ln(openwk)	1	(exposure)				
/soldwk						
lnalpha	-1.761846	.050425			-1.860678	-1.663015
/eventrate						
lnalpha	1.01437	.1028234			.8128398	1.2159
var(e.durrate)	190.6538	35.85085			131.8816	275.6175

GSEM of isotonics category below

(Std. Err. adjusted for 248 clusters in storeupc)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
soldwk						
c.durratehati#c.eventratehati	-.0124763	.0025079	-4.97	0.000	-.0173918	-.0075609
durratehati	.0631585	.0060718	10.40	0.000	.0512581	.0750589
durratehat2i	.0002375	.0001878	1.27	0.206	-.0001305	.0006056
regprice	-.2747764	.0233371	-11.77	0.000	-.3205162	-.2290367
percdisc	-.0190092	.0084759	-2.24	0.025	-.0356217	-.0023967
shelfno						
2	-.0021966	.1229927	-0.02	0.986	-.2432578	.2388646
3	-.0159347	.1122588	-0.14	0.887	-.2359578	.2040885
4	-.1497357	.1397455	-1.07	0.284	-.4236319	.1241604
5	.1436468	.1290794	1.11	0.266	-.1093442	.3966377
6	.2480206	.132666	1.87	0.062	-.0120001	.5080412
7	.1650726	.1309031	1.26	0.207	-.0914927	.421638
store						
6515	-.4704022	.0861782	-5.46	0.000	-.6393083	-.301496
6520	-.5259628	.0837061	-6.28	0.000	-.6900237	-.3619018
6539	-.7177733	.0879922	-8.16	0.000	-.8902349	-.5453118
week						
1	0	(empty)				
2	-.3760767	.0462666	-8.13	0.000	-.4667576	-.2853959
3	-.2071481	.0440171	-4.71	0.000	-.2934199	-.1208762
4	.0575368	.0367225	1.57	0.117	-.014438	.1295116
5	-.0716756	.0412022	-1.74	0.082	-.1524305	.0090793
6	-.2517473	.0431581	-5.83	0.000	-.3363356	-.1671589
7	-.1104077	.0492615	-2.24	0.025	-.2069583	-.013857
8	0	(omitted)				
5.vendor	-.6512163	.0769051	-8.47	0.000	-.8019475	-.5004851
_cons	-3.795033	.2236171	-16.97	0.000	-4.233315	-3.356752
ln(openwk)	1	(exposure)				
durrate						
eventratehati	31.02856	5.819949	5.33	0.000	19.62167	42.43545
eventratehat2i	-7.112925	2.216356	-3.21	0.001	-11.4569	-2.768946
casepack						
4	8.416574	8.223595	1.02	0.306	-7.701376	24.53452
8	-11.5548	8.744407	-1.32	0.186	-28.69352	5.583922
12	-24.01213	12.47778	-1.92	0.054	-48.46812	.4438637
15	-15.86799	11.67275	-1.36	0.174	-38.74616	7.010169
casepack#c.capacityr						
3	-1.622802	1.031903	-1.57	0.116	-3.645295	.3996914
4	-2.428336	.2982795	-8.14	0.000	-3.012953	-1.843719
8	0	(omitted)				
12	.5018692	.4303861	1.17	0.244	-.341672	1.34541
15	.1863941	.3922602	0.48	0.635	-.5824217	.9552099
pingint						
1	.8487062	3.041169	0.28	0.780	-5.111876	6.809288
2	10.58209	3.926377	2.70	0.007	2.886533	18.27765
3	1.803567	2.58534	0.70	0.485	-3.263607	6.870741
5.vendor	5.839442	3.260506	1.79	0.073	-.5510316	12.22992
store						
6515	7.876807	2.037996	3.86	0.000	3.882409	11.87121
6520	9.244153	3.124211	2.96	0.003	3.120811	15.36749
6539	7.556991	3.544003	2.13	0.033	.6108726	14.50311
_cons	-35.6128	8.504204	-4.19	0.000	-52.28074	-18.94487
ln(openwk)	1	(exposure)				
eventrate						
casepack#c.capacityr						
4	-.0268496	.0267203	-1.00	0.315	-.0792205	.0255213
8	-.1991695	.0437389	-4.55	0.000	-.2848962	-.1134429
12	-.1746633	.0554144	-3.15	0.002	-.2832736	-.066053
15	-.1468822	.0525934	-2.79	0.005	-.2499634	-.043801
capacityr	.1339041	.0800245	1.67	0.094	-.022941	.2907491
soldhati						
L1.	.0372405	.0072403	5.14	0.000	.0230498	.0514313
5.vendor	.8907706	.3216041	2.77	0.006	.2604382	1.521103
week						
1	0	(empty)				
2	0	(omitted)				
3	-.6494756	.1964653	-3.31	0.001	-1.03454	-.2644107
4	-.1904643	.169157	-1.13	0.260	-.522006	.1410774
5	-.056953	.1810418	-0.31	0.753	-.4117883	.2978824

6	-.405322	.1851956	-2.19	0.029	-.7682987	-.0423452
7	-.4415171	.1719652	-2.57	0.010	-.7785628	-.1044715
8	0	(omitted)				
_cons	-11.02508	.5982408	-18.43	0.000	-12.19761	-9.852548
ln(openwk)	1	(exposure)				
/soldwk						
lnalpha	-1.346003	.066592			-1.476521	-1.215486
/eventrate						
lnalpha	.4309983	.1960927			.0466637	.815333
var(e.durrate)	310.8433	45.70905			233.0095	414.6765

3SLS of overall data below Three-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
Sales	4,144	24	.5316718	0.5162	4423.81	0.0000
Duration	4,144	20	6.682142	0.0559	315.97	0.0000
Frequency	4,144	16	2.084855	0.0360	254.94	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Sales						
c.durratehattob#c.eventratehattob	-.0155254	.0015258	-10.18	0.000	-.0185158	-.0125349
durratehattob	.0890182	.0035437	25.12	0.000	.0820728	.0959637
durratehat2tob	.003295	.0003186	10.34	0.000	.0026706	.0039193
regprice	-.2283972	.0078348	-29.15	0.000	-.2437531	-.2130413
percdisc	-.0038958	.0030003	-1.30	0.194	-.0097764	.0019848
shelfno						
2	-.1420456	.0505781	-2.81	0.005	-.2411768	-.0429144
3	-.1723017	.0483488	-3.56	0.000	-.2670636	-.0775397
4	-.0796815	.0487783	-1.63	0.102	-.1752853	.0159223
5	-.2673021	.047369	-5.64	0.000	-.3601437	-.1744605
6	-.2361926	.0491212	-4.81	0.000	-.3324685	-.1399167
7	.0088085	.0532072	0.17	0.869	-.0954756	.1130926
store						
6515	-.2474233	.0266005	-9.30	0.000	-.2995593	-.1952873
6520	-.0545089	.0236277	-2.31	0.021	-.1008184	-.0081994
6539	-.3933783	.0316297	-12.44	0.000	-.4553714	-.3313853
week						
3	.4159881	.0332153	12.52	0.000	.3508873	.4810888
4	.1670642	.0313146	5.34	0.000	.1056888	.2284397
5	.0775368	.0310192	2.50	0.012	.0167403	.1383332
6	.2208672	.0315531	7.00	0.000	.1590242	.2827102
7	.6794686	.0370909	18.32	0.000	.6067717	.7521655
8	.3017665	.0321003	9.40	0.000	.2388512	.3646819
vendor						
2	-.150305	.0471246	-3.19	0.001	-.2426674	-.0579426
3	-1.050835	.0344197	-30.53	0.000	-1.118297	-.9833739
4	-.676149	.0289808	-23.33	0.000	-.7329503	-.6193477
5	-1.059674	.0335619	-31.57	0.000	-1.125454	-.9938941
_cons	3.072414	.0735519	41.77	0.000	2.928255	3.216573
Duration						
eventratehattob	.9040629	.1453879	6.22	0.000	.619108	1.189018
eventratehat2tob	.0478198	.0306015	1.56	0.118	-.012158	.1077976
casepack						
4	1.560266	2.212534	0.71	0.481	-2.776222	5.896753
8	-1.804566	1.481137	-1.22	0.223	-4.70754	1.098409
12	-1.805038	1.203716	-1.50	0.134	-4.164279	.554202
15	1.568203	1.500104	1.05	0.296	-1.371947	4.508353
casepack#c.capacityr						
3	-.0242728	.1642494	-0.15	0.883	-.3461957	.29765
4	-.0532545	.2387119	-0.22	0.823	-.5211212	.4146122
8	0	(omitted)				
12	-.0292175	.0124455	-2.35	0.019	-.0536103	-.0048247
15	-.1100899	.0482394	-2.28	0.022	-.2046374	-.0155424
pingint						
1	-.4247695	.5456164	-0.78	0.436	-1.494158	.644619
2	.3092766	.5482344	0.56	0.573	-.7652432	1.383796
3	-.5867818	.3526966	-1.66	0.096	-1.278054	.1044908
vendor						
2	-1.641367	.9075356	-1.81	0.071	-3.420104	.1373699
3	.2491322	.4181822	0.60	0.551	-.5704898	1.068754
4	-.1742002	.3989341	-0.44	0.662	-.9560966	.6076962
5	-1.00841	.6852112	-1.47	0.141	-2.351399	.3345798
store						
6515	.9431445	.2334189	4.04	0.000	.4856519	1.400637
6520	.2517912	.3669885	0.69	0.493	-.4674929	.9710754
6539	1.399242	.3664369	3.82	0.000	.6810394	2.117445
_cons	5.861338	1.352043	4.34	0.000	3.211382	8.511295
Frequency						
casepack#c.capacityr						
4	.0608583	.0210984	2.88	0.004	.0195062	.1022104
8	-.1371298	.0309496	-4.43	0.000	-.1977899	-.0764698
12	-.1274727	.0244421	-5.22	0.000	-.1753783	-.0795671
15	-.1001426	.0196999	-5.08	0.000	-.1387538	-.0615315

capacityr	.0971081	.0254375	3.82	0.000	.0472516	.1469645
soldwk						
L1.	.0132419	.0016676	7.94	0.000	.0099735	.0165103
vendor						
2	-1.669743	.2480004	-6.73	0.000	-2.155815	-1.183672
3	.2024433	.1315274	1.54	0.124	-.0553457	.4602322
4	-.3040025	.1228165	-2.48	0.013	-.5447183	-.0632866
5	-1.180368	.2028546	-5.82	0.000	-1.577956	-.7827802
week						
3	.1787818	.1008221	1.77	0.076	-.0188258	.3763894
4	.0413556	.0949851	0.44	0.663	-.1448118	.2275229
5	.0360902	.0943696	0.38	0.702	-.1488707	.2210512
6	-.0851911	.0948015	-0.90	0.369	-.2709986	.1006165
7	.1114438	.1013865	1.10	0.272	-.0872702	.3101577
8	.168126	.0958074	1.75	0.079	-.0196531	.3559051
_cons	1.664121	.1991809	8.35	0.000	1.273733	2.054508

Endogenous variables: ratsold durrate ratfrq

3SLS model of yogurt category below Three-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
Sales	2,408	19	.5733566	0.4558	2019.05	0.0000
Duration	2,408	11	5.856693	0.0240	106.07	0.0000
Frequency	2,408	10	2.222085	0.0090	71.55	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Sales						
c.durratehatytob#c.eventratehatytob	-.038272	.0046214	-8.28	0.000	-.0473298	-.0292143
durratehatytob	.137383	.005661	24.27	0.000	.1262876	.1484783
durratehat2ytob	.0084605	.0011408	7.42	0.000	.0062245	.0106965
regprice	-.3794771	.0881427	-4.31	0.000	-.5522336	-.2067206
percdisc	.0071354	.0087005	0.82	0.412	-.0099173	.0241881
shelfno						
4	.0818151	.0356651	2.29	0.022	.0119128	.1517174
5	-.1396099	.0324888	-4.30	0.000	-.2032867	-.0759331
6	-.1782687	.0387615	-4.60	0.000	-.2542397	-.1022976
store						
6515	-.211128	.0483496	-4.37	0.000	-.3058915	-.1163645
6520	.0268053	.0700216	0.38	0.702	-.1104346	.1640451
6539	-.5766393	.0373434	-15.44	0.000	-.649831	-.5034476
week						
3	.6122884	.0494961	12.37	0.000	.5152778	.709299
4	.1519197	.0445986	3.41	0.001	.0645081	.2393314
5	.123292	.0445965	2.76	0.006	.0358844	.2106995
6	.4165239	.0461646	9.02	0.000	.3260431	.5070048
7	1.27411	.068327	18.65	0.000	1.140192	1.408028
8	.5275981	.0487541	10.82	0.000	.4320417	.6231544
vendor						
3	-1.162602	.0446226	-26.05	0.000	-1.250061	-1.075144
4	-.7138951	.0510323	-13.99	0.000	-.8139166	-.6138736
_cons	3.289359	.1150435	28.59	0.000	3.063878	3.51484
Duration						
eventratehatytob	1.018951	.188668	5.40	0.000	.6491686	1.388734
eventratehat2ytob	.0787645	.0450059	1.75	0.080	-.0094455	.1669745
12.casepack	0	(omitted)				
casepack#c.capacityr						
12	-.041056	.0112352	-3.65	0.000	-.0630767	-.0190354
pingint						
1	-1.251655	.6276571	-1.99	0.046	-2.48184	-.0214697
2	-.1878627	.6307011	-0.30	0.766	-1.424014	1.048289
3	-.7978674	.4098833	-1.95	0.052	-1.601224	.005489
vendor						
3	.1742404	.3674657	0.47	0.635	-.5459791	.8944598
4	-.3778871	.3511504	-1.08	0.282	-1.066129	.310355
store						
6515	1.015096	.2785356	3.64	0.000	.4691765	1.561016
6520	-.1833323	.421044	-0.44	0.663	-1.008563	.6418987
6539	.1478079	.4206899	0.35	0.725	-.6767292	.972345
_cons	5.288707	.7050983	7.50	0.000	3.90674	6.670674
Frequency						
casepack#c.capacityr						
12	0	(omitted)				
capacityr	-.0305332	.0044819	-6.81	0.000	-.0393176	-.0217488
soldwk						
L1.	.0134669	.0019812	6.80	0.000	.0095838	.01735
vendor						
3	.205869	.1407614	1.46	0.144	-.0700183	.4817562
4	-.3039151	.1321128	-2.30	0.021	-.5628514	-.0449788
week						

3	.0689278	.1409221	0.49	0.625	-.2072744	.3451299
4	.0765051	.1341228	0.57	0.568	-.1863709	.339381
5	.0414562	.1324574	0.31	0.754	-.2181555	.301068
6	.0137861	.133057	0.10	0.917	-.2470009	.274573
7	-.1349018	.1472284	-0.92	0.360	-.4234642	.1536606
8	.0430276	.13493	0.32	0.750	-.2214304	.3074856
_cons	1.709608	.2255545	7.58	0.000	1.26753	2.151687

3SLS model of yogurt category below Three-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
Sales	1,736	21	.3545712	0.5019	1778.57	0.0000
Duration	1,736	17	7.688901	0.0725	216.16	0.0000
Frequency	1,736	13	1.861855	0.0971	253.01	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Sales						
c.durratehatitob#c.eventratehatitob	-.007	.0008325	-8.41	0.000	-.0086316	-.0053684
durratehatitob	.0293583	.0020781	14.13	0.000	.0252854	.0334313
durratehat2itob	.0008158	.0001291	6.32	0.000	.0005628	.0010688
regprice	-.114067	.0059714	-19.10	0.000	-.1257708	-.1023632
percdisc	-.0042301	.0021683	-1.95	0.051	-.0084798	.0000196
shelfno						
2	-.0628825	.0335115	-1.88	0.061	-.1285638	.0027989
3	-.0018555	.0364846	-0.05	0.959	-.0733639	.0696529
4	-.0505452	.0387365	-1.30	0.192	-.1264674	.025377
5	.1505467	.0361948	4.16	0.000	.0796063	.2214872
6	.2691815	.0385856	6.98	0.000	.1935552	.3448079
7	.2298183	.0362604	6.34	0.000	.1587492	.3008874
store						
6515	-.3662353	.0354877	-10.32	0.000	-.4357901	-.2966806
6520	-.2185752	.0345319	-6.33	0.000	-.2862566	-.1508939
6539	-.2475886	.0358757	-6.90	0.000	-.3179038	-.1772734
week						
3	.2820107	.0342846	8.23	0.000	.2148141	.3492073
4	.1811149	.0322494	5.62	0.000	.1179072	.2443227
5	.0577984	.0317951	1.82	0.069	-.0045189	.1201156
6	.0530624	.031942	1.66	0.097	-.0095429	.1156676
7	.3221233	.0334887	9.62	0.000	.2564868	.3877599
8	.1744754	.031991	5.45	0.000	.1117742	.2371767
5.vendor	-.6188997	.0295688	-20.93	0.000	-.6768534	-.560946
_cons	1.794211	.0747053	24.02	0.000	1.647791	1.94063
Duration						
eventratehatitob	1.27102	.264899	4.80	0.000	.7518272	1.790212
eventratehat2itob	.091128	.0452887	2.01	0.044	.0023639	.1798921
casepack						
4	.8750119	2.458627	0.36	0.722	-3.943808	5.693832
8	-.1722104	1.772388	-0.10	0.923	-3.646027	3.301606
12	-2.526931	2.599635	-0.97	0.331	-7.622122	2.568261
15	3.414562	1.658986	2.06	0.040	.1630098	6.666115
casepack#c.capacityr						
3	.1359556	.2000643	0.68	0.497	-.2561632	.5280744
4	.1310574	.2726617	0.48	0.631	-.4033497	.6654644
8	0	(omitted)				
12	.0885588	.1060755	0.83	0.404	-.1193454	.296463
15	-.1286785	.060439	-2.13	0.033	-.2471367	-.0102202
pingint						
1	1.166486	.9400402	1.24	0.215	-.6759585	3.008932
2	.5069073	.9475194	0.53	0.593	-1.350197	2.364011
3	.0758145	.6099878	0.12	0.901	-1.11974	1.271369
store						
6515	-.4239262	.3988043	-1.06	0.288	-1.205568	.3577159
6520	-.1387679	.6284708	-0.22	0.825	-1.370548	1.093012
6539	2.139641	.6295139	3.40	0.001	.9058167	3.373466
5.vendor	.3378963	.6379194	0.53	0.596	-.9124027	1.588195
_cons	3.240858	1.753155	1.85	0.065	-.1952621	6.676978
Frequency						
casepack#c.capacityr						
4	.027961	.0194482	1.44	0.151	-.0101568	.0660789
8	-.1383095	.0285239	-4.85	0.000	-.1942153	-.0824037
12	-.134534	.0268429	-5.01	0.000	-.1871452	-.0819228
15	-.1120758	.0257752	-4.35	0.000	-.1625943	-.0615572
capacityr	.1032825	.0355349	2.91	0.004	.0336353	.1729297
soldwk						
L1.	.0238002	.0034344	6.93	0.000	.0170688	.0305316
5.vendor	.5351283	.1213183	4.41	0.000	.2973487	.7729078
week						
3	.2537189	.1360028	1.87	0.062	-.0128416	.5202794
4	-.036342	.127832	-0.28	0.776	-.2868881	.2142041
5	.026783	.1275816	0.21	0.834	-.2232724	.2768384
6	-.2006957	.1274064	-1.58	0.115	-.4504077	.0490163
7	.4582082	.1297495	3.53	0.000	.203904	.7125125
8	.3968424	.1293148	3.07	0.002	.1433901	.6502947
_cons	-.1961274	.3106033	-0.63	0.528	-.8048987	.4126439

V. Stockout duration (daily item stockout duration in minutes) compared over groups

Appendices IV thru VII are Bonferroni-corrected (5%) pairwise comparisons of means. This specific set of tests shows that the Isotonics product category has significantly different SHO duration in all but two stores. Store 6515 and 6520 have mean SHO durations that do not differ from one another. In the yogurt product category, only store 6539 has a significantly different mean SHO duration and is much larger than the other three stores.

```
Pairwise comparisons of means with equal variances

over         : yogurt store
```

	Number of Comparisons
yogurt#store	28

SOweekmin	Mean	Std. Err.	Bonferroni Groups
yogurt#store			
0 6504	177.8679	17.23585	A
0 6515	331.5939	17.4352	BC
0 6520	382.2758	18.71875	C
0 6539	537.3191	17.89341	
1 6504	170.7961	27.70729	A
1 6515	258.537	26.84169	AB
1 6520	253.6156	28.43977	AB
1 6539	697.4373	28.13987	

```
Note: Means sharing a letter in the group label
      are not significantly different at the 5%
      level.
```

VI. Stockout frequency (number of stockout events/day) compared over groups

Appendices IV thru VII are Bonferroni-corrected (5%) pairwise comparisons of means. This specific set of tests shows that the Isotonics product category has significantly different mean SHO frequencies for all but two stores. Store 6515 and 6520 have mean SHO frequencies that do not significantly differ from one another. For Yogurt, stores 6504 and 6520 have significantly different mean SHO frequency, while the other stores have greater mean SHO frequencies.

```
Pairwise comparisons of means with equal variances

over         : yogurt store
```

	Number of Comparisons
yogurt#store	28

soskuweek	Mean	Std. Err.	Bonferroni Groups
yogurt#store			
0 6504	.3948404	.0293081	
0 6515	.8577181	.0296471	B
0 6520	.8994327	.0318297	B
0 6539	1.235627	.0304263	
1 6504	.6011299	.047114	A
1 6515	.9247084	.0456421	B
1 6520	.6547619	.0483595	A
1 6539	2.212121	.0478495	

```
Note: Means sharing a letter in the group label
      are not significantly different at the 5%
      level.
```

VII. Daily item sales (number units sold/day) compared over groups

Appendices IV thru VII are Bonferroni-corrected (5%) pairwise comparisons of means. This specific set of tests shows that for the Isotonics product category the number of units sold per day is not significantly different between stores 6504 and 6520. For Yogurt, each store has a significantly different mean daily sales.

Pairwise comparisons of means with equal variances

over : yogurt store

	Number of Comparisons
yogurt#store	28

sumSKUwksold	Mean	Std. Err.	Bonferroni Groups
yogurt#store			
0 6504	12.82903	.337948	A
0 6515	21.07919	.3418568	
0 6520	13.87983	.3670237	A
0 6539	18.31008	.350841	
1 6504	27.1096	.5432645	
1 6515	57.55992	.5262925	
1 6520	32.73929	.5576264	
1 6539	47.52098	.5517462	

Note: Means sharing a letter in the group label are not significantly different at the 5% level.

VIII. Weekly category sales (number units sold/day) compared over groups

Appendices IV thru VII are Bonferroni-corrected (5%) pairwise comparisons of means. This specific set of tests shows that the mean weekly sales of the Isotonics product category has two stores (6515 and 6539) where they are not significantly different from one another. Yogurt's weekly category sales significantly varies from store to store.

```
Pairwise comparisons of means with equal variances

over         : yogurt store
```

	Number of Comparisons
yogurt#store	28

catsales	Mean	Std. Err.	Bonferroni Groups
yogurt#store			
0 6504	1538.725	18.78374	
0 6515	1869.013	19.001	A
0 6520	1257.536	20.39982	
0 6539	1896.957	19.50035	A
1 6504	6205.49	30.19559	
1 6515	12185.93	29.25225	
1 6520	7642.076	30.99385	
1 6539	9529.381	30.66702	

```
Note: Means sharing a letter in the group label
      are not significantly different at the 5%
      level.
```

IX. Weeks during which each category was tracked

week	Isotonics	Yogurt	Total
1	62	8	70
2	147	152	299
3	200	267	467
4	200	266	466
5	200	264	464
6	200	264	464
7	200	264	464
8	200	257	457
9	200	256	456
10	201	256	457
11	201	256	457
12	201	252	453
13	200	243	443
14	200	237	437
15	200	233	433
16	200	225	425
17	200	221	421
18	200	220	420
19	200	209	409
20	200	162	362
21	199	139	338
22	199	126	325
23	198	86	284
24	198	34	232
25	198	29	227
26	198	3	201
27	198	0	198
28	197	0	197
29	195	0	195
30	194	0	194
31	193	0	193
32	192	0	192
33	191	0	191
34	190	0	190
35	189	0	189
36	188	0	188
37	185	0	185
38	185	0	185
39	172	0	172
40	170	0	170
41	166	0	166
42	165	0	165
43	154	0	154
44	152	0	152
45	145	0	145
46	135	0	135
47	127	0	127
48	105	0	105
49	91	0	91
50	65	0	65
51	53	0	53
52	43	0	43
53	26	0	26
Total	9,068	4,929	13,997

This table lists the number of unique items (SKUs) that are tracked by the shelf liner technology on a weekly basis by product category. While there are more Yogurt items that are tracked (267 unique items, maximum) than Isotonics items (201 items, maximum), the Isotonics product category is tracked for a longer period of time than the Yogurt category. The data is unbalanced because of these differences.

X. Capacity and case pack size combinations, shared and excess space combinations

Case packsize (units) \ Dedicated space (units)	4	7	8	10	11	16	18	24	30	36	48	54	72
3	170	2,134			325								
4			1,444			50							
8				200									
12							2,904	1,155	557	1,971	18	426	71
15								505	179				

This table above provides the combinations of dedicated space (shelf space that is set aside for one specific item) and case pack size for all of the items in this study sample. Shelf height and depth within the standard 4-shelf foot panels tend to vary in dimensions very little if at all. Similarly, items within a product category usually have items varying in dimension very little or in multiples (for example, one item offered as a single package and another item offered with 6 of the same single packets as one consumer unit. i.e. six-pack of drinks, etc.). This is why studies simulating all combination within an observed range of case pack size or shelf capacity may yield unrealistic functional estimates.

		Excess dedicated space in units				
		0	1	2	6	9
Units of case pack in shared space	0	4,888				
	1			325		
	2		2,304			
	6			200	3,887	505

The table above shows how case packs can be split between shared and dedicated (shelf) space. In the case of (0,0), the shelf capacity is an exact multiple of case pack size. In this study sample, 40.37% of the observations consist of items that are not postponed. These are 62 unique items out of the total 102 items, or 60.78% of the items in the study sample, as opposed to the often cited Gruen & Corsten (2008) figure of 91% of items being an exact multiple of case pack size.

XI. Testing for whether random effects is appropriate (Ch 4)

```
. xtreg soldwk c.durratehat#c.eventratehat durratehat durratehat2 regprice percdisc i.shelfno i.store i.week i.vendor, re

Random-effects GLS regression                   Number of obs     =      4,144
Group variable: storeupc                        Number of groups  =        592

R-sq:                                           Obs per group:
     within  = 0.1936                                         min =          7
     between = 0.5109                                         avg =        7.0
     overall = 0.4493                                         max =          7

                                                Wald chi2(24)     =    1584.90
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000
```

soldwk	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
c.durratehat#c.eventratehat	-.1821025	.0802339	-2.27	0.023	-.339358	-.024847
durratehat	.842215	.1642551	5.13	0.000	.520281	1.164149
durratehat2	.0263798	.0085267	3.09	0.002	.0096678	.0430917
regprice	-3.343056	.4649158	-7.19	0.000	-4.254274	-2.431838
percdisc	.0796464	.0654617	1.22	0.224	-.0486561	.2079489
shelfno						
2	.5735929	3.269215	0.18	0.861	-5.833951	6.981137
3	3.366985	3.08281	1.09	0.275	-2.675212	9.409181
4	6.60953	3.106976	2.13	0.033	.5199685	12.69909
5	-1.852992	3.044276	-0.61	0.543	-7.819664	4.113681
6	-.6361054	3.161275	-0.20	0.841	-6.832091	5.55988
7	1.100863	3.443391	0.32	0.749	-5.648058	7.849784
store						
6515	13.8092	1.569416	8.80	0.000	10.73321	16.8852
6520	2.928772	1.510391	1.94	0.052	-.0315397	5.889083
6539	7.62717	1.650744	4.62	0.000	4.391771	10.86257
week						
3	-4.21169	.6692962	-6.29	0.000	-5.523486	-2.899893
4	3.124541	.6071168	5.15	0.000	1.934614	4.314468
5	3.930507	.5915656	6.64	0.000	2.771059	5.089954
6	-.3917448	.6231078	-0.63	0.530	-1.613014	.8295241
7	-8.338811	.833182	-10.01	0.000	-9.971817	-6.705804
8	-1.045142	.6590106	-1.59	0.113	-2.336779	.2464947
vendor						
2	-24.59427	2.619517	-9.39	0.000	-29.72843	-19.46011
3	-22.98921	2.15836	-10.65	0.000	-27.21952	-18.7589
4	-17.77137	1.793959	-9.91	0.000	-21.28746	-14.25527
5	-28.20084	2.08629	-13.52	0.000	-32.28989	-24.11178
_cons	51.88196	4.015371	12.92	0.000	44.01197	59.75194
sigma_u	11.950863					
sigma_e	9.7713146					
rho	.5993374	(fraction of variance due to u_i)				

```
.
end of do-file

. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

        soldwk[storeupc,t] = Xb + u[storeupc] + e[storeupc,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                  soldwk |   531.5767       23.05595
                       e |   95.47859       9.771315
                       u |   142.8231       11.95086

        Test:   Var(u) = 0
                             chibar2(01) =  3820.59
                          Prob > chibar2 =   0.0000
```

Panel data can be studied with fixed effects or random effects models. The antecedents to SHO in Chapter 4 (case pack size and shelf capacity) do not vary over time. Additionally, the hypotheses include endogenous regressors and different dependent variable distributions (discrete and continuous). As such, GSEM is used as the primary specification (followed by a transformation of the discrete variable to be able to use 3SLS as an alternate specification). In the GSEM the errors are clustered by item (UPC) groups to account for random effects. However, traditionally a panel data regression is first run as a random effects model and then the Breusch and Pagan Lagrangian (BPL) multiplier test shows if a random effects model is appropriate. Below the significant (χ^2=3820, p<0.01) BPL test result after a random effects model on the transformed dependent variables and fitted endogenous regressors shows that there is a panel effect. There is a significant difference across items so that having robust errors clustered by item is appropriate.

Panel data can be studied with fixed effects or random effects models. The antecedents to SHO in Chapter 4 (case pack size and shelf capacity) do not vary over time. Additionally, the hypotheses include endogenous regressors and different dependent variable distributions (discrete and

XII. Centered variables for quadratic relationship

A comparison of using centered (demeaned) frequency variables. Columns 1, 4, and 7 do not have centering. Columns 2, 5, and 8 have both frequency and predicted frequency squared centered. When generating the frequency squared term, since the frequency is estimated with a negative binomial, it cannot be centered. For columns 3, 6, and 9 only the squared term is centered. The squared term is centered by first centering the predicted linear term and then squaring it. Since they worsen fit (higher AIC scores) and do not qualitatively change the results (all coefficients have the same sign and significance) centered terms are not included in the final model.

	(1) 1 All	(2) 4 AC	(3) 7 AC2	(4) 2 Yogurt	(5) 5 YC	(6) 8 YC2	(7) 3 Isoton~s	(8) 6 IC	(9) 9 IC2
Lagged weekly sales									
Duration	-0.002	-0.003	-0.002	-0.007	-0.008	-0.007	0.008	0.004	0.008
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.005)	(0.006)
Frequency	0.139***		0.143***	0.095***		0.106***	0.213***		0.215***
	(0.015)		(0.016)	(0.014)		(0.015)	(0.030)		(0.030)
Duration # Frequency	-0.002**		-0.002**	-0.000		-0.000	-0.006***		-0.006***
	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)
Frequency squared	0.039***								
	(0.008)								
Centered frequency		0.143***			0.106***			0.215***	
		(0.016)			(0.015)			(0.030)	
Duration # Centered		-0.002**			-0.000			-0.006***	
frequency		(0.001)			(0.001)			(0.001)	
Centered frequency squared		0.041***	0.041***						
		(0.007)	(0.007)						
Frequency squared				0.109***					
				(0.024)					
Centered frequency squared					0.122***	0.122***			
					(0.028)	(0.028)			
Frequency squared							0.030**		
							(0.009)		
Centered frequency squared								0.029**	0.029**
								(0.009)	(0.009)
Item shelf capacity									
Shipping box size	1.483***	1.483***	1.483***				1.226***	1.226***	1.226***
	(0.113)	(0.113)	(0.113)				(0.088)	(0.088)	(0.088)
Duration									
Shipping box size	-0.107	-0.108	-0.108				0.030	0.035	0.035
	(0.067)	(0.067)	(0.067)				(0.055)	(0.053)	(0.053)
Frequency	8.265***		8.253***	6.911***		6.890***	10.270***		10.265***
	(0.471)		(0.470)	(0.635)		(0.630)	(0.509)		(0.507)
Frequency squared	-0.288***								
	(0.087)								
Centered frequency		8.253***			6.890***			10.265***	
		(0.470)			(0.630)			(0.507)	
Centered frequency squared		-0.350***	-0.350***						
		(0.094)	(0.094)						
Frequency squared				-0.584**					
				(0.227)					
Centered frequency squared					-0.815**	-0.815**			
					(0.309)	(0.309)			
Frequency squared							-0.214		
							(0.312)		
Centered frequency squared								-0.289	-0.289
								(0.424)	(0.424)
L.soldwk									
Item shelf capacity	0.029***	0.029***	0.029***	0.019***	0.019***	0.019***	0.048***	0.048***	0.048***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)	(0.007)
Frequency									
L.Lagged weekly sales	0.014***	0.014***	0.014***	0.013***	0.013***	0.013***	0.029***	0.029***	0.029***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)
Item shelf capacity	-0.018**	-0.018**	-0.018**	-0.004	-0.004	-0.004	-0.102***	-0.102***	-0.102***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.019)	(0.019)	(0.019)
Observations	4736.000	4736.000	4736.000	2752.000	2752.000	2752.000	1984.000	1984.000	1984.000
AIC	144718.91	144771.6	144771.6	72432.514	72539.483	72539.483	53528.557	53539.56	53539.56
BIC	145087.3	145139.99	145139.99	72710.758	72817.727	72817.727	53830.572	53841.575	53841.575

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

XIII. Goodness of fit of dependent variable data in Chapter 4

There are multiple approaches to see what distribution observed data takes on: graphical approach, least squares comparisons and Chi-square or Kolmogorov-Smirnov (KS) tests. While the KS test is preferred over the Chi-square test when desiring a more conservative (higher) p value, the KS test gives valid values for discrete distributions. Even if testing a discrete distribution, if the sample size is less than 20 observations Chi square must be used instead of the KS test. In contrast, least squares comparison is an approach where observed data is compared against several theoretical distributions where test results are listed in ascending order of least square error. In this approach the distribution with the smallest least square error may still not be a good fit of the observed data. Last of all, the least statistically robust is the graphical approach. In the graphical approach the observed data is superimposed on a graph of a single theoretical distribution. The observed data may be displayed as a histogram along with the theoretical distribution, or as points on a Q-Q (or P-P) plot where the quality of fit is as good as the proximity of graph points to a 45-degree reference line.

After obtaining at least two possible theoretical distributions that may fit the observed data, the next step is running the econometric model assuming each distribution and comparing how tightly the models fit the observed data. The lower the AIC and BIC scores, the tighter the model fits observed data.

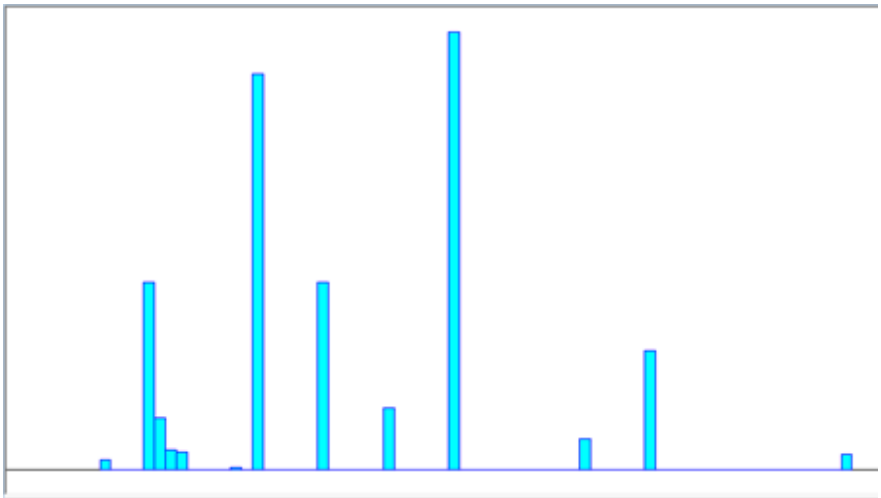
The following pages of this appendix section provides the various commands and outputs obtained from the fitting process described above of the following variables: Shelf capacity, SHO frequency, SHO duration, and Total weekly item sales. Discrete variables are presented with P-P plots instead of Q-Q plots.

Shelf Capacity

Shelf capacity has integer values calculated by multiplying each item's facing, height and depth for a volumetric measure of the number of units which fit into the dedicated space of each item. In the models in Chapter 4, the best fit is assuming a negative binomial distribution of the observed data.

Descriptive statistics and histogram

Descriptive statistics and histogram of shelf capacity provided by ARENA's Input Analyzer software:



```
        Data Summary

Number of Data Points   = 4736
Min Data Value          = 4
Max Data Value          = 72
Sample Mean             = 26.5
Sample Std Dev          = 14

        Histogram Summary

Histogram Range         = 3.5 to 72.5
Number of Intervals     = 69
```

Descriptive statistics of shelf capacity provided by EasyFit

Descriptive Statistics

Statistic	Value
Sample Size	4736
Range	68
Mean	26.331
Variance	201.73
Std. Deviation	14.203
Coef. of Variation	0.53941
Std. Error	3.9393
Skewness	0.58653
Excess Kurtosis	-0.01596

Percentile	Value
Min	4
5%	7
10%	7
25% (Q1)	18
50% (Median)	24
75% (Q3)	36
90%	48
95%	54
Max	72

There are slight differences in the mean and standard deviation between ARENA and EasyFit due to rounding.

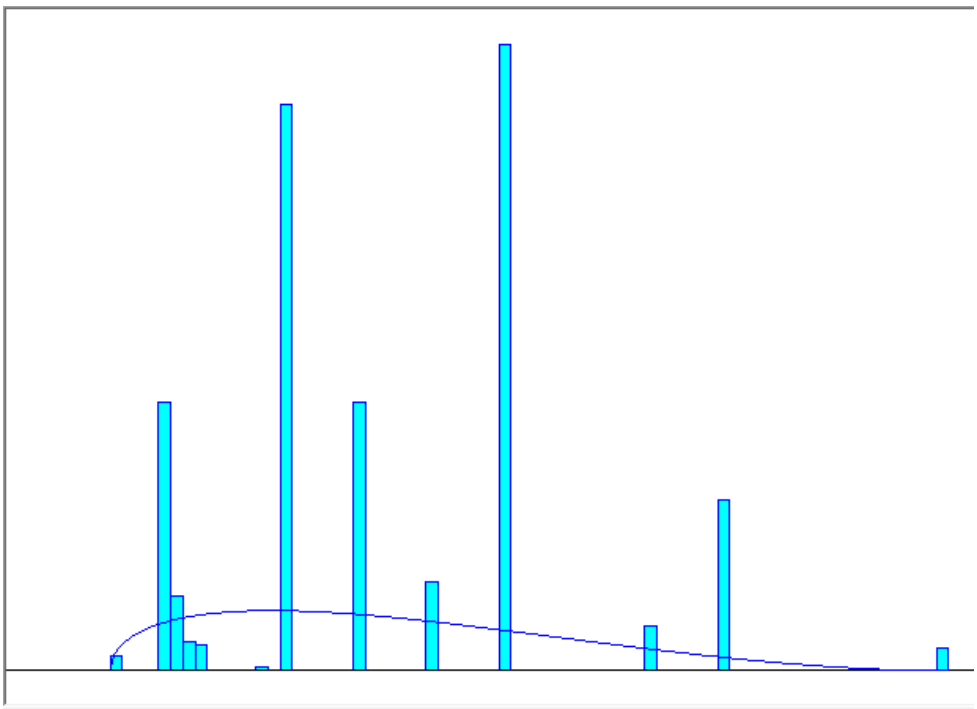
Least squares comparison approach

Function	Sq Error
Beta	0.165
Erlang	0.165
Gamma	0.165
Normal	0.166
Weibull	0.167
Lognormal	0.167
Exponential	0.17
Triangular	0.171
Uniform	0.171
Poisson	0.2

The least of the square errors for the capacity data appears to be a three-way tie between Beta, Erlang, and Gamma distributions—all 3 of which are closely related to one another.

These statistics are comparable to the sum of squares due to error statistics measuring the random error component of an econometric model.

Graphical approach and Chi-square test

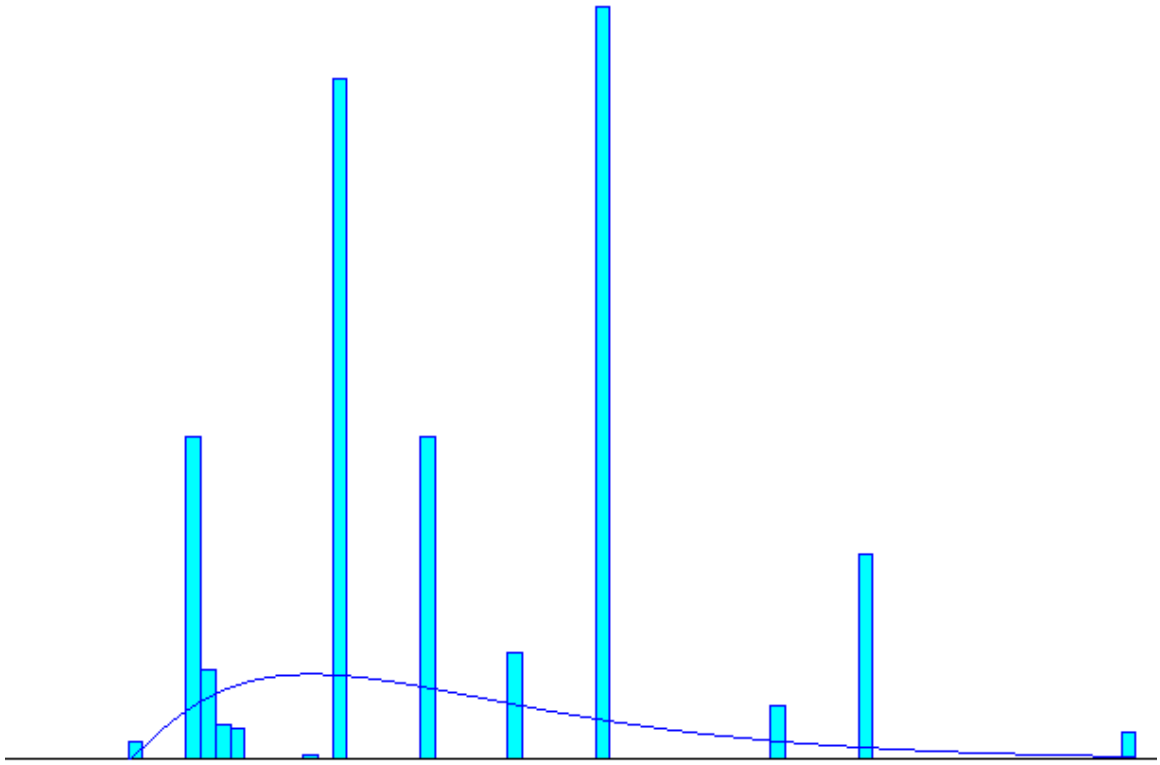


```
        Distribution Summary

Distribution:   Beta
Expression:     3.5 + 69 * BETA(1.46, 2.92)
Square Error:   0.164752

Chi Square Test
  Number of intervals   = 61
  Degrees of freedom    = 58
  Test Statistic        = 3.93e+004
  Corresponding p-value < 0.005
```

ARENA shows that the distribution with the lowest square error (Beta) is still significantly different than the observed data.

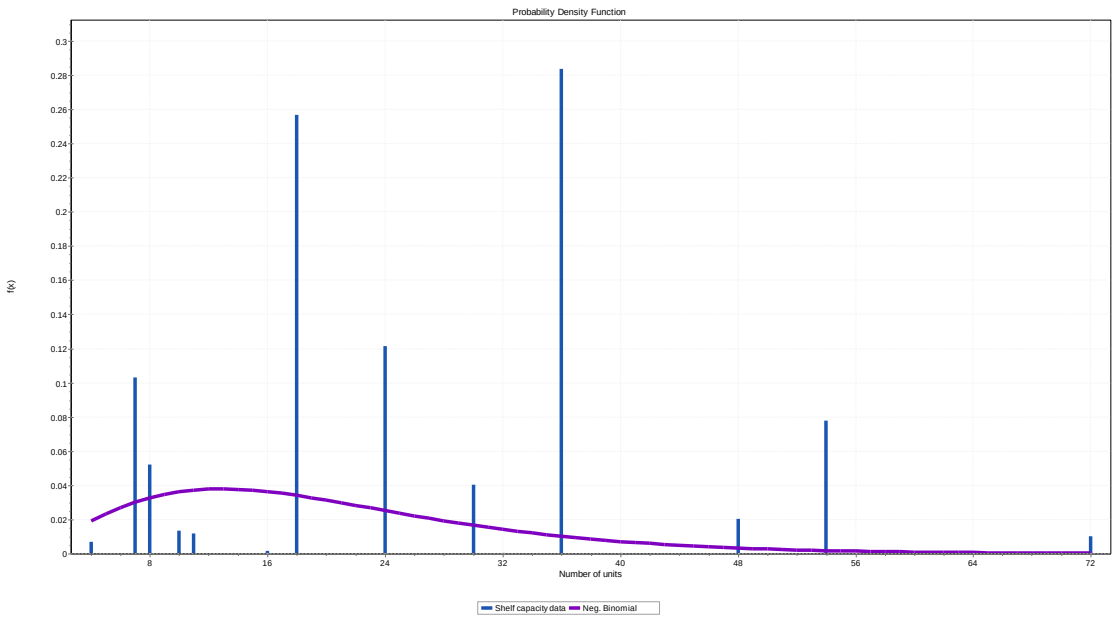
Here the graphical approach shows the Beta distribution

underestimates the mid-range observed values. Chi-square has a significant p value showing the observed data does not fit the Beta distribution. Similarly, the observed data also does not fit the gamma distribution at 5% (below).

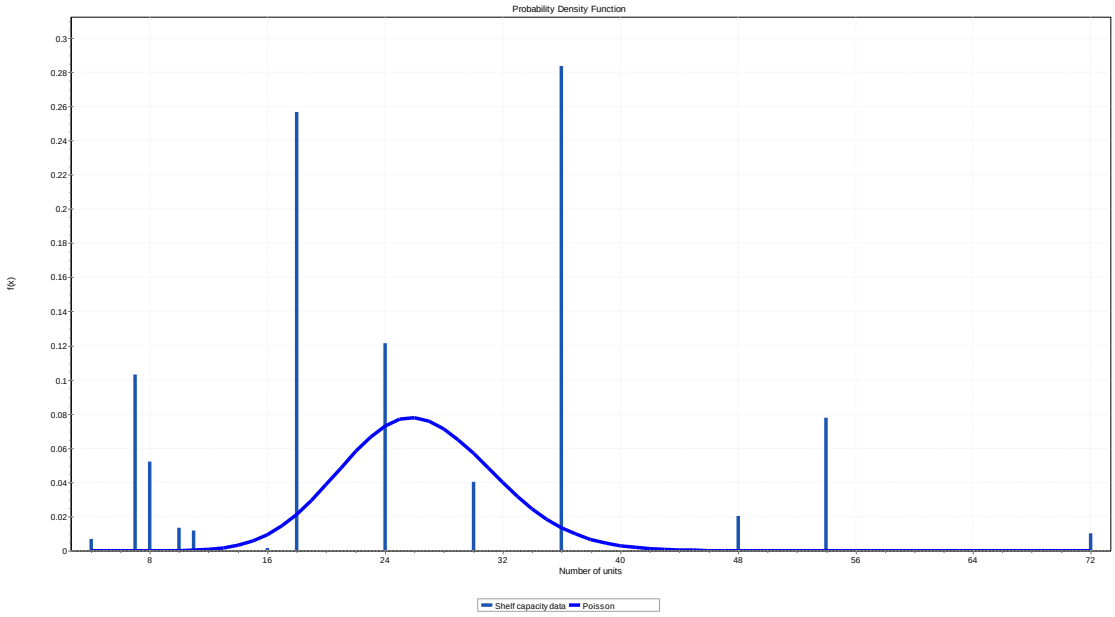


```
        Distribution Summary

Distribution:   Gamma
Expression:     3.5 + GAMM(10.5, 2.19)
Square Error:   0.165385

Chi Square Test
  Number of intervals   = 65
  Degrees of freedom    = 62
  Test Statistic        = 4.41e+004
  Corresponding p-value < 0.005
```

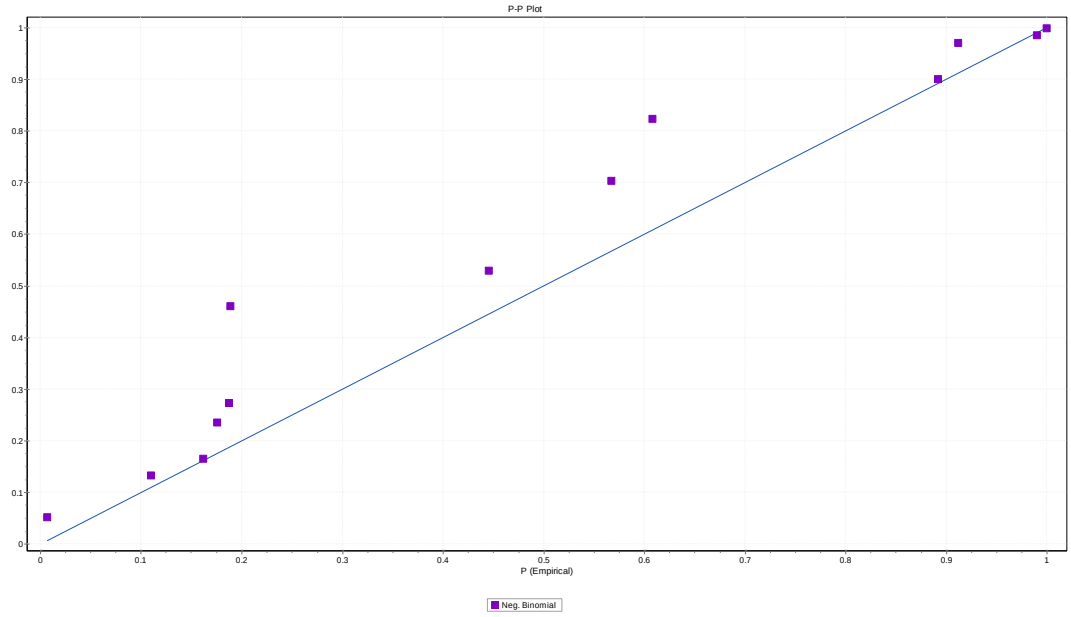
In comparison, EasyFit attempts to fit only discrete distributions to the shelf capacity data since it takes on only integer values.

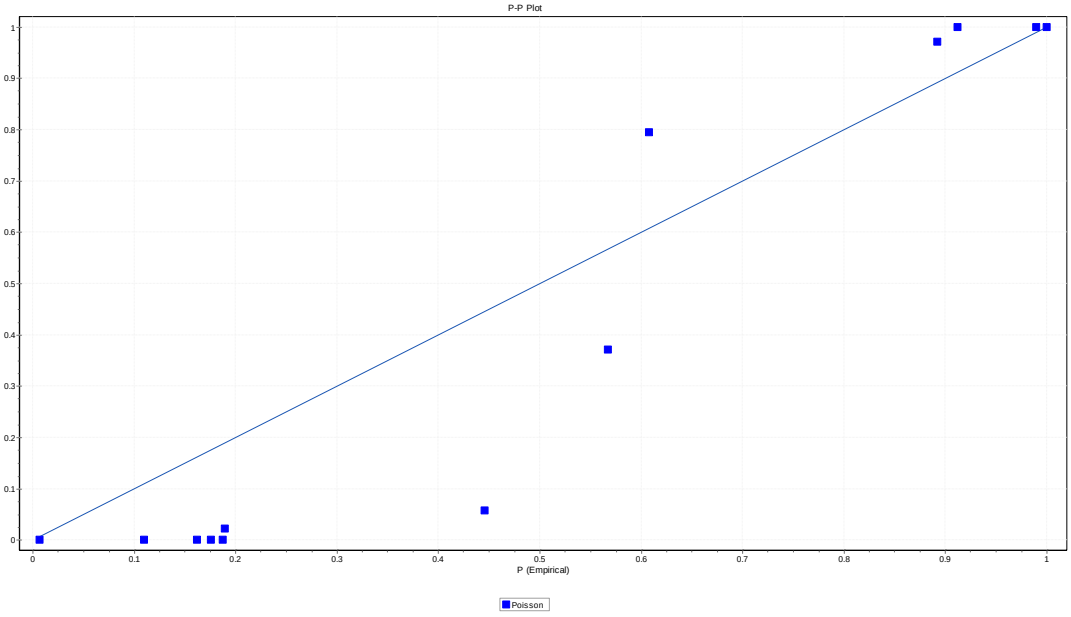


A histogram (above) of shelf capacity data with a negative binomial curve super-imposed in purple. The histogram (below) has a blue curve for the Poisson distribution.



Instead of Q-Q plots, since shelf capacity carries integer values, a P-P plot can compare the observed data to negative binomial (top) and Poisson (bottom) distributions.





The graph points should fall along the diagonal line and for shelf capacity the Poisson graph's data points are further away from the diagonal than the graph of negative binomial. In both cases the data appears to fit the theoretical distributions better at the ends of the reference lines. In contrast, a non-linear plot is acceptable if the graph points are further away from the reference line ends and are on or very close to being on the reference line in the middle range.

EasyFit provides the below test results for the binomial and the Poisson distributions. The test statistic is larger than the critical values (at every level of alpha) meaning that there are enough observations which show that the data does not fit the negative binomial and Poisson distributions.

Neg. Binomial [#4]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	4736 0.33976 0 3				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.01559	0.01777	0.01973	0.02206	0.02367
Reject?	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	4736 784.85 2				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	Yes	Yes	Yes	Yes

Poisson [#5]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	4736 0.38872 0 4				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.01559	0.01777	0.01973	0.02206	0.02367
Reject?	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	4736 3585.2 5				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	Yes	Yes	Yes	Yes

Comparing models with different distribution assumptions

Despite both software packages showing a significant difference between tested distributions and observed data, the shelf capacity equation is run in Stata assuming a negative binomial and also assuming a gamma distribution.

The Stata code assuming the data follows a negative binomial distribution is:

nbreg capacityr i.casepack propCAT SKUsd, exposure(openwk) vce(cluster UPC).

The Stata output is as below.

```
Negative binomial regression                    Number of obs     =      4,736
                                                Wald chi2(6)      =    2130.44
Dispersion           = mean                     Prob > chi2       =     0.0000
Log pseudolikelihood = -16996.327               Pseudo R2         =     0.1179

                                 (Std. Err. adjusted for 148 clusters in UPC)
------------------------------------------------------------------------------
             |               Robust
   capacityr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    casepack |
           4 |   .1064685   .0574247     1.85   0.064    -.0060818    .2190188
           8 |   .3189543   .0451607     7.06   0.000      .230441    .4074677
          12 |   1.423839   .0706711    20.15   0.000     1.285326    1.562351
          15 |   1.079087   .0700718    15.40   0.000     .9417487    1.216425
             |
     propCAT |   .0009437   .0267986     0.04   0.972    -.0515806     .053468
       SKUsd |    .012269   .0041157     2.98   0.003     .0042024    .0203355
       _cons |  -7.004533   .0482834  -145.07   0.000    -7.099167     -6.9099
  ln(openwk) |          1  (exposure)
             |
    /lnalpha |  -2.217532   .0802041                     -2.374729   -2.060335
             |
       alpha |   .1088775   .0087324                      .0930397    .1274113
------------------------------------------------------------------------------
```

The Stata code for gamma and its output are below. The parameter estimates from each are the same in terms of sign and significance. However the AIC and BIC values differ. Each model's AIC and BIC statistics are obtained by using the command *estat ic* after each model is run.

glm capacityr i.casepack propCAT SKUsd, family(nbinomial 1) exposure(openwk) vce(cluster UPC)

```
Generalized linear models                          No. of obs      =      4,736
Optimization     : ML                              Residual df     =      4,729
                                                   Scale parameter =          1
Deviance         =  641.2805594                    (1/df) Deviance =    .135606
Pearson          =  665.5509721                    (1/df) Pearson  =   .1407382

Variance function: V(u) = u+(1)u^2                 [Neg. Binomial]
Link function    : g(u) = ln(u)                    [Log]

                                                   AIC             =   8.389199
Log pseudolikelihood = -19858.62219                BIC             =     -39380

                                 (Std. Err. adjusted for 148 clusters in UPC)
```

capacityr	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
casepack						
4	.1101054	.0584465	1.88	0.060	-.0044476	.2246584
8	.3173361	.0453689	6.99	0.000	.2284148	.4062574
12	1.415993	.0707779	20.01	0.000	1.27727	1.554715
15	1.075854	.0678581	15.85	0.000	.9428546	1.208853
propCAT	-.0003024	.0259444	-0.01	0.991	-.0511525	.0505478
SKUsd	.0119627	.0041067	2.91	0.004	.0039138	.0200116
_cons	-6.985213	.0488209	-143.08	0.000	-7.080901	-6.889526
ln(openwk)	1	(exposure)				

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	4,736	-19267.26	-16996.33	8	34008.65	34060.36

The above are the statistics for the negative binomial assumption while the bottom AIC and BIC values assume gamma distribution of shelf capacity values. Negative binomial has a better fit (34,008.65<39,731.24). This process is repeated for the Yogurt and Isotonics subsets with similar results.

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	4,736	.	-19858.62	7	39731.24	39776.49

SHO Frequency

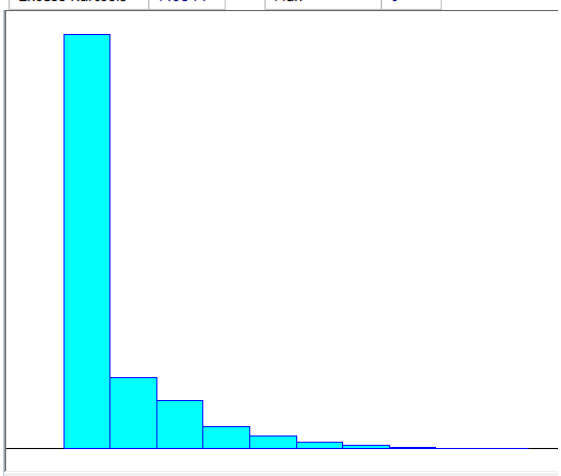
Descriptive Statistics and least square error approach

Descriptive Statistics

Statistic	Value
Sample Size	4736
Range	9
Mean	0.6231
Variance	1.5431
Std. Deviation	1.2422
Coef. of Variation	1.9936
Std. Error	0.39283
Skewness	2.5742
Excess Kurtosis	7.6844

Percentile	Value
Min	0
5%	0
10%	0
25% (Q1)	0
50% (Median)	0
75% (Q3)	1
90%	2
95%	3
Max	9

EasyFit shows half of the data set has a weekly item SHO frequency of 0, while the most frequently SHO item had a total of 9 SHOs per week. There is overdispersion where the variance is greater than the mean.



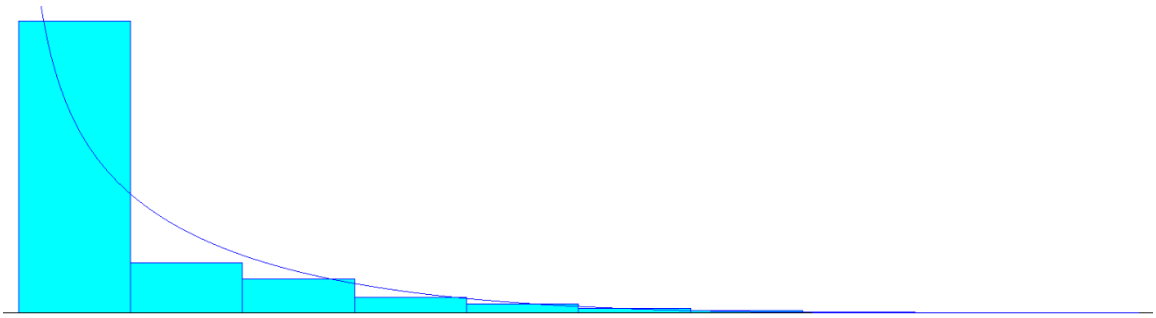
```
        Data Summary

Number of Data Points   = 4736
Min Data Value          = 0
Max Data Value          = 9
Sample Mean             = 0.623
Sample Std Dev          = 1.24

        Histogram Summary

Histogram Range         = -0.5 to 9.5
Number of Intervals     = 10
```

ARENA's histogram show the large number of zeroes in the data set. The descriptive stats match the summary from EasyFit.

```
Function       Sq Error
-----------------------
Beta           0.0103
Weibull        0.0186
Lognormal      0.0287
Exponential    0.0294
Gamma          0.0575
Poisson        0.0766
Erlang         0.0801
Normal         0.228
Triangular     0.36
Uniform        0.43
```

Comparison (from ARENA, above) of square errors between theoretical and empirical distributions show those with the least error are Beta and Weibull.

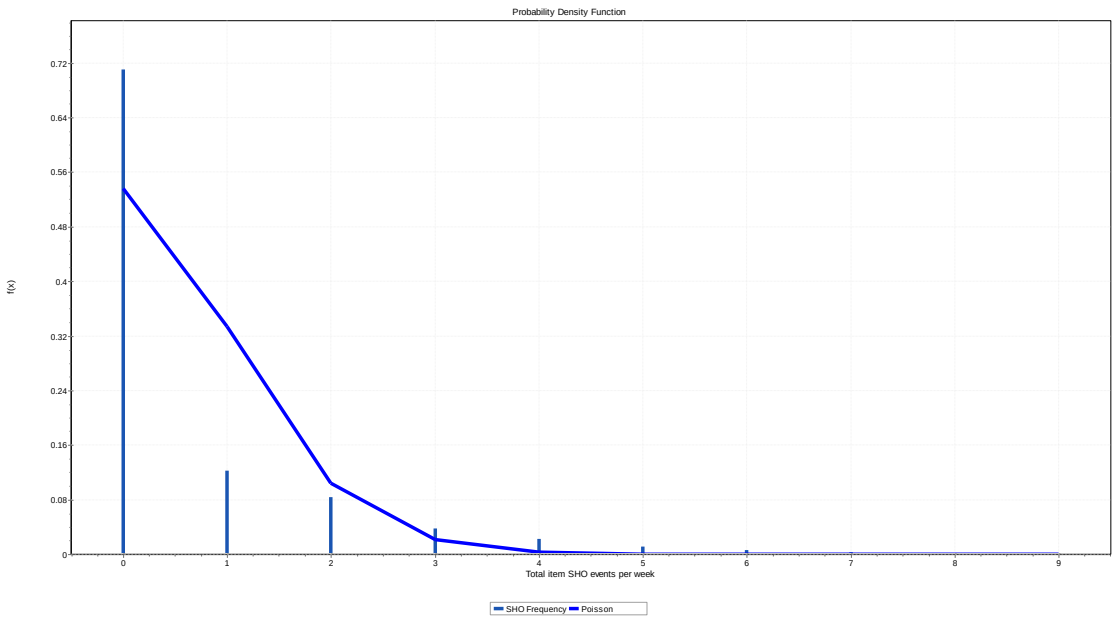
Graphical approach and Chi square

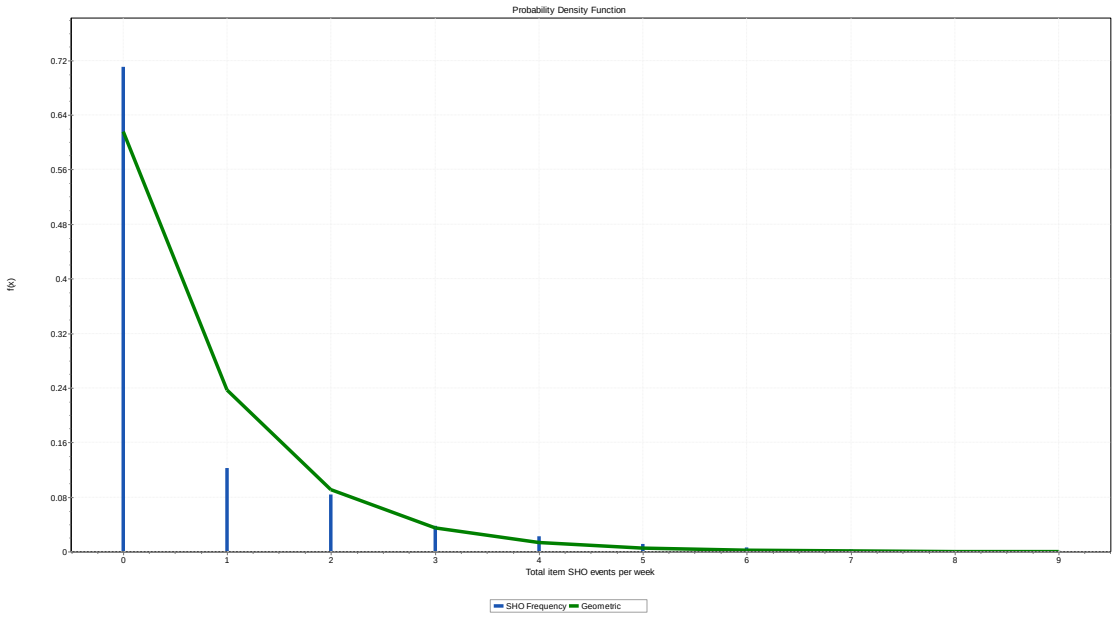


```
        Distribution Summary

Distribution:   Beta
Expression:     -0.5 + 10 * BETA(0.613, 4.85)
Square Error:   0.010275

Chi Square Test
  Number of intervals   = 7
  Degrees of freedom    = 4
  Test Statistic        = 243
  Corresponding p-value < 0.005
```

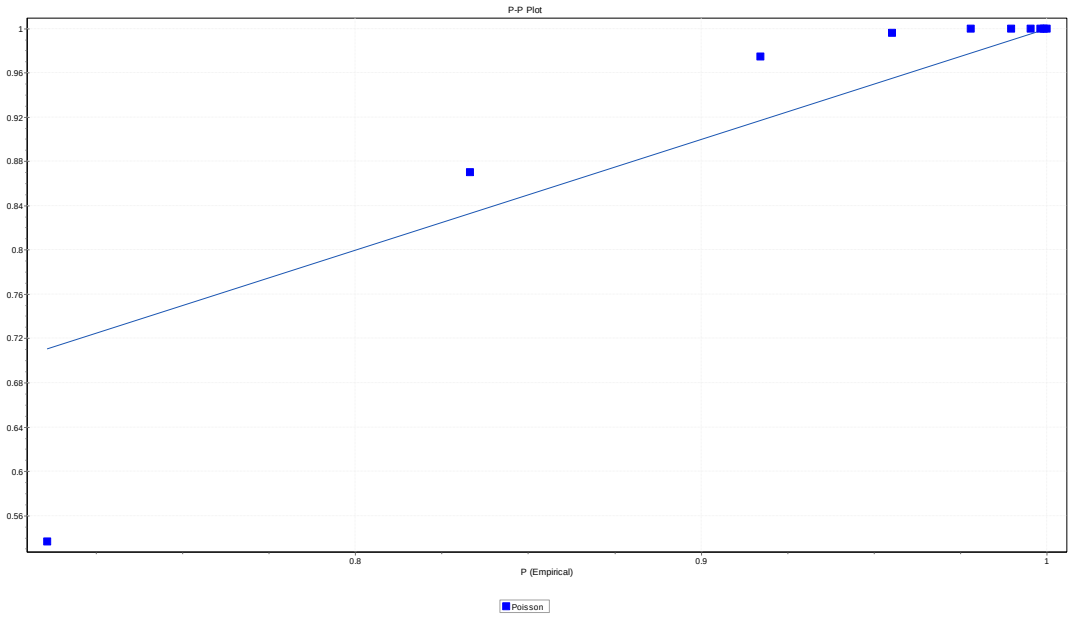
As the distribution with least error, the Chi square test still has a p value less than 0.05, showing there is observed data that does not fit the beta distribution.

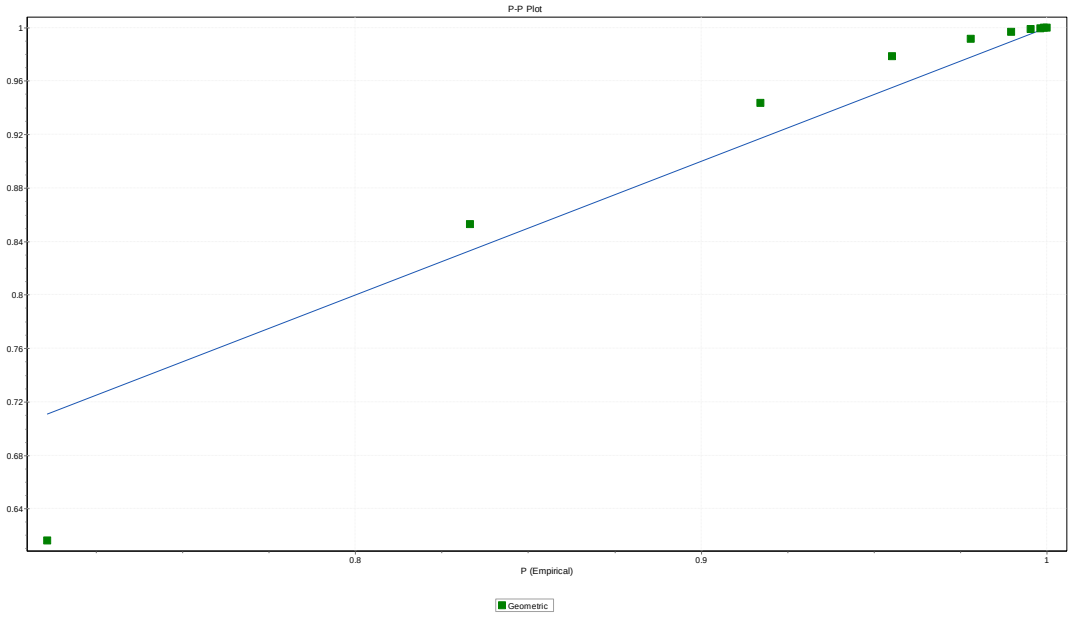
EasyFit produces graphs for Poisson and geometric distributions below.





The probability plots for Poisson and geometric distributions do not greatly differ in appears. Both show the greatest error at the lower bound (zero) of the data set.





Poisson [#3]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	4736 0.53628 0 2				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.01559	0.01777	0.01973	0.02206	0.02367
Reject?	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	4736 1345.6 1				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	Yes	Yes	Yes	Yes

Geometric [#2]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	4736 0.61611 0 3				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.01559	0.01777	0.01973	0.02206	0.02367
Reject?	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	4736 1698.2 2				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	Yes	Yes	Yes	Yes

The test statistics for Poisson and geometric are both greater than the critical values at all levels of alpha shown above. The observed data has evidence to reject the null hypothesis that it fits either distribution.

Comparing SHO frequency models assuming different distributions

No model is fitted with the beta, gamma or geometric distributions because of the characteristics of the data set. Beta is a continuous distribution and the discrete SHO frequency could be assumed to be continuous however it is time series data. Currently there is no preprogrammed module for time series data assuming a beta distribution. The SHO frequency equation includes lagged sales values as well as factor variables which currently cannot easily be modelled. Gamma distribution does not include zeroes, whereas SHO frequency data values have many. Finally, geometric distribution is just a

specific incidence of negative binomial and of the discrete compound Poisson contribution. The two models compared in Stata are thus the negative binomial and Poisson.

The Stata code for the comparison is:

```
nbreg eventrate L.soldwk capacityrhat i.week i.vendor, exposure(openwk) vce(cluster UPC)
predict eventratehat //save predicted values to fit into GSEM
estat ic
poisson eventrate L.soldwk capacityrhat i.week i.vendor, exposure(openwk) vce(cluster UPC)
estat ic
```

The resulting model output and AIC and BIC values are:

```
Negative binomial regression                    Number of obs     =      4,144
                                                Wald chi2(12)     =     141.84
Dispersion           = mean                     Prob > chi2       =     0.0000
Log pseudolikelihood = -4043.3271               Pseudo R2         =     0.0257

                                (Std. Err. adjusted for 148 clusters in UPC)
```

eventrate	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
soldwk						
L1.	.016065	.0025026	6.42	0.000	.01116	.0209699
capacityrhat	-.0106973	.0084269	-1.27	0.204	-.0272137	.0058191
week						
3	-.3996636	.0966715	-4.13	0.000	-.5891363	-.2101909
4	.0690681	.086138	0.80	0.423	-.0997592	.2378954
5	.0288051	.0795048	0.36	0.717	-.1270215	.1846316
6	-.1820674	.0815271	-2.23	0.026	-.3418577	-.0222771
7	-.6615396	.116547	-5.68	0.000	-.8899675	-.4331116
8	-.0503601	.1014603	-0.50	0.620	-.2492185	.1484984
vendor						
2	-.8390372	.5985869	-1.40	0.161	-2.012246	.3341715
3	.3965682	.1620589	2.45	0.014	.0789386	.7141978
4	.0363802	.149203	0.24	0.807	-.2560524	.3288127
5	.3686103	.2292505	1.61	0.108	-.0807124	.817933
_cons	-9.700703	.3179164	-30.51	0.000	-10.32381	-9.077599
ln(openwk)	1	(exposure)				
/lnalpha	.8041213	.0883536			.6309515	.9772912
alpha	2.234732	.1974466			1.879398	2.657248

The number of observations are lower than the total 4,736 data points in the sample because some items never SHO. They are still included into the study because the overall model looks at both the antecedents to and effects of SHO. Item demand and supply characteristics are linked to SHO frequency and duration which are then linked to item sales. Including items that never SHO give more conservative coefficient estimates of item-based parameters affecting SHO. For example, if an item with pack size 12 and shelf capacity 48 never SHO during the sample period and another item with the same pack size and shelf capacity SHO multiple times, not including the non-SHO-item would make the pack size and shelf capacity parameters of the SHO-item have a stronger link to SHO occurrence.

```
Poisson regression                              Number of obs     =      4,144
                                                Wald chi2(12)     =     172.00
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -4626.5754               Pseudo R2         =     0.0585

                                 (Std. Err. adjusted for 148 clusters in UPC)
```

eventrate	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
soldwk						
L1.	.0150015	.0021299	7.04	0.000	.010827	.0191761
capacityrhat	-.0089622	.0082751	-1.08	0.279	-.0251812	.0072568
week						
3	-.4255279	.0937153	-4.54	0.000	-.6092066	-.2418493
4	.0593008	.0834174	0.71	0.477	-.1041942	.2227959
5	.0348619	.0785992	0.44	0.657	-.1191898	.1889136
6	-.1791715	.0828778	-2.16	0.031	-.3416089	-.0167341
7	-.7020508	.1210799	-5.80	0.000	-.9393631	-.4647386
8	-.0859106	.1069065	-0.80	0.422	-.2954434	.1236222
vendor						
2	-.8463845	.6013157	-1.41	0.159	-2.024942	.3321727
3	.4168995	.1642224	2.54	0.011	.0950296	.7387695
4	-.0580573	.151823	-0.38	0.702	-.355625	.2395104
5	.3357915	.1893354	1.77	0.076	-.035299	.7068821
_cons	-9.643269	.2998113	-32.16	0.000	-10.23089	-9.055649
ln(openwk)	1	(exposure)				

Again, some items over time do not show any SHO events, so this non-varying dependent variable (SHO frequency) drops out since the model uses maximum likelihood estimators.

```
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	4,144	-4149.934	-4043.327	14	8114.654	8203.266

The AIC and BIC values for negative binomial (above) and Poisson (below) so that negative binomial distribution is a better fit for the data since 8114 is less than 9279.

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	4,144	-4914.258	-4626.575	13	9279.151	9361.433

SHO Duration

Unlike item shelf capacity and SHO frequency, SHO duration is a percentage between 0 and 100. It has a continuous but censored distribution of values that are calculated by taking the total SHO duration for the week and dividing it by that week's open store hours (which vary from store to store as well as over time for each store).

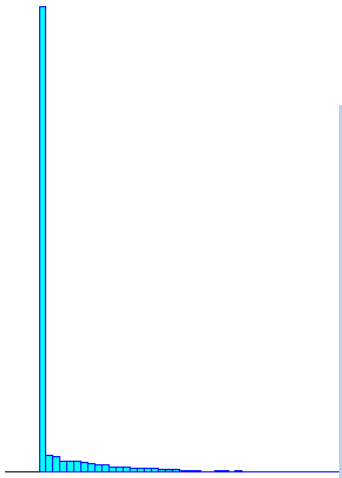
Descriptive statistics and least square error approach

Descriptive Statistics

Statistic	Value
Sample Size	4736
Range	80.941
Mean	2.6829
Variance	52.469
Std. Deviation	7.2435
Coef. of Variation	2.6999
Std. Error	0.10526
Skewness	4.729
Excess Kurtosis	30.789

Percentile	Value
Min	0
5%	0
10%	0
25% (Q1)	0
50% (Median)	0
75% (Q3)	0.74074
90%	9.3907
95%	15.856
Max	80.941

Since half of the data is zero for SHO frequency, it makes sense that half of the data for SHO duration is also zero. The maximum duration of SHO for an item is about 80% of the store's open hours.



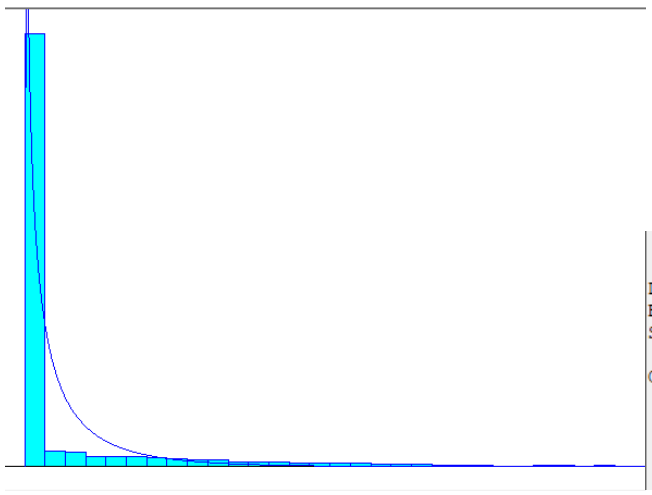
```
        Data Summary

Number of Data Points   = 4736
Min Data Value          = 0
Max Data Value          = 81
Sample Mean             = 2.68
Sample Std Dev          = 7.24

        Histogram Summary

Histogram Range         = -0.5 to 81.5
Number of Intervals     = 82
```

The histogram of SHO duration is similar to the histogram of SHO frequency. Since the 75th percentile of the SHO frequency is larger than the 75th percentile of SHO duration, it makes sense that the histogram cell with value of less than 1 is much taller (or higher) in duration than in frequency.

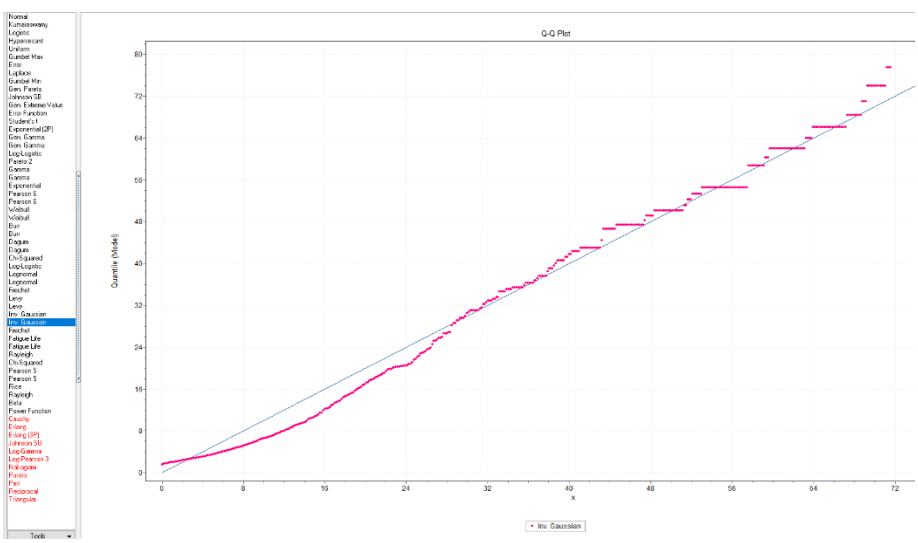
Function	Sq Error
Weibull	0.0586
Lognormal	0.0977
Beta	0.135
Gamma	0.175
Exponential	0.279
Erlang	0.279
Normal	0.487
Triangular	0.531
Uniform	0.541
Poisson	0.588

The 4 distributions with the least square error are checked using quantile-quantile plots below. The Chi square test below for Weibull shows a significant p value below 0.05, rejecting the null hypothesis that the empirical distribution fits the Weibull distribution.

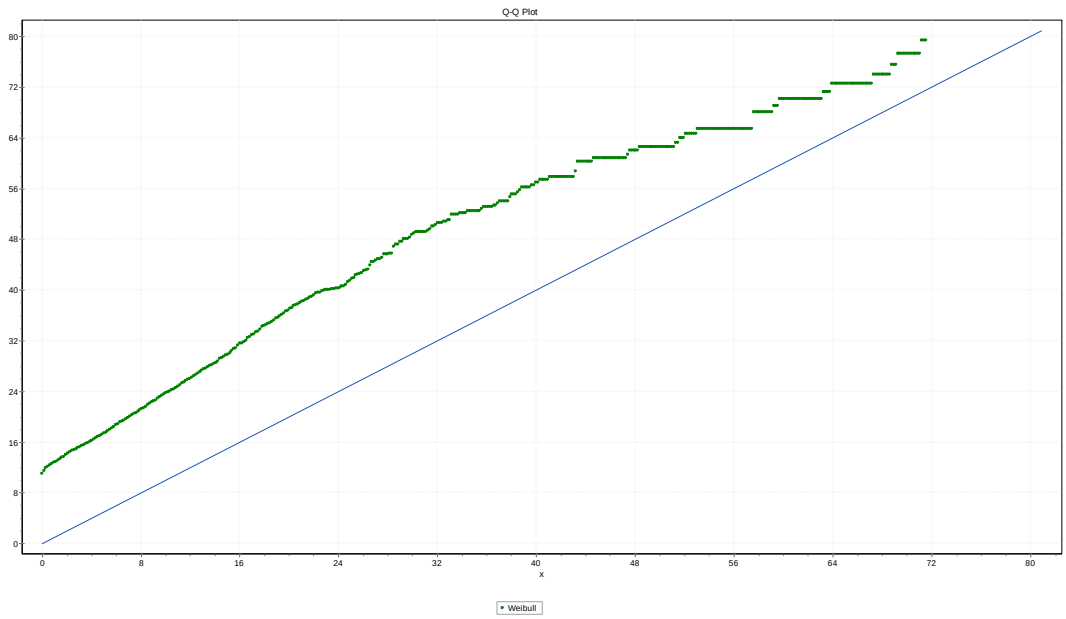
Graphical approach and fit tests



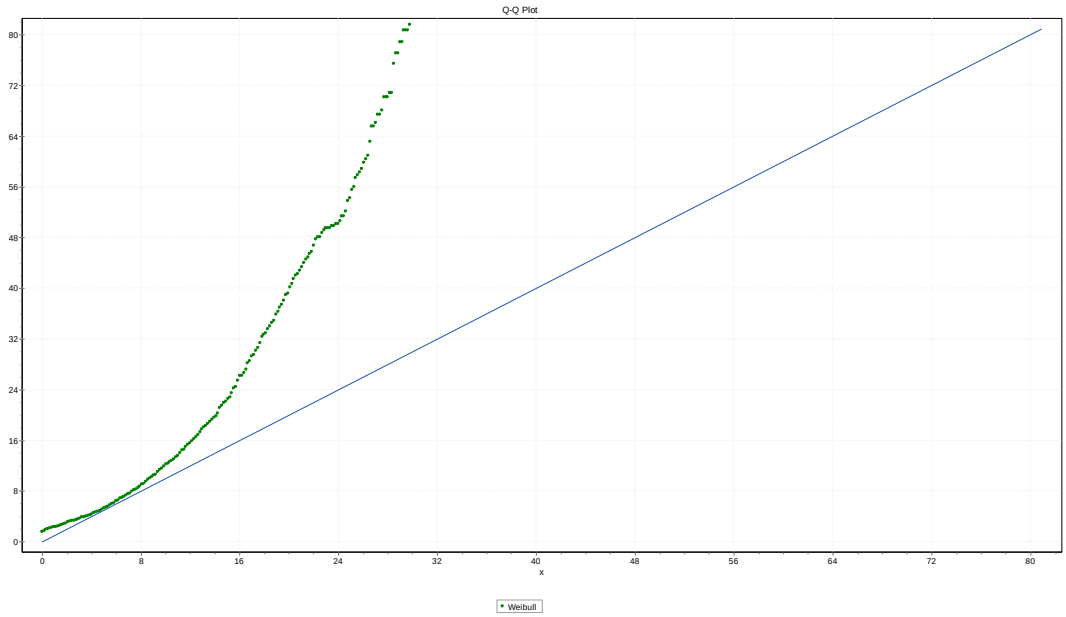
```
        Distribution Summary

Distribution:   Weibull
Expression:     -0.5 + WEIB(1.34, 0.664)
Square Error:   0.058639

Chi Square Test
  Number of intervals   = 17
  Degrees of freedom    = 14
  Test Statistic        = 4.11e+003
  Corresponding p-value < 0.005
```

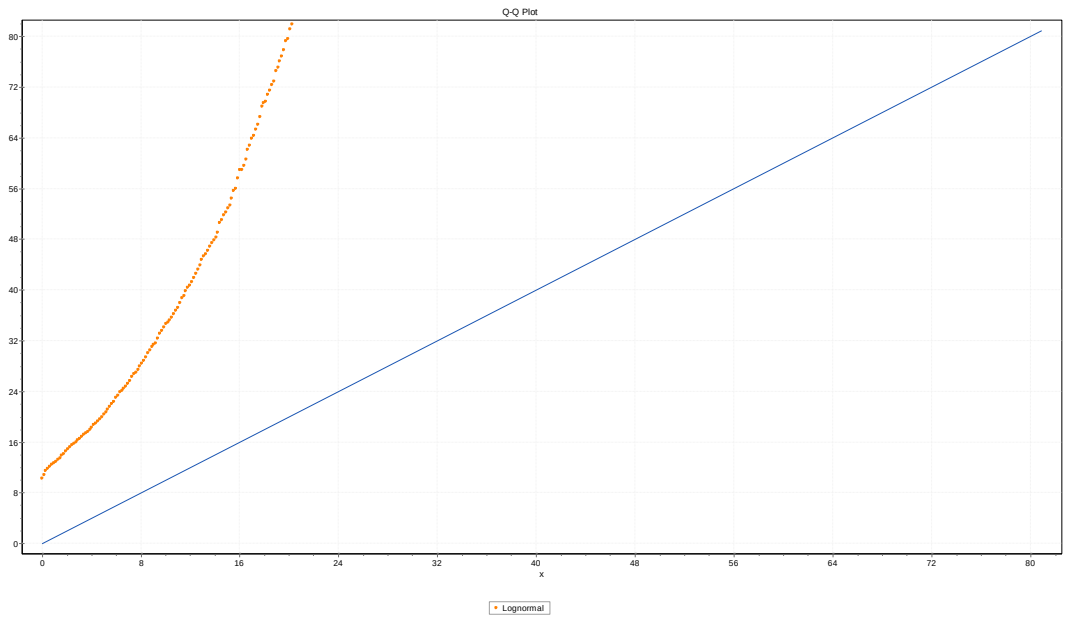


Since SHO duration is a continuous distribution, EasyFit has many distributions for which Q-Q plots are possible.

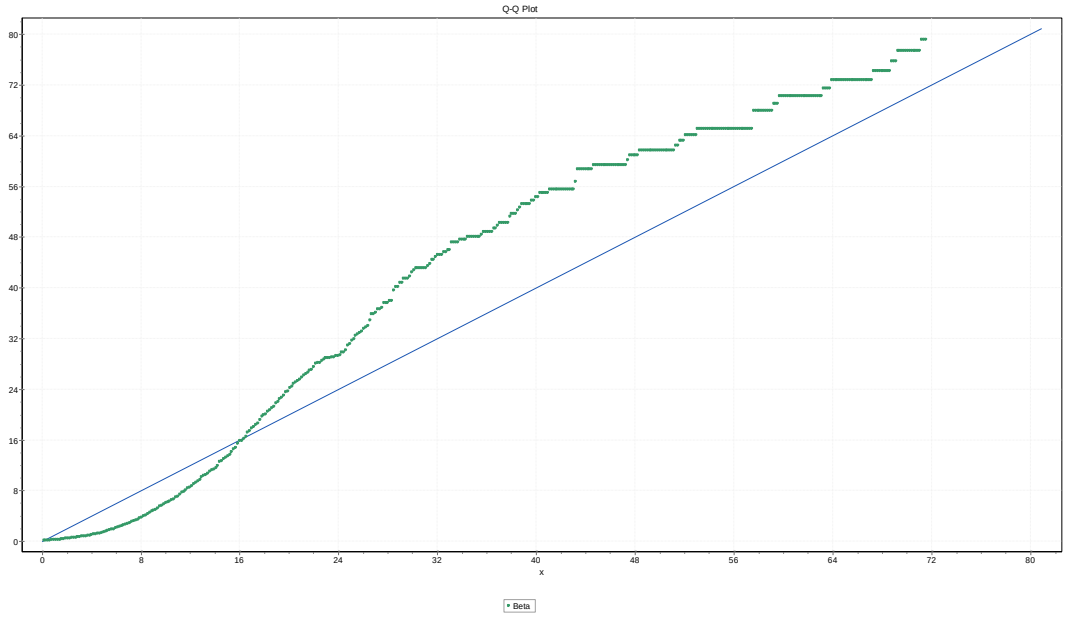


Weibull with parameters: α=0.84507 β=8.5208 (above) and different parameters (α=0.31478 β=0.77705) below. This is a good example of how Weibull is often a distribution which statistical software suggests instead of one based on economic theory. Different parameter values can give Weibull a completely different shape from appearing somewhat like the normal distribution to the exponential distribution.



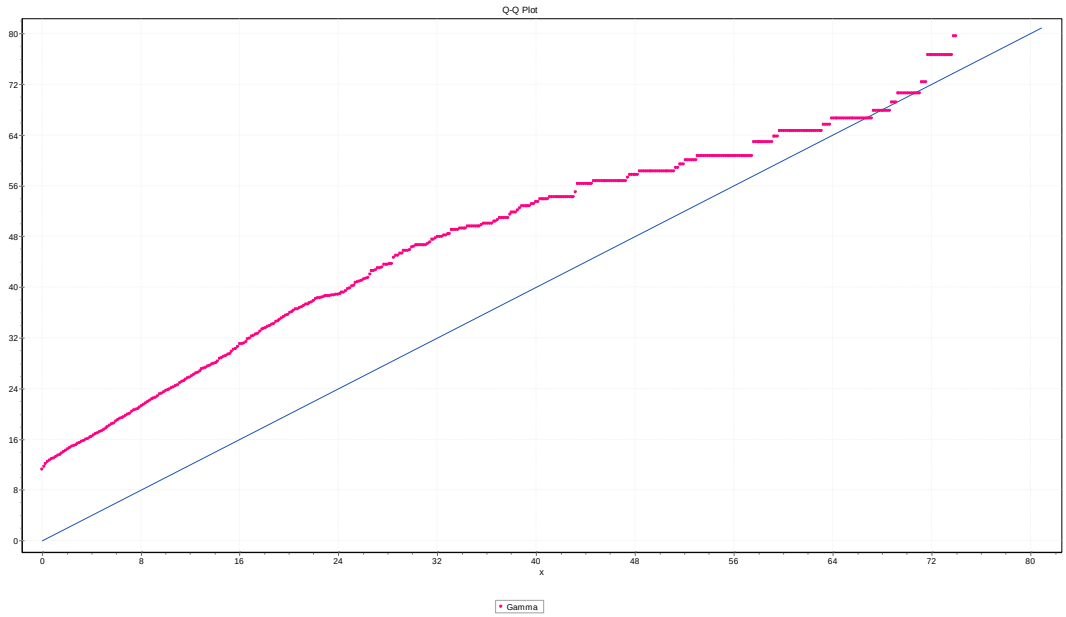


The lognormal Q-Q plot (above) shows how poor the observed data fit this distribution. The low least square error score for this distribution may be due to the large number of 0 values in the SHO duration data.

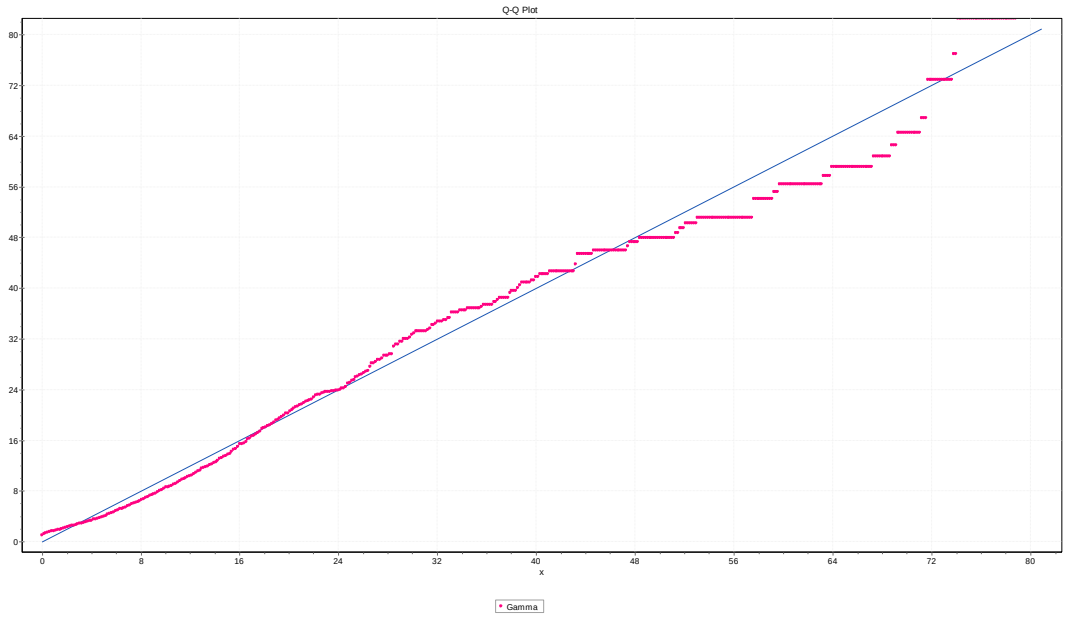


The Q-Q plot above shows a closer fit of the observed data to the beta distribution.

Depending on the parameter values of the distribution, the Q-Q plot shows the observed data to be distributed similar to the gamma distribution.

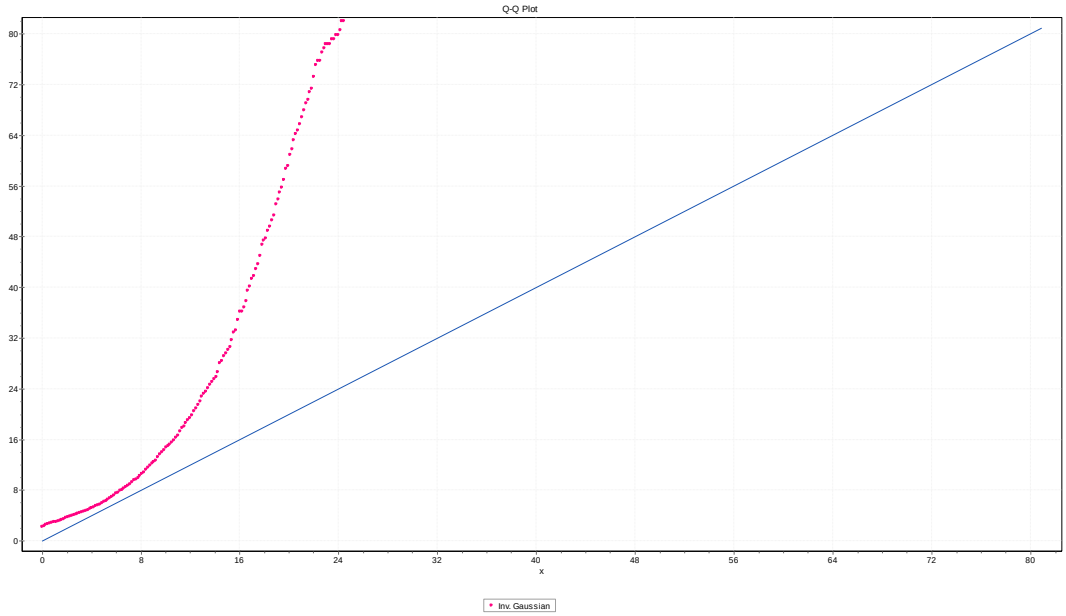


Above is gamma distribution with parameters: α=0.76338 β=12.158.

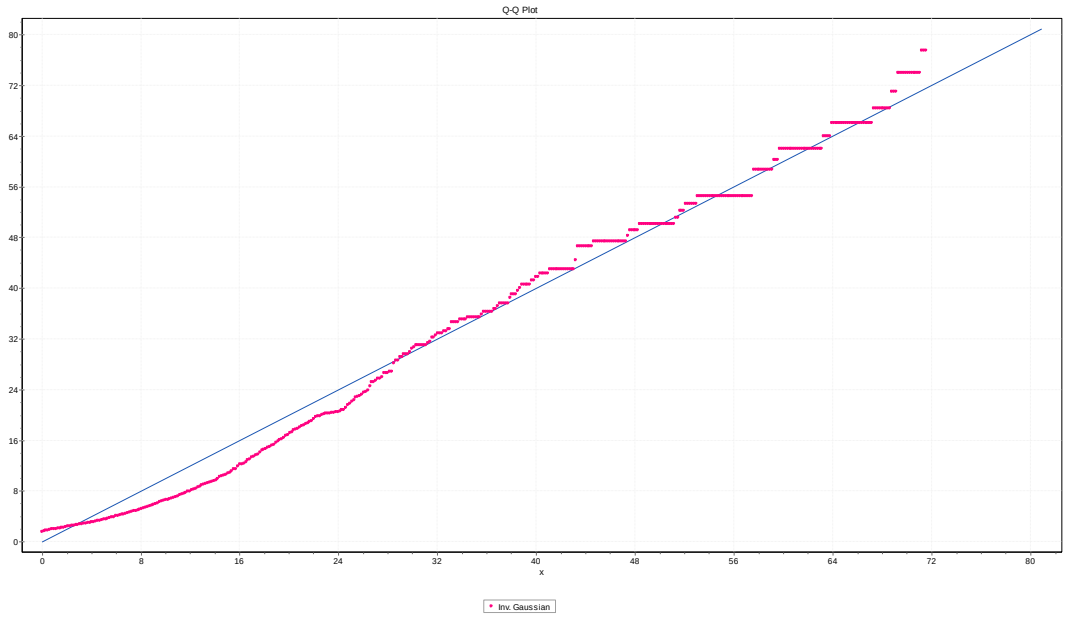


Above is gamma with parameters: α=0.13718 β=19.557.

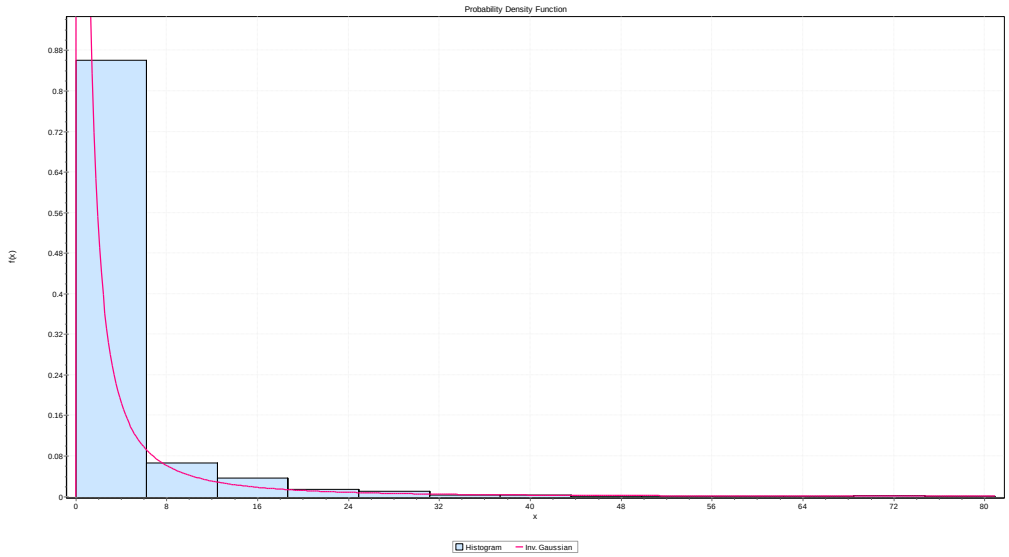
The inverse Gaussian also takes on very different shapes depending on its parameters. The Q-Q plots showing how well SHO duration fits this distribution also varies greatly by parameter values.



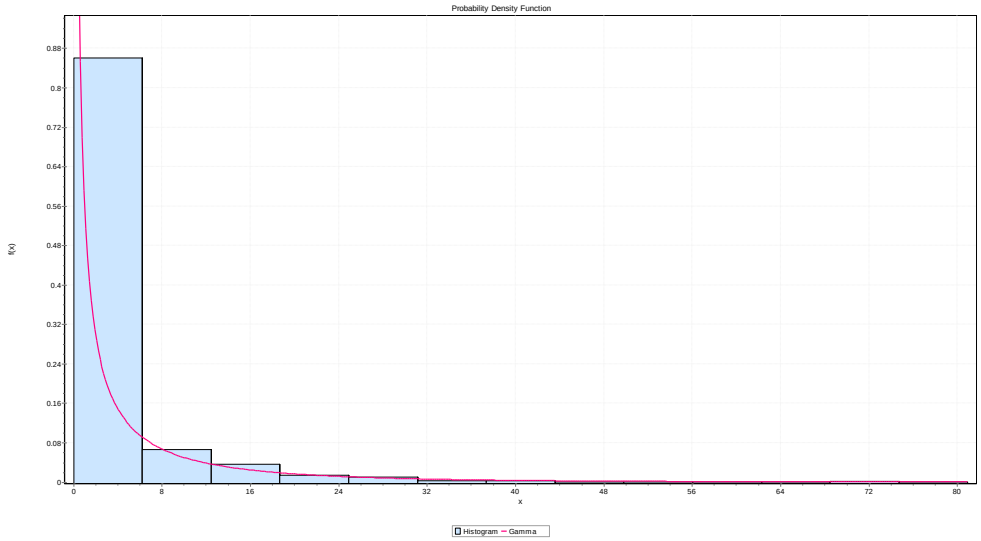
Q-Q plot above for inverse Gaussian with parameters: λ=0.3605 μ=9.2813. Inverse Gaussian Q-Q plot below with parameters: l=0.36805 m=2.6829.



Although the Q-Q plots suggest that the inverse Gaussian is the best distribution for SHO duration, it is more appropriate for non-zero valued distributions skewed to larger values. Histograms of the observed data and inverse Gaussian versus gamma distributions illustrate the issue. Even if the sorted observed data fall into quantiles that are more similar to inverse Gaussian rather than gamma, the superimposed histogram and curves show that the inverse Gaussian has a f(x) of zero when the SHO duration is zero.



In contrast the gamma PDF plot for SHO duration at zero approaches infinity.



Comparing SHO duration models with different distribution assumptions

```
. glm durrate pingsused i.vendor eventratehat eventratehat2 i.store, family(igaussian) link(power -2) vce(cluster UPC)

initial values not feasible
```

Despite being distributed similar to the inverse Gaussian, the values that SHO duration takes on is as important at its shape (distribution). The zero values in the data set make it impossible for the GLM model to converge (above), even after setting its initial values to parameter estimates from another regression (below).

```
Tobit regression                                Number of obs     =      4,144
                                                   Uncensored     =      1,171
Limits: lower = 0                                  Left-censored  =      2,973
        upper = 100                                Right-censored =          0

                                                F(  10,   4134)   =      18.93
                                                Prob > F          =     0.0000
Log pseudolikelihood = -6148.4892               Pseudo R2         =     0.0258

                                  (Std. Err. adjusted for 148 clusters in UPC)
------------------------------------------------------------------------------
             |               Robust
     durrate |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   pingsused |  -.5395732   1.062221    -0.51   0.612    -2.622097    1.542951
             |
      vendor |
          2  |  -1.839476   5.295488    -0.35   0.728    -12.22148    8.542529
          3  |   2.414156    1.69081     1.43   0.153    -.9007415    5.729053
          4  |   2.415349   1.605902     1.50   0.133    -.7330829    5.563781
          5  |   5.147854   1.769784     2.91   0.004     1.678125    8.617584
             |
 eventratehat |   17.86756   2.470401     7.23   0.000     13.02424    22.71087
eventratehat2 |  -2.736299   .5459243    -5.01   0.000    -3.806604   -1.665994
             |
       store |
      6515   |   .8984625   1.498049     0.60   0.549     -2.03852    3.835445
      6520   |  -.4612506   1.808467    -0.26   0.799     -4.00682    3.084318
      6539   |   2.337116    1.78863     1.31   0.191     -1.16956    5.843792
             |
       _cons |  -22.88885    2.64877    -8.64   0.000    -28.08186   -17.69583
-------------+----------------------------------------------------------------
var(e.durrate)|   257.3539     28.727                      206.7702    320.3122
------------------------------------------------------------------------------
.
end of do-file

. matrix b1=e(b)

. do "C:\Users\findu\AppData\Local\Temp\STD6144_000000.tmp"

. glm durrate pingsused i.vendor eventratehat eventratehat2 i.store, family(igaussian) link(power -2) vce(cluster UPC) from(b1, skip)

initial values not feasible
```

Set initial values of this GLM to the coefficient estimates from the Tobit regression.

The gamma distribution also has issues with data points at zero. Even if the model converges to a solution as below, the model cannot consistently fit the observed data.

```
Generalized linear models                          No. of obs      =      4,144
Optimization     : ML                              Residual df     =      4,133
                                                   Scale parameter =   1.24e+15
Deviance         =   6204303532                    (1/df) Deviance =    1501162
Pearson          =  5.11938e+18                    (1/df) Pearson  =   1.24e+15

Variance function: V(u) = u^2                      [Gamma]
Link function    : g(u) = 1/u                      [Reciprocal]

                                                   AIC             =    1497180
Log pseudolikelihood =  -3102156631                BIC             =   6.20e+09

                                 (Std. Err. adjusted for 148 clusters in UPC)
```

durrate	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
pingsused	.0086847	.0355264	0.24	0.807	-.0609458	.0783151
vendor						
2	.2887327	.6614743	0.44	0.662	-1.007733	1.585199
3	-.0853677	.0667466	-1.28	0.201	-.2161887	.0454533
4	-.0719322	.0733399	-0.98	0.327	-.2156757	.0718114
5	-.2517601	.0633091	-3.98	0.000	-.3758438	-.1276765
eventratehat	-.5575325	.101699	-5.48	0.000	-.756859	-.3582061
eventratehat2	.1112727	.0285143	3.90	0.000	.0553857	.1671597
store						
6515	-.0736072	.0767706	-0.96	0.338	-.2240748	.0768604
6520	-.0873853	.0888029	-0.98	0.325	-.2614359	.0866652
6539	-.1381242	.0781105	-1.77	0.077	-.291218	.0149697
_cons	.9404614	.0994924	9.45	0.000	.7454598	1.135463

```
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	4,144	.	-3.10e+09	11	6.20e+09	6.20e+09

The GLM model assuming a gamma distribution of SHO Duration converges to the above coefficient estimates but the fit is extremely poor with AIC and BIC values in the billions.

The AIC and BIC values of the tobit model are comparatively lower than when assuming a gamma distribution of SHO duration.

```
Tobit regression                                Number of obs     =      4,144
                                                   Uncensored     =      1,171
Limits: lower = 0                                  Left-censored  =      2,973
        upper = 100                                Right-censored =          0

                                                F(  10,    4134)   =      18.93
                                                Prob > F          =     0.0000
Log pseudolikelihood = -6148.4892               Pseudo R2         =     0.0258

                                  (Std. Err. adjusted for 148 clusters in UPC)
```

durrate	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
pingsused	-.5395732	1.062221	-0.51	0.612	-2.622097	1.542951
vendor						
2	-1.839476	5.295488	-0.35	0.728	-12.22148	8.542529
3	2.414156	1.69081	1.43	0.153	-.9007415	5.729053
4	2.415349	1.605902	1.50	0.133	-.7330829	5.563781
5	5.147854	1.769784	2.91	0.004	1.678125	8.617584
eventratehat	17.86756	2.470401	7.23	0.000	13.02424	22.71087
eventratehat2	-2.736299	.5459243	-5.01	0.000	-3.806604	-1.665994
store						
6515	.8984625	1.498049	0.60	0.549	-2.03852	3.835445
6520	-.4612506	1.808467	-0.26	0.799	-4.00682	3.084318
6539	2.337116	1.78863	1.31	0.191	-1.16956	5.843792
_cons	-22.88885	2.64877	-8.64	0.000	-28.08186	-17.69583
var(e.durrate)	257.3539	28.727			206.7702	320.3122

```
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	4,144	-6311.172	-6148.489	12	12320.98	12396.93

Weekly item sales (units)

The dependent variable in the main equation in Chapter 4 is the weekly unit sales of each item in each store.

Descriptive statistics and least square error approach

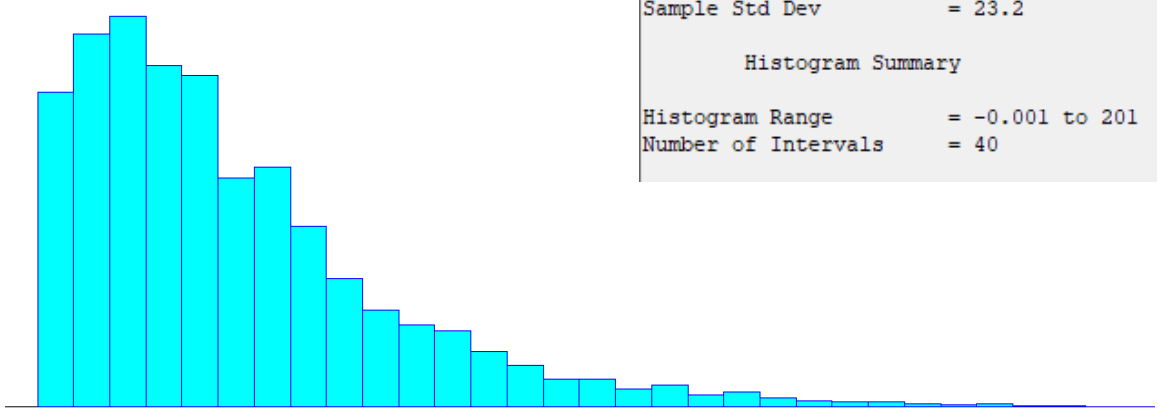
Descriptive Statistics

Statistic	Value
Sample Size	4736
Range	201
Mean	27.822
Variance	538.79
Std. Deviation	23.212
Coef. of Variation	0.83428
Std. Error	1.9411
Skewness	1.8084
Excess Kurtosis	5.0402

Percentile	Value
Min	0
5%	3
10%	5
25% (Q1)	11
50% (Median)	22
75% (Q3)	37
90%	58
95%	73
Max	201

Weekly item sales varies from 0 to 201 units.

```
        Data Summary

Number of Data Points   = 4736
Min Data Value          = 0
Max Data Value          = 201
Sample Mean             = 27.8
Sample Std Dev          = 23.2

        Histogram Summary

Histogram Range         = -0.001 to 201
Number of Intervals     = 40
```

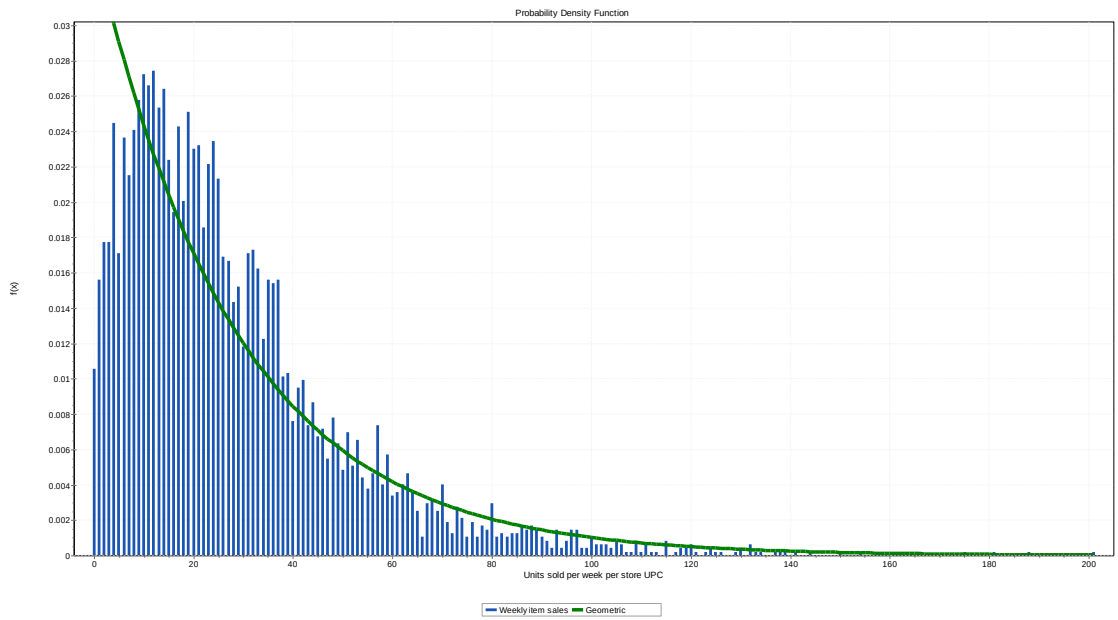


```
Function        Sq Error
------------------------
Gamma          0.00116
Beta           0.00133
Exponential    0.00612
Erlang         0.00612
Lognormal      0.0135
Normal         0.0152
Triangular     0.0426
Weibull        0.0583
Uniform        0.0616
```

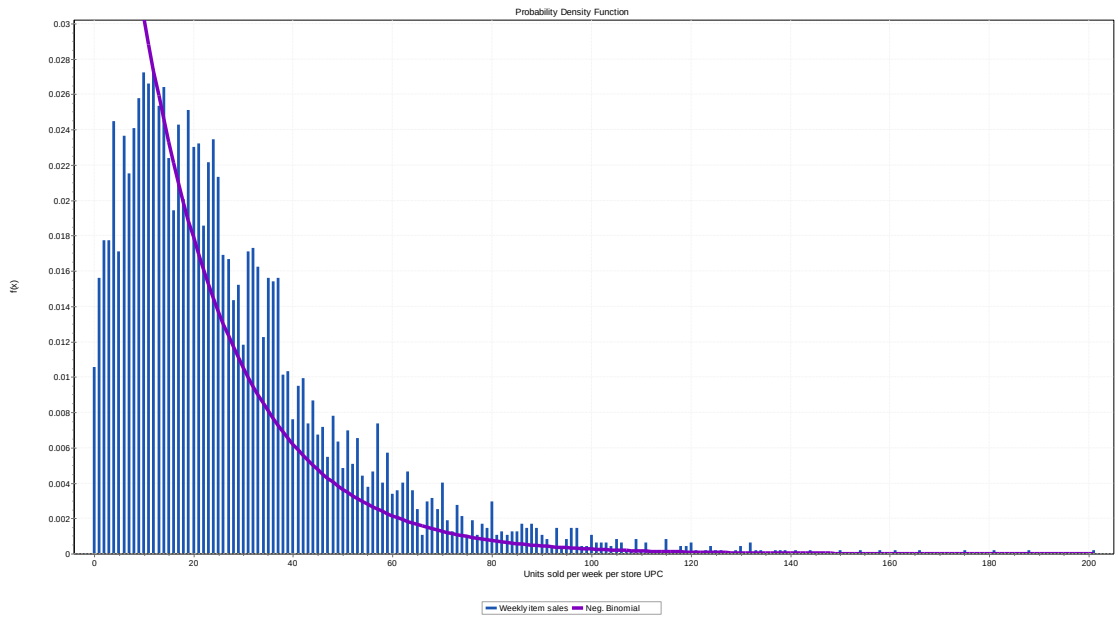
Once again ARENA does not allow for testing only discrete distributions. Given the number of observations and the range of values, it may be possible to consider the data as continuous instead of discrete. Gamma will be considered since it has the least square error among other distributions.

Graphical approach and fit tests

Below is the histogram of weekly sales with the geometric distribution and the KS test results. The test statistic is larger than any of the critical values from the geometric distribution. This suggests that there are observations (as seen by histogram cells that go up higher than the geometric curve) that shows the weekly sales data does not fit the geometric distribution.

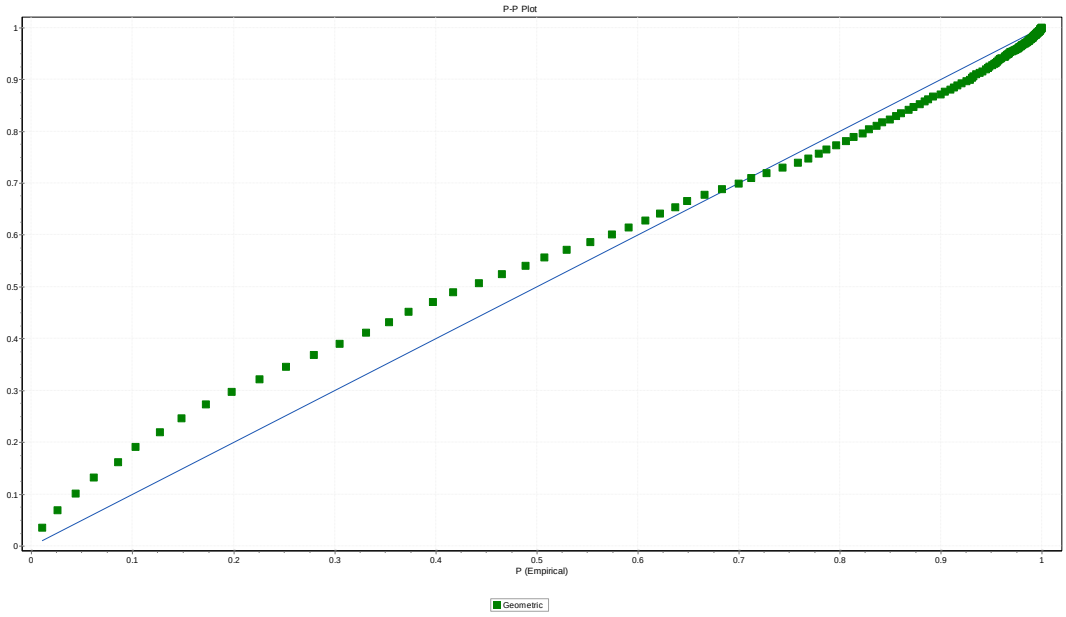


Geometric [#2]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	4736 0.12499 0 1				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.01559	0.01777	0.01973	0.02206	0.02367
Reject?	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	4736 123.56 1				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	Yes	Yes	Yes	Yes

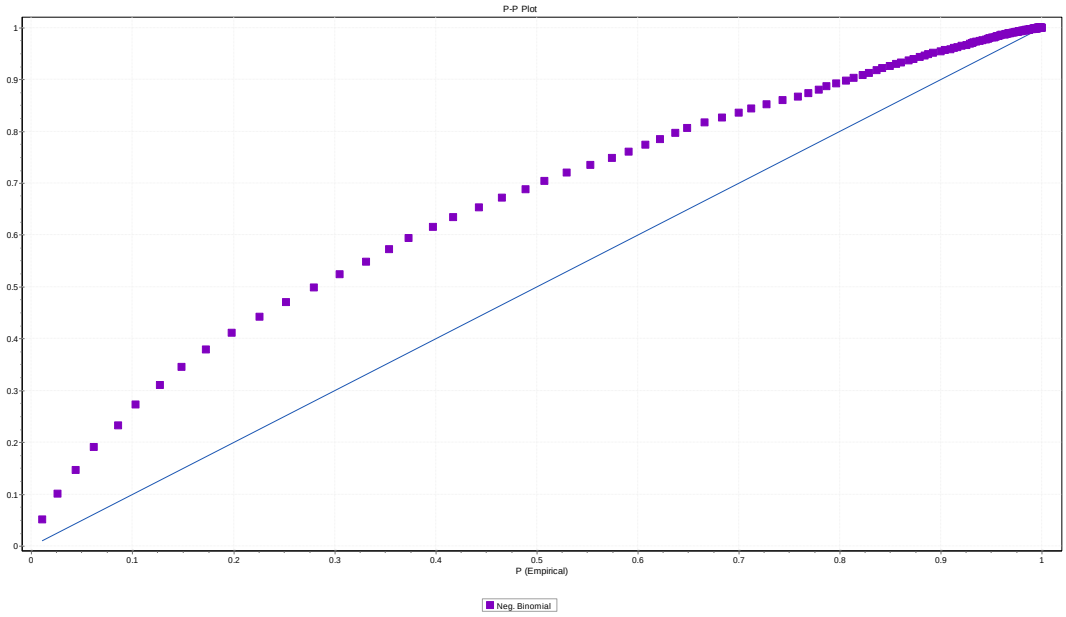


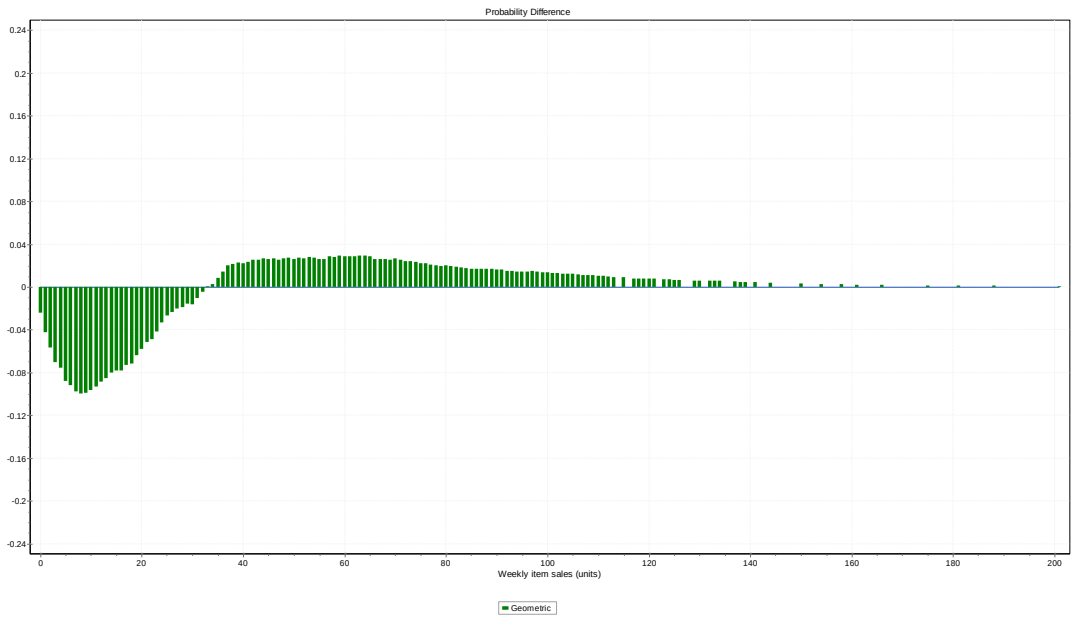
Neg. Binomial [#3]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	4736 0.24594 0 3				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.01559	0.01777	0.01973	0.02206	0.02367
Reject?	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	4736 723.57 2				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	Yes	Yes	Yes	Yes

The histogram of the observed data also has the negative binomial curve above. The KS test has an even larger test statistic so that the null hypothesis is rejected for negative binomial as well.

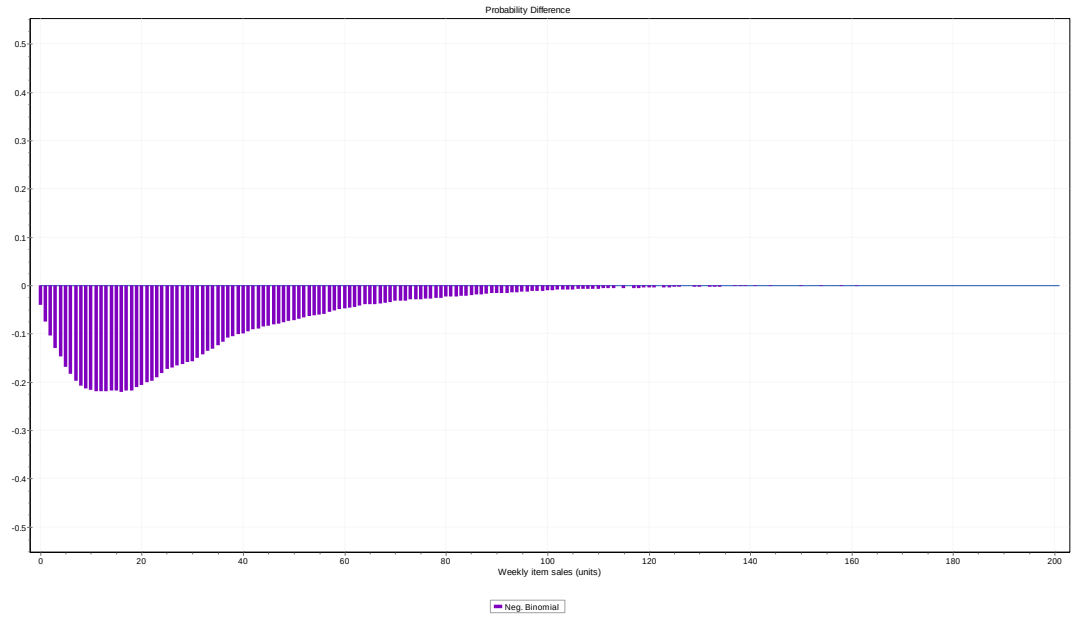


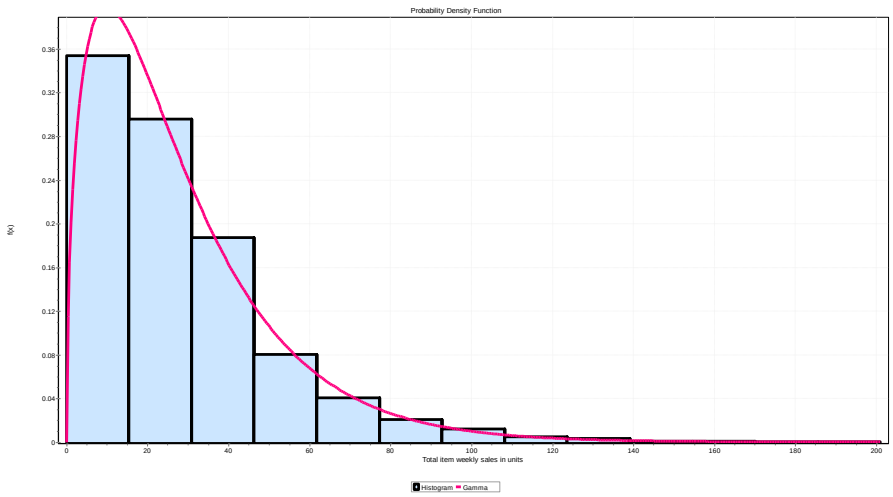
The Q-Q plots show that the observed data is distributed more like the geometric curve (above) rather than the negative binomial (below).

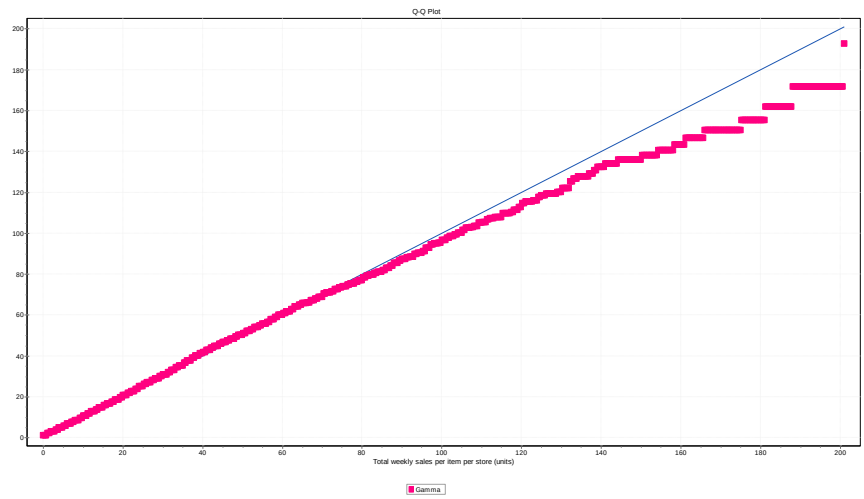




Probability difference plots are included for this variable to get a better view of where or how the theoretical distributions differ from the observed data in terms of the cumulative distribution function. While both distributions can more accurately model weekly sales values greater than 100, the geometric distribution is more accurate around the mean value of sales.







Gamma [#19]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	4736 0.0256 0.00396 2				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.01559	0.01777	0.01973	0.02206	0.02367
Reject?	Yes	Yes	Yes	Yes	Yes
Anderson-Darling					
Sample Size Statistic Rank	4736 99.137 8				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	Yes	Yes	Yes	Yes
Chi-Squared					
Deg. of freedom Statistic P-Value Rank	12 62.97 6.4550E-9 8				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	15.812	18.549	21.026	24.054	26.217
Reject?	Yes	Yes	Yes	Yes	Yes

The gamma distribution appears to have a closer fit to weekly sales data than the discrete distributions listed above. The p-value of 0.00396 of the KS test means that there is evidence in the observed data where it does not fit the gamma distribution at any level of alpha tested. However, the test statistic is 0.0256 which is relative close to the critical values compared to other KS test results.

Comparing Weekly sales models with different distribution assumptions

To save space, the 3 different distribution assumptions (negative binomial, geometric, gamma) are being listed below in a chart. The command to bring all of the estimation results together in a table is:

```
. estout NBRE GEOM GAMA, replace cells(b(star fmt(%5.3f)) se(par)) stats(N aic bic,
fmt(%9.3f %10.0g) labels(Observations AIC BIC)) legend label postfoot(Standard errors in
parentheses) numbers collabels(none) wrap varlabels(_cons Constant) modelwidth(10)
varwidth(26)
```

	(1) Neg.Bi.	(2) Geometric	(3) Gamma
L.Fitted Duration	-0.836***	-0.828***	-0.164***
	(0.061)	(0.064)	(0.001)
L.Fitted Frequency	16.583***	16.451***	2.841***
	(1.116)	(1.171)	(0.018)
L.Fitted Duration #	-0.031***	-0.032***	0.003***
L.Fitted Frequency	(0.005)	(0.006)	(0.000)
L.Fitted frequency	-2.521***	-2.499***	-0.435***
squared	(0.173)	(0.182)	(0.003)
Fitted item	0.032***	0.033***	-0.006***
shelf capacity	(0.006)	(0.006)	(0.000)
L.percdisc	-0.021**	-0.024**	0.001
	(0.007)	(0.007)	(0.001)
vendor=1	0.000	0.000	0.000
	(.)	(.)	(.)
vendor=2	-1.246***	-1.240***	-0.396***
	(0.157)	(0.165)	(0.002)
vendor=3	1.521***	1.503***	0.391***
	(0.162)	(0.170)	(0.003)
vendor=4	1.836***	1.818***	0.391***
	(0.156)	(0.164)	(0.003)
vendor=5	3.892***	3.870***	0.787***
	(0.370)	(0.390)	(0.006)
shelfno=1	0.000	0.000	0.000
	(.)	(.)	(.)
shelfno=2	-0.007	0.001	0.049***
	(0.178)	(0.179)	(0.004)
shelfno=3	0.568***	0.556**	0.020***
	(0.172)	(0.181)	(0.002)
shelfno=4	0.571***	0.557**	0.026***

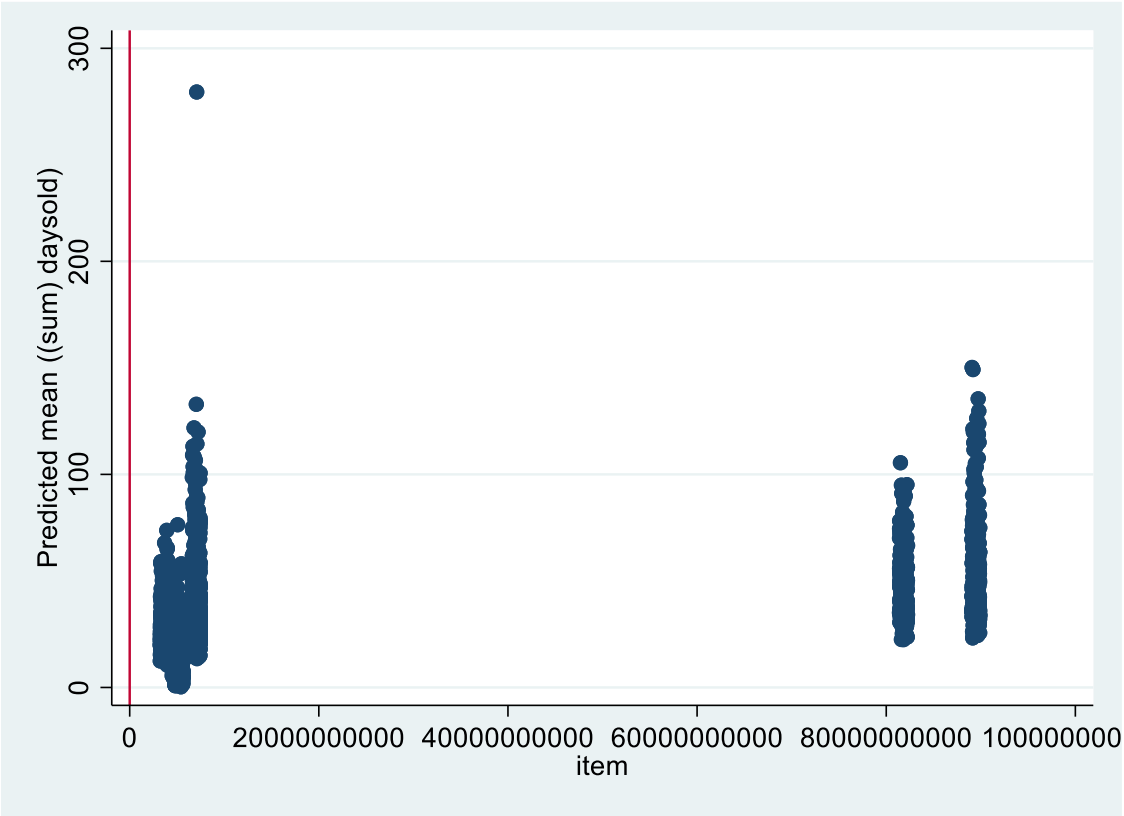
	(0.172)	(0.181)	(0.002)
shelfno=5	0.540***	0.534**	0.018***
	(0.158)	(0.164)	(0.001)
shelfno=6	0.491**	0.481**	0.024***
	(0.166)	(0.174)	(0.002)
shelfno=7	0.398	0.373	0.022***
	(0.229)	(0.247)	(0.002)
store=6504	0.000	0.000	0.000
	(.)	(.)	(.)
store=6515	0.455***	0.454***	-0.040***
	(0.045)	(0.047)	(0.001)
store=6520	-0.445***	-0.437***	-0.182***
	(0.039)	(0.042)	(0.001)
store=6539	1.286***	1.268***	0.063***
	(0.136)	(0.143)	(0.003)
Constant	-26.960***	-26.809***	-12.269***
	(1.507)	(1.582)	(0.026)
/			
lnalpha	-1.611***		
	(0.078)		
Observations	3552.000	3552.000	3552.000
AIC	26816.994	29262.623	4.580e+12
BIC	26946.675	29386.128	4.580e+12

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

The L. prefix in front of the variables denote lagged values for this lagged sales equation. The predicted values of this model are later fitted into the GSEM.

Negative binomial once again is the assumed distribution which neither overfits or underfits the observed data (the AIC and BIC values are the lowest).

XIV. Visual heteroscedasticity check (Ch. 4)



There is a single outlier among the predicted sales values across items, above. 278 units are predicted for sales of a French vanilla yogurt in a single store during a single week. The item has a mean weekly sales of 159 and standard deviation of 29 units. The actual sales for this item during this week in this store is 188 units, with the highest sales for this item occurring at a different week of 201 units. 201 units is the maximum units sold per week of any item.

XV. Postestimation comparison of sales by SHO attributes

The Stata code for the postestimation tests in Ch. 4.

```
******For post-estimation high-low analysis
gen matrix=.
gen matrixy=.
gen matrixi=.


***high-low cut-off values for duration and frequency, including subsets

egen durcut=mean(durrate) if durrate!=0
gen durcutc=durcut
egen frqcut=mean(eventrate) if eventrate!=0
gen frqcutc=frqcut
egen durcuty=mean(durrate) if (durrate!=0   & cat=="Yogurt")
gen durcutyc=durcuty
egen frqcuty=mean(eventrate) if (eventrate!=0  & cat=="Yogurt")
gen frqcutyc=frqcuty
egen durcuti=mean(durrate) if (durrate!=0  & cat=="Isotonics")
gen durcutic=durcuti
egen frqcuti=mean(eventrate) if (eventrate!=0 & cat=="Isotonics")
gen frqcutic=frqcuti

**low duration low frequency
replace matrix=1 if durrate<=durcutc & eventrate<=frqcutc & durrate!=0 & eventrate!=0
**low duration high frequency
replace matrix=2 if eventrate>frqcutc & durrate<=durcutc & durrate!=0
**high duration low frequency
replace matrix=3 if eventrate<=frqcutc & durrate>durcutc & eventrate!=0
**high duration high frequency
replace matrix=4 if eventrate>frqcutc & durrate>durcutc
**no stockout
replace matrix=0 if durrate==0|eventrate==0

**same as above for product category subsets
**low duration low frequency
replace matrixy=1 if durrate<=durcutyc & eventrate<=frqcutyc & durrate!=0 & eventrate!=0 &
cat=="Yogurt"
**low duration high frequency
replace matrixy=2 if eventrate>frqcutyc & durrate<=durcutyc & durrate!=0 & cat=="Yogurt"
**high duration low frequency
replace matrixy=3 if eventrate<=frqcutyc & durrate>durcutyc & eventrate!=0 & cat=="Yogurt"
```

```
**high duration high frequency
replace matrixy=4 if eventrate>frqcutyc & durrate>durcutyc & cat=="Yogurt"
**no stockout
replace matrixy=0 if (durrate==0|eventrate==0) & cat=="Yogurt"
**low duration low frequency
replace matrixi=1 if durrate<=durcutic & eventrate<=frqcutic & durrate!=0 & eventrate!=0 &
cat=="Isotonics"
**low duration high frequency
replace matrixi=2 if eventrate>frqcutic & durrate<=durcutic & durrate!=0 & cat=="Isotonics"
**high duration low frequency
replace matrixi=3 if eventrate<=frqcutic & durrate>durcutic & eventrate!=0 & cat=="Isotonics"
**high duration high frequency
replace matrixi=4 if eventrate>frqcutic & durrate>durcutic & cat=="Isotonics"
**no stockout
replace matrixi=0 if (durrate==0|eventrate==0) & cat=="Isotonics"

*********figure out high-low frequency and duration matrix
separate saleshat, by(matrix)
separate saleshaty, by(matrixy)
separate saleshati, by(matrixi)
separate thrslshat, by(matrix)
separate thrslshaty, by(matrixy)
separate thrslshati, by(matrixi)

pwmean saleshat, over(matrix) effects sort mcompare(bonferroni) groups
pwmean saleshaty, over(matrixy) effects sort mcompare(bonferroni) groups
pwmean saleshati, over(matrixi) effects sort mcompare(bonferroni) groups
pwmean thrslshat, over(matrix) effects sort mcompare(bonferroni) groups
pwmean thrslshaty, over(matrixy) effects sort mcompare(bonferroni) groups
pwmean thrslshati, over(matrixi) effects sort mcompare(bonferroni) groups
```

After running the above Stata code, the below output is obtained.

```
. pwmean saleshat, over(matrix) effects sort mcompare(bonferroni) groups

Pairwise comparisons of means with equal variances

over         : matrix
```

	Number of Comparisons
matrix	10

saleshat	Mean	Std. Err.	Bonferroni Groups
matrix			
3	23.23357	1.319365	A
0	26.11192	.3246413	A
1	28.83846	.6859053	
4	33.64001	1.177462	
2	47.91574	1.779032	

Note: Means sharing a letter in the group label are not significantly different at the 5% level.

	Number of Comparisons
matrix	10

saleshat	Contrast	Std. Err.	Bonferroni t	Bonferroni P>|t|	Bonferroni [95% Conf.	Interval]
matrix						
3 vs 2	-24.68217	2.214876	-11.14	0.000	-30.90274	-18.4616
4 vs 2	-14.27573	2.133394	-6.69	0.000	-20.26746	-8.284011
3 vs 1	-5.604891	1.487007	-3.77	0.002	-9.78121	-1.428572
3 vs 0	-2.878349	1.358719	-2.12	0.342	-6.694365	.9376664
1 vs 0	2.726542	.7588531	3.59	0.003	.5952725	4.857811
4 vs 1	4.801549	1.362675	3.52	0.004	.9744211	8.628677
4 vs 0	7.528091	1.221397	6.16	0.000	4.097749	10.95843
4 vs 3	10.40644	1.768373	5.88	0.000	5.439894	15.37299
2 vs 1	19.07728	1.906678	10.01	0.000	13.7223	24.43226
2 vs 0	21.80382	1.80841	12.06	0.000	16.72483	26.88282

The comparisons based on predictions from the overall GSEM model

```
. pwmean saleshaty, over(matrixy) effects sort mcompare(bonferroni) groups

Pairwise comparisons of means with equal variances

over          : matrixy
```

	Number of Comparisons
matrixy	10

saleshaty	Mean	Std. Err.	Bonferroni Groups
matrixy			
0	33.83535	.4665983	A
3	34.7287	1.917166	AB
1	38.85148	.9958957	B
4	48.39118	1.63061	C
2	52.51496	2.401613	C

Note: Means sharing a letter in the group label are not significantly different at the 5% level.

	Number of Comparisons
matrixy	10

saleshaty	Contrast	Std. Err.	Bonferroni t	Bonferroni P>|t|	Bonferroni [95% Conf.	Bonferroni Interval]
matrixy						
3 vs 2	-17.78627	3.072991	-5.79	0.000	-26.42023	-9.152303
4 vs 2	-4.123779	2.902867	-1.42	1.000	-12.27976	4.0322
3 vs 1	-4.122781	2.160401	-1.91	0.565	-10.19271	1.947145
3 vs 0	.8933425	1.973129	0.45	1.000	-4.650419	6.437104
1 vs 0	5.016123	1.099783	4.56	0.000	1.926142	8.106105
4 vs 1	9.539707	1.91068	4.99	0.000	4.171404	14.90801
4 vs 3	13.66249	2.516826	5.43	0.000	6.591139	20.73384
2 vs 1	13.66349	2.599914	5.26	0.000	6.358691	20.96828
4 vs 0	14.55583	1.696055	8.58	0.000	9.790543	19.32112
2 vs 0	18.67961	2.44652	7.64	0.000	11.8058	25.55342

The comparisons of predictions from the yogurt GSEM model

```
. pwmean saleshati, over(matrixi) effects sort mcompare(bonferroni) groups

Pairwise comparisons of means with equal variances

over          : matrixi
```

	Number of Comparisons
matrixi	10

saleshati	Mean	Std. Err.	Bonferroni Groups
matrixi			
3	14.15367	3.065322	A
0	15.57309	.3466629	A
4	16.36423	.9945216	A
1	16.91794	.8522183	A
2	18.03896	1.169068	A

Note: Means sharing a letter in the group label are not significantly different at the 5% level.

	Number of Comparisons
matrixi	10

saleshati	Contrast	Std. Err.	Bonferroni t	Bonferroni P>|t|	Bonferroni [95% Conf.	Bonferroni Interval]
matrixi						
3 vs 2	-3.885286	3.280688	-1.18	1.000	-13.10611	5.33554
3 vs 1	-2.764266	3.181583	-0.87	1.000	-11.70654	6.178012
4 vs 2	-1.674729	1.534859	-1.09	1.000	-5.988663	2.639204
3 vs 0	-1.41942	3.084862	-0.46	1.000	-10.08985	7.251009
4 vs 1	-.553709	1.309713	-0.42	1.000	-4.234839	3.127421
4 vs 0	.7911367	1.053209	0.75	1.000	-2.169051	3.751325
2 vs 1	1.12102	1.446719	0.77	1.000	-2.945183	5.187223
1 vs 0	1.344846	.9200278	1.46	1.000	-1.241019	3.930711
4 vs 3	2.210557	3.222618	0.69	1.000	-6.847056	11.26817
2 vs 0	2.465866	1.219383	2.02	0.433	-.9613777	5.89311

The comparison of predictions from the isotonics GSEM model

```
. pwmean thrslshat, over(matrix) effects sort mcompare(bonferroni) groups

Pairwise comparisons of means with equal variances

over          : matrix
```

	Number of Comparisons
matrix	10

thrslshat	Mean	Std. Err.	Bonferroni Groups
matrix			
3	.8705333	.0405474	A
0	.9849455	.009977	AB
1	1.011249	.0210796	B
4	1.084813	.0361864	B
2	1.418198	.0546741	

Note: Means sharing a letter in the group label are not significantly different at the 5% level.

	Number of Comparisons
matrix	10

thrslshat	Contrast	Std. Err.	Bonferroni t	Bonferroni P>|t|	Bonferroni [95% Conf.	Bonferroni Interval]
matrix						
3 vs 2	-.547665	.0680687	-8.05	0.000	-.7388386	-.3564914
4 vs 2	-.3333851	.0655645	-5.08	0.000	-.5175256	-.1492445
3 vs 1	-.1407155	.0456994	-3.08	0.021	-.2690641	-.0123668
3 vs 0	-.1144122	.0417568	-2.74	0.062	-.2316879	.0028634
1 vs 0	.0263032	.0233214	1.13	1.000	-.039196	.0918024
4 vs 1	.0735645	.0418784	1.76	0.791	-.0440527	.1911817
4 vs 0	.0998677	.0375366	2.66	0.078	-.0055552	.2052907
4 vs 3	.21428	.0543465	3.94	0.001	.0616456	.3669143
2 vs 1	.4069496	.058597	6.94	0.000	.2423777	.5715214
2 vs 0	.4332528	.0555769	7.80	0.000	.2771628	.5893428

Comparison of predictions from the overall 3SLS model

```
. pwmean thrslshaty, over(matrixy) effects sort mcompare(bonferroni) groups

Pairwise comparisons of means with equal variances

over          : matrixy
```

	Number of Comparisons
matrixy	10

thrslshaty	Mean	Std. Err.	Bonferroni Groups
matrixy			
3	1.267303	.0511376	A
0	1.269404	.0124458	A
1	1.32536	.0265641	A
4	1.489721	.0434942	
2	1.723533	.0640596	

Note: Means sharing a letter in the group label are not significantly different at the 5% level.

	Number of Comparisons
matrixy	10

thrslshaty	Contrast	Std. Err.	Bonferroni t	Bonferroni P>|t|	Bonferroni [95% Conf.	Interval]
matrixy						
3 vs 2	-.4562298	.0819676	-5.57	0.000	-.6865284	-.2259313
4 vs 2	-.2338121	.0774298	-3.02	0.026	-.4513612	-.0162631
3 vs 1	-.0580569	.0576256	-1.01	1.000	-.2199635	.1038496
3 vs 0	-.0021011	.0526304	-0.04	1.000	-.149973	.1457707
1 vs 0	.0559558	.0293351	1.91	0.566	-.026465	.1383767
4 vs 1	.1643607	.0509646	3.22	0.013	.021169	.3075525
4 vs 0	.2203166	.0452398	4.87	0.000	.0932094	.3474237
4 vs 3	.2224177	.0671327	3.31	0.009	.0337996	.4110357
2 vs 1	.3981729	.069349	5.74	0.000	.203328	.5930178
2 vs 0	.4541287	.0652574	6.96	0.000	.2707796	.6374778

Comparison of predictions from yogurt 3SLS model

```
. pwmean thrslshati, over(matrixi) effects sort mcompare(bonferroni) groups

Pairwise comparisons of means with equal variances

over          : matrixi
```

	Number of Comparisons
matrixi	10

thrslshati	Mean	Std. Err.	Bonferroni Groups
matrixi			
2	.5081936	.0340447	A
3	.5088245	.0892659	A
4	.5327593	.0289617	A
1	.5495131	.0248176	A
0	.5971087	.0100952	A

Note: Means sharing a letter in the group label are not significantly different at the 5% level.

	Number of Comparisons
matrixi	10

thrslshati	Contrast	Std. Err.	Bonferroni t	Bonferroni P>|t|	Bonferroni [95% Conf.	Interval]
matrixi						
2 vs 0	-.088915	.0355099	-2.50	0.124	-.1887205	.0108904
3 vs 0	-.0882842	.0898349	-0.98	1.000	-.3407776	.1642092
4 vs 0	-.0643494	.0306707	-2.10	0.360	-.1505537	.0218549
1 vs 0	-.0475956	.0267923	-1.78	0.758	-.1228991	.0277079
2 vs 1	-.0413194	.0421302	-0.98	1.000	-.1597322	.0770933
3 vs 1	-.0406886	.0926516	-0.44	1.000	-.3010985	.2197214
4 vs 1	-.0167538	.0381404	-0.44	1.000	-.1239527	.0904452
3 vs 2	.0006308	.0955376	0.01	1.000	-.2678908	.2691525
4 vs 3	.0239348	.0938465	0.26	1.000	-.2398339	.2877034
4 vs 2	.0245656	.044697	0.55	1.000	-.1010613	.1501926

Comparison of predictions from isotonics 3SLS model

References

(2016, 11 23). Retrieved from Online Etymology Dictionary: http://etymonline.com/index.php?allowed_in_frame=0&search=store

Aastrup, J., & Kotzab, H. (2010). Forty years of out-of-stock research - and shelves are still empty. *The international review of retail, distribution and consumer research*, 147-164.

ACNielsen. (2006). *Consumer-centric category management: how to increase profits by managing categories based on consumer needs.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Alderson, W. (1950). Marketing efficiency and the principle of postponement. *Cost and Profit Outlook*, 15-18.

Ali, O. G., & Pinar, E. (2016). Multi-period-ahead forecasting with residual extrapolation and information sharing -- Utilizing a mutitude of retail series. *International Journal of Forecasting*, 502-517.

Anderson, E. T., Fitzsimons, G. J., & Simester, D. (2006). Measuring and mitigating the costs of stockouts. *Management Science*, 1751-1763.

Anderson, G., & Blundell, R. (1984). Consumer Non-Durables in the U.K.: A Dynamic Demand System. *The Economic Journal* (pp. 35-44). Wiley.

Anupindi, R., Dada, M., & Gupta, S. (1998). Estimation of Consumer Demand with Stock-Out Based Substitution: An Application to Vending Machine Products . *Marketing Science*, 406-423.

Arndt, J., & Olsen, L. (1975). A research note on economies of scale in retailing. *Swedish Journal of Economics*, 207-228.

Avegliano, P., & Cardonha, C. (2014). Investigating the Hidden Losses Caused by Out-of-Shelf Events: A Multi-Agent-Based Simulation. *Proceedings of the 2014 Winter Simulation Conference*, (pp. 242-251).

Avlijaš, G., Simićević, A., Avlijaš, R., & Prodanović, M. (2015). Measuring the impact of stock-keeping unit attributes on retail stock-out performance. *Operations Management Research*, 131-141.

Awang, Z. (2012). *Structural Equation Modeling.*

Azzi, A., Battini, D., Persona, A., & Sgarbossa, F. (2012). Packaging Design: General Framework and Research Agenda. *Packaging Technology and Science*, 435-456.

Bailey, J. P., & Rabinovich, E. (2005). Internet book retailing and supply chain management: an analytical study of inventory location speculation and postponement. *Transportation Research: Part E*, 159-177.

Balakrishnan, A., Pangburn, M. S., & Stavrulaki, E. (2004). "Stack Them High, Let 'em Fly": Lot-Sizing Policies When Inventories Stimulate Demand. *Management Science*, 630-644.

Ballou, R. H. (1976). Improved Stock Location in the Physical Distribution Channel. In P. van Buijtenen, M. Christopher, & G. Wills, *Business Logistics* (pp. 190-199). The Hague: Martinus-Nijhoff.

Bartholdi, J. J., & Hackman, S. T. (2016). *Warehousing & Distribution Science.* Atlanta.

Betancourt, R., & Gautschi, D. (1988). The Economics of Retail Firms. *Managerial and Decision Economics*, 133-144.

Bonomi Santos, J., & D'Antone, S. (2014). Reinventing the wheel? A critical view of demand-chain management. *Industrial Marketing Management*, 1012-1025.

Boone, C. A., Craighead, C. W., & Hanna, J. B. (2007). Postponement: an evolving supply chain concept. *International Journal of Physical Distribution & Logistics Management*, 594-611.

Boone, T., & Ganeshan, R. (2015, January 1). Commentary on The Keys to Demand-Supply Integration Extension Beyond Fast-Moving Consumer Goods. *Foresight*, pp. 21-23.

Borin, N., Farris, P. W., & Freeland, J. R. (1994). A model for determining retail product category assortment and shelf space allocation. *Decision Sciences*, 359-384.

Bouzaabia, O., van Riel, A., & Semeijn, J. (2013). Managing in-store logistics: A fresh prespective on retail service. *Journal of Service Management*, 112-129.

Bowersox, D. J., & Closs, D. J. (1996). *Logistical Management.* New York: McGraw-Hill Companies, Inc.

Broekmeulen, Sternbeck, Donselaar, v., & Kuhn. (2016). Decision support for selecting hte optimal product unpacking location in a retail supply chain. *European Journal of Operational Research*.

Brunk, M. E. (1953). Controlled Experiments in Retail Merchandising. *Journal of Farm Economics*, 916-923.

Bucklin, L. P. (1965). Postponement, speculation and the structure of distribution channels. *Journal of Marketing Research*, 26-31.

Buzek, G. (2015). *Retailers and the Ghost Economy - $1.75 Trillion Reasons to be Afraid.* Franklin: IHL Services.

Buzek, G. (2017, April 25). *NewsFlash: Retail in US grew 4.1% in Q1 2017*. Retrieved from IHL Group: http://www.ihlservices.com/news/analyst-corner/2017/04/newsflash-retail-in-us-grew-4-1-in-q1-2017/

Cachon, G. (2001). Managing a Retailer's Shelf Space, Inventory, and Transportation. *Manufacturing & Service Operations Management*, 211-229.

Calderon-Cortes, S. J., & Morales-Arroyo, M. A. (2015). The imact of out of stock on laundry sales in a supermarket chain in Latin America using vector autoregressive analysis (VAR). *The Business & Management Review*, 100-109.

Campo, K., & Gijsbreachts, E. (2005). Retail assortment, shelf and stockout management: issues, interplay and future challenges. *Applied Stochastic Models in Business and Industry*, 383-392.

Campo, K., Gijsbrechts, E., & Nisol, P. (2003). The impact of retailer stockouts on whether, how much, and what to buy. *International Journal of Research in Marketing*, 273-286.

Campo, K., Gijsbrechts, E., & Nisol, P. (2004). Dynamics in consumer response to product unavailability: do stock-out reactions signal response to permanent assortment reductions? *Journal of Business Research*, 834-843.

Cardos, M., & Garcia-Sabater, J. P. (2006). Designing a consumer products retail chain inventory replenishment policy with the consideration of transportation costs. *International Journal of Production Economics*, 525-535.

Chan, A. T., Ngai, E. W., & Moon, K. K. (2017). The effects of strategic and manufacturing flexibilities and supply chain agility on firm performance in the fashion industry. *European Journal of Operational Research*, 486-499.

Chernev, A. (2011). Product assortment and consumer choice: An interdisciplinary review. *Foundations and Trends in Marketing*, 1-61.

Chiang, W.-y. K. (2010). Product availability in competitive and cooperative dual-channel distirubtion with stock-out based substitution. *European Journal of Operational Research*, 111-126.

Chuang, H. H.-C., Oliva, R., & Liu, S. (2015). On-Shelf Availability, Retail Performance, and External Audits: A Field Experiment. *Production and Operations Management*, 935-951.

Closs, D. J., Nyaga, G. N., & Voss, M. D. (2010). The differential impact of product complexity, inventory level, and configuration capacity on unit and order fill rate performance. *Journal of Operations Management*, 47-57.

Cohen, M. A., Agrawal, N., & Agrawal, V. (2006). Achieving breakthrough service delivery through dynamic asset deployment strategies. *Interfaces*, 259-271.

Condea, C., Thiesse, F., & Fleisch, E. (2012). RFID-enabled shelf replenishment with backroom monitoring in retail stores. *Decision Support Systems*, 839-849.

Cooper, L. G., Baron, P., Levy, W., Swisher, M., & Gogos, P. (1999). PromoCast ™: A New Forecasting Method for Promotion Planning. *Marketing Science*, 301-316.

Cooper, W. L., Homem-de-Mello, T., & Kleywegt, A. J. (2006). Models of the Spiral-Down Effect in Revenue Management. *Operations Research*, 968-987.

Corsten, D., & Gruen, T. (2003). Desperately seeking shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks. *International Journal of Retail & Distribution Management*, 605-617.

Corsten, D., & Gruen, T. (2003). Desperately seeking shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks. *International Journal of Retail & Distribution Management*, 605-617.

Corsten, D., & Gruen, T. (2005). On shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks. *Consumer Driven Electronic Transformation*, 131-149.

Curhan, R. C. (1973). Shelf Space Allocation and Profit Maximization in Mass Retailing. *Journal of Marketing*, 54-60.

Curseu, A., Woensel, T., Fransoo, J., Donselaar, K., & Broekmeulen, R. (2009). Modeling handling operations in retail stores: An empirical analysis. *The Journal of the Operational Research Society*, 200-214.

David, D. K. (1922). *Retail store management problems.* New York: A. W. Shaw Company.

Decker, C., Kubach, U., & Beigl, M. (2003). Revealing the Retail Black Box by Interaction Sensing. *ICDCS Workshops*, (pp. 328-333).

Defee, C. C., Randall, W. S., & Gibson, B. J. (2009). Roles and Capabilities of the Retail Supply Chain Organization. *Journal of Transportation Management*, 31-47.

Donselaar, K. H., Gaur, V., Woensel, T. v., Broekmeulen, R. A., & Fransoo, J. C. (2010). Ordering Behavior in Retail Stores and Implications for Automated Replenishment. *Management Science*, 766-784.

Dreyer, H., Dukovska-Popovska, I., Kiil, K., & Kaipia, R. (2017). Retail Tactical Planning: An Aligned Process? In N. I. (eds), *Advances in Production Management Systems. Initiatives for a Sustainable World* (pp. 415-422). Springer, Cham.

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

Dubelaar, C., Chow, G., & Larson, P. D. (2001). Relationships between inventory, sales and service in a retail chain store operation. *International Journal of Physical Distribution & Logistics Management*, 96-108.

Dujak, D., Ferencic, M., & Franjkovic, J. (2014). Retail Ready Packaging -- What's in it for Food Manufacturers? (pp. 31-42). Osijek, Croatia: Business Logistics in Modern Management.

Ehrenthal, J. C., & Stolzle, W. (2013). An examination of the causes for retail stockouts. *International Journal of Physical Distribution & Logistics Management*, 54-69.

Eisand, M. (2014). Shelf space elasticity: A meta-analysis. *Journal of Retailing*, 168-181.

Ellram, L., & Cooper, M. (1990). Supply chain management, partnership, and the shipper-third party relationship. *International Journal of Logistics Management*, 1-10.

Eroglu, C., & Hofer, C. (2011). Lean, leaner, too lean? The inventory-performance link revisted. *Journal of Operations Management*, 356-369.

Eroglu, C., Williams, B. D., & Waller, M. A. (2011). Consumer-drive retail operations: The moderating effects of consumer demand and case pack quantity. *International Journal of Physical Distribution & Logistics Management*, 420-434.

Eroglu, C., Williams, B. D., & Waller, M. A. (2013). The Backroom Effect in Retail Operations. *Production and Operations Management*, 915-923.

Esper, T. L., Ellinger, A. E., Stank, T. P., Flint, D. J., & Moon, M. (2010). Demand and supply integration: a conceptual framework of value creation through knowledge management. *Journal of the Academy of Marketing Science*, 5-18.

Ettouzani, Y., Yates, N., & Mena, C. (2012). Examining retail on shelf availability: promotional impact and a call for research. *International Journal of Physical Distribution & Logistics Management*, 213-243.

Evers, P. R. (1997). Hidden benefits of emergency transshipments. *Journal of Business Logistics*, 55-76.

Evers, P. T., & Wan, X. (2012). Systems analysis using simulation. *Journal of Business Logistics*, 80-89.

Eynan, A., & Fouque, T. (2003). Capturing the Risk-Pooling Effect Through Demand Reshape. *Management Science*, 704-717.

Fawcett, S. E., & Magnan, G. M. (2002). The rhetoric and reality of supply chain integration. *International Journal of Physical Distribution & Logistics*, 339-361.

Fawcett, S. E., Waller, M. A., & Fawcett, A. M. (2010). Elaborating a dynamic systems theory to understand collaborative inventory successes and failures. *International Journal of Logistics Management*, 510-537.

Federgruen, A., & Zheng, Y.-S. (1992). The joint replenishment problem with general joint cost structures. *Operations Research*, 384-403.

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry*, 105-109.

Fernie, J., & Grant, D. B. (2008). On-shelf availability: the case of a UK grocery retailer. *The International Journal of Logistics Management*, 293-308.

Fisher, M. L. (1997, March-April). What Is the Right Supply Chain for Your Product? *Harvard Business Review 75*, pp. 105-117.

Fleischmann, B., Meyr, H., & Wagner, M. (2008). Advanced Planning. In H. Stadler, & C. Kilger, *Supply Chain Management and Advanced Planning : Concepts, Models, Software, and Case Studies* (pp. 81-106). Berlin: Springer.

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 39-50.

Freimuth, R., & Lam, L. (1992). Active Walker Models for Filamentary Growth Patterns. *Woodward Conference.* New York, NY: Springer.

Frontoni, E., Mancini, A., & Zingaretti, P. (2014). Real time out of shelf detection using embedded sensor network. *Mechatronic and Embedded Systems and Applications (MESA)*, 1-6.

Frontoni, E., Marinelli, F., & Rosetti, R. (2017). Shelf space re-allocation for out of stock reduction. *Computers & Industrial Engineering*, 32-40.

Garica-Dastugue, S. J., & Lambert, D. M. (2007). Interorganizational Time-Based Postponement in the Supply Chain. *Journal of Business Logistics*, 57-81.

Gattorna, J. (1998). *Strategic Supply Chain Alignment: Best Practice in Supply Chain Management.* Gower Publishing, Ltd.

Gaur, V., Fisher, M. L., & Raman, A. (2005). An econometric analysis of inventory turnover performance in retail services. *Management Science*, 181-194.

Gelders, L. F., & van Looy, P. M. (1978). An Inventory Policy for Slow and Fast Movers in a Petrochemical Plant: A Case Study. *The Journal of the Operational Research Society*, 867-874.

Gerchak, Y., & Mossman, D. (1992). On the effect of demand randomness on inventories and costs. *Operations Research*, 804-807.

Gilland, W. G., & Heese, H. S. (2013). Sequence Matters: Shelf-Space Allocation under Dynamic Customer-Driven Substitution. *Production and Operations Management*, *22*(4), 875-887.

Goyal, S., Hardgrave, B. C., Aloysius, J. A., & DeHoratius, N. (2016). The effectiveness of RFID in backroom and sales floor inventory management. *The International Journal of Logistics Management*, 795-815.

Grewal, D., Janakiraman, R., Kalyanam, K., Kannan, P., Ratchford, B., Song, R., & Tolerico, S. (2010). Strategic Online and Offline Retail Pricing: A Review and Research Agenda. *Journal of Interactive Marketing*, 138-154.

Grubor, A., & Milicevic, N. (2015). Measuring On-shelf Availability of FMCG Products. *Industrija*, 53-71.

Grubor, A., & Milicevic, N. (2015). The analysis of FMCG product availability in retail stores. *Engineering Economics*, 67-74.

Gruen, T. W., & Corsten, D. (2008). *A Comprehensive Guide to Retail Out-of-Stock Reduction in the Fast-Moving Consumer Goods Industry.* Grocery Manufacturers of America, Food Marketing Institute, National Association of Chain Drug Stores, The Procter & Gamble Company, University of Colorado at Colorado Springs.

Gruen, T. W., Corsten, D., & Bharadwaj, S. (2002). *Retail Out of Stocks: A Worldwide Examination of Causes, Rates, and Consumer Responses.* Washington DC: Grocery Manufacturers of America.

Gunzler, D., Chen, T., Wu, P., & Zhang, H. (2013). Introduction to mediation analysis with structural equation modeling. *Shanghai Archives of Psychiatry*, 390-394.

Gurler, U., & Yilmaz, A. (2010). Inventory and coordination issues with two substitutable products. *Applied Mathematical Modelling*, 539-551.

Hair Jr, J. F., Babin, B. J., & Krey, N. (2017). Covariance-Based Structural Equation Modeling in the Journal of Advertising: Review and Recommendations. *Journal of Advertising*, 163-177.

Hausruckinger, G. (2005). *Approaches to measuring on-shelf availability at the point of sale.* München: ECR Europe.

Hellstrom, D., & Saghir, M. (2006). Packaging and Logistics Interactions in Retail Supply Chains. *Packaging Technology and Science*, 197-216.

Hendricks, K., & Singhal, V. (2003). The effect of supply chain glitches on shareholder wealth. *Journal of Operations Management*, 501-522.

Hise, R. T., Kelly, J. P., Gable, M., & McDonald, J. B. (1983). Factors affecting the performance of individual chain store units: An empirical analysis. *Journal of Retailing*, 22-39.

Hochrein, S., Glock, C. H., Bogaschewsky, R., & Heider, M. (2015). Literature reviews in supply chain management: a tertiary study. *Management Review Quarterly*, 239–280.

Honhon, D., Gaur, V., & Seshadri, S. (2010). Assortment planning and inventory decisions under stockout-based substitution. *Operations Research*, 1364-1379.

Honhon, D., Johnalagedda, S., & Pan, X. A. (2012). Optimal Algorithms for Assortment Selection Under Ranking-Based Consumer Choice Models. *Manufacturing & Service Operations Management, 14*(2), 279-289.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal* , 1-55.

Huang, Y., & Zhang, Y. C. (2016). The out-of-stock (OOS) effect on choice shares of available options. *Journal of Retailing*, 13-24.

Hubner, A. H., & Kuhn, H. (2012). Retail category management: State-of-the-art review of quantitative research and software applications in assortment and shelf space management. *Omega*, 199-209.

Hubner, A. H., Kuhn, H., & Sternbeck, M. G. (2013). Demand and supply chain planning in grocery retail: an operations planning framework. *International Journal of Retail & Distribution Management*, 512-530.

Hubner, A., & Schaal, K. (2017). Effect of replenishment and backroom on retail shelf-space planning. *Business Research*, 1-34.

Jafari, H., Nyberg, A., & Hilletofth, P. (2016). Postponement and logistics flexibility in retailing: a multiple case study from Sweden. *Industrial Management & Data Systems*, 445-465.

Jishan, Z., & Yajun, Y. (2011). Applying Postponement Strategy to a Multi-Warehouse Inventory Planning. *Proceedings of the Northeast Business & Economics Association*, (pp. 593-596).

Johnston, F. R., Boylan, J. E., & Shale, E. A. (2003). An Examination of the Size of Orders from Customers, Their Characterisation and the Implications for Inventory Control of Slow Moving Items. *The Journal of the Operational Research Society*, 833-837.

Kacena, J. J., & Hessa, J. D. (2012). Spontaneous selection: The influence of product and retailing factors on consumer impulse purchases. *Journal of Retailing and Consumer Services*, 578-588.

Kaldor, N. (1939). Speculation and Economic Stability. *The Review of Economic Studies*, 1-27.

Karakul, M., & Chan, L. (2008). Analytical and managerial implications of integrating product substitutability in the joint pricing and procurement problem. *European Journal of Operational Research*, 179-204.

Karampatsa, M., Grigoroudis, E., & Matsatsinis, N. F. (2017). Retail Category Management: A Review on Assortment and Shelf-Space Planning Models. *Operational Research in Business and Economics* (pp. 35-67). Springer International Publishing.

Khouja, M., Mehrez, A., & Rabinowitz, G. (1996). A two-item newsboy problem with substitutability. *International Journal of Production Economics*, 267-275.

Kim, M., & Lennon, S. J. (2011). Consumer Response to Online Apparel Stockouts. *Psychology & Marketing*, 115-144.

Kim, S.-H. (2014). Postponement for designing mass-customized supply chains: categorization and framework for strategic decision making. *International Journal of Supply Chain Management*, 1-11.

Kotzab, H., & El-Jafli, S. (2015). The role of instore logistics and shelf ready packaging for on-shelf availability. *Logistics Management*, 415-426.

Kuhn, H. (2013). Integrative retail logistics: An exploratory study. *Operations management research: advancing practice through research*, 2-18.

Kumar, V., Anand, A., & Song, H. (2017). Future of Retailer Profitability: An Organizing Framework. *Journal of Retailing*, 96-119.

Kurtulus, M., & Toktay, L. B. (2011). Category Captainship vs. Retailer Category Management under Limited Retail Shelf Space. *Production and Operations Management*, 47-56.

Lam, S. Y. (2001). The Effects of Store Environment on Shopping Behaviors: A Critical Review. *Advances in Consumer Research*, 190-197.

Laroche, M., Pons, R., Zgolli, N., Cervellon, M.-C., & Kim, C. (2003). A model of consumer response to two retail sales promotion techniques. *Journal of Business Research*, 513-522.

Levy, M., & Weitz, B. A. (2004). *Retailing Management.* New York: McGraw-Hill/Irwin.

Li, S., Rao, S. S., Ragu-Nathan, T., & Ragu-Nathan, B. (2005). Development and validation of a measurement instrument for studying supply chain management practices. *Journal of Operations Management*, 618-641.

Liptak, A. (2017, May 20). *Amazon could be planning to bring its checkout-free grocery stores to Europe*. Retrieved from The Verge: https://www.theverge.com/2017/5/20/15669982/amazon-go-checkout-free-grocery-stores-europe-trademark

Lo, S. C., Tung, J., & Huang, K.-P. (2017). Customer Perception and Preference on Product Packaging. *Internatonal Journal of Organizational Innovation*, 3-15.

Mahajan, S., & van Ryzin, G. (2001). Inventory competition under dynamic consumer choice. *Operations Research*, 646-657.

Maister, D. (1976). Centralisation of inventories and the "square root law". *International Journal of Physical Distribution*, 124-134.

Mangiaracina, R., Song, G., & Perego, A. (2015). Distribution network design: a literature review and a research agenda. *International Journal of Physical Distribution & Logistics Management*, 506-531.

McGillivray, A. R., & Silver, E. A. (1978). Some Concepts for Inventory Control under Substitutable Demand. *Infor, 16*(1), 47-63.

Mckinnon, A., Mendes, D., & Nababteh, M. (2007). In-store logistics: an analysis of on-shelf availability and stockout responses for three product groups. *International Journal of Logistics Research and Applications: A Leading Journal of Supply Chain Management*, 251-268.

Mertler, C. A., & Rachel, A. (2005). *Advanced and multivariate statistical methods*. Glendale, AZ: Pyrczak Publishing.

Milicevic, N., & Grubor, A. (2015). The effect of backroom size on retail product availability - operational and technological solutions. *Amfiteatru Economic* , 661-675.

Mishra, B. K., & Prasad, A. (2006). Minimizing retail shrinkage due to employee theft. *International Journal of Retail & Distribution Management*, 817-832.

Moorthy, R., Behera, S., & Verma, S. (2015). On-Shelf Availability in Retailing. *International Journal of Computer Applications*, 47-51.

Moorthy, R., Bhargave, S., Behera, S., Ramanathan, P., & Verma, S. (2015). Applying Image Processing for Detecting On-Shelf Availability and Product Positioning in Retail Stores. *Proceedings of the Third International Symposium on Women in Computing and Informatics* (pp. 451-457). ACM.

Mou, S., Robb, D. J., & DeHoratius, N. (2018). Retail Store Operations: Literature Review and Research Directions. *European Journal of Operational Research*, 399-422.

Moussaoui, I., Williams, B., Hofer, C., Aloysius, J. A., & Waller, M. (2016). Drivers of retail on-shelf availability: systematic review, critical assessment, and reflections on the road ahead. *International Journal of Physical Distribution & Logistics Management*, 516-535.

Musalem, A., Olivares, M., Bradlow, E. T., Terwiesch, C., & Corsten, D. (2010). Strutural Estimation fo the Effect of Out-of-Stocks. *Management Science*, 1180-1197.

Nachtmann, H., Waller, M. A., & Rieske, D. W. (2010). THE IMPACT OF POINT-OF-SALE DATA INACCURACY AND INVENTORY RECORD DATA ERRORS. *Journal of Business Logistics*, 149-158.

Nagarajan, M., & Rajagopalan, S. (2008). Inventory Models for Substitutable Products: Optimal Policies and Heuristics. *Management Science*, 1453-1466.

Nair, A. (2005). Linking manufacturing postponement, centralized distribution and value chain flexibility with performance. *International Journal of Production Research*, 447-463.

Netessine, S., & Rudi, N. (2003). Centralized and competitive inventory models with demand substitution. *Operations Research*, 329-335.

Nieto-Barajas, L. E., & Sinha, T. (2015). Bayesian interpolation of unequally spaced time series. *Stochastic Environmental Research & Risk Assessment*, 577-587.

Odell, J. (2002). Agent-based manufacturing: a case study. *Journal of Object Technology*, 51-61.

Oeser, G. (2010, September). *Methods of Risk Pooling in Businses Logistics and Their Application.* Retrieved December 2014, from https://opus4.kobv.de/opus4-euv/files/43/Dissertation_Gerald_Oeser.pdf

Oeser, G. (2015). *Risk-Pooling Essentials - Reducing Demand and Lead-Time Uncertainty.* New York: Cham : Springer.

Olander-Rose, M., & Nilsson, F. (2009). Competitive Advantage Through Packaging Design - Propositions for Supply Chain Effectiveness and Efficiency. *International Conference on Engineering Design* (pp. 1/279-1/290). Stanford: Design Society.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological Methods*, 434-447.

Pagh, J. D., & Cooper, M. C. (1998). Supply Chain Postponement and Speculation Strategies: How to Choose the Right Strategy. *Journal of Business Logistics*, 13-33.

Pal, J. W., & Byrom, J. W. (2003). The five Ss of retail operations: a model and tool for improvement. *International Journal of Retail & Distribution Management*, 518-528.

Panayides, P. (2013). Coefficient alpha: interpret with caution. *Europe's Journal of Psychology*, 687-696.

Papkiriakopoulos, D., & Doukidis, G. (2011). Classification performance for making decisions about products missing from the shelf. *Advances in Decision Sciences*.

Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies*, 435-480.

Piercy, N., & Ellinger, A. (2015). Demand- and supply-side cross-functional relationships: an application of disconfirmation theory. *Journal of Strategic Marketing*, 49-71.

Pires, M., Pratas, J., Liz, J., & Amorim, P. (2017). A framework for designing backroom areas in grocery stores. *International Journal of Retail & Distribution Management*, 230-252.

Pizzi, G., & Scarpi, D. (2016). The effect of shelf layout on satisfaction and perceived assortment size: an empirical assessment. *Journal of Retailing and Consumer Services*, 67-77.

Plouffe, C. R., Bolander, W., Cote, J. A., & Hochstein, B. (2016). Does the Customer Matter Most? Exploring Strategic Frontline Employees' Influence of Customers, the Internal Business Team, and External Business Partners. *Journal of Marketing*, 106-123.

Pogson, P. W. (1923). The Practical Application of the Theory of Accounting for Supplies. *Journal of Accountancy*, 430-439.

Pomerantz, L. (2014, May 6). Time For Retailers To Re-Evaluate Their Store Footprint -- One Size Does Not Fit All. *Forbes*.

Rabinovich, E., & Evers, P. T. (2003). Postponement effects on inventory performance and the impact of information systems. *International Journal of Logistics Management*, 33-48.

Rajaram, K., & Tang, C. S. (2001). The impact of product substitution on retail merchandising. *European Journal of Operational Research*, 582-601.

Raman, A., DeHoratius, N., & Ton, Z. (2001). Execution: The missing link in retail operations. *California Management Review*, pp. 136-152.

Rand, W. (2013). The future applications of agent-based modeling in marketing. In *The Routledge Companion to the Future of Marketing* (pp. 379-392). Abindon: Routledge.

Rand, W., & Rust, R. T. (2011). Agent-based modeling in marketing: Guidelines for rigor. *International Journal of Research in Marketing*, 181-193.

Randall, W. S., Gibson, B. J., Defee, C. C., & Williams, B. D. (2011). Retail supply chain management: key priorities and practices. *The International Journal of Logistics Management*, 390-402.

Regan, W. J. (1960). Self-service in retailing. *Journal of Marketing*, 43-48.

Regattieri, A., & Santarelli, G. (2013). The Important Role of Packaging in Operations Management. In P. A. (Ed.), *Operations Management* (pp. 183-220). Intech Open Science.

Reimann, M. (2012). Accurate response by postponement. *European Journal of Operational Research*, 619-628.

Reiner, G., Teller, C., & Kotzab, H. (2013). Analyzing the efficient execution of in-store logistics processes in grocery retailing - the case of dairy products. *Production and Operations Management*, 924-939.

RILA proposes to redefine 'shrinkage'. (2016). *MMR*, p. 48.

Robinson, S. (2015). Modelling without queues: adapting discrete-event simulation for service operations. *Journal of Simulation*, 195-205.

Rockwell Automation. (2013). ARENA 14.5. Coraopolis, PA, USA.

Rong, Y., Chen, Y.-J., & Shen, Z.-J. M. (2015). The Impact of Demand Uncertainty on Product Line Design Under Endogenous Substitution. *Naval Research Logistics*, 143-157.

Rosado, L., Goncalves, J., Costa, J., Ribeiro, D., & Soares, F. (2016). Supervised Learning for Out-of-Stock Detection in Panoramas of Retail Shelves. *IEEE International COnference on Imaging Systems and Techniques (IST)* (pp. 406-411). IEEE.

Rosales, C., Whipple, J. M., & Blackhurst, J. (2018). The Impact of Out-of-Stocks and Supply Chain Design on Manufacturers: Insights from an Agent-Based Model. *Transportation Journal*, 137-162.

Saghir, M., & Jonson, G. (2001). Packaging Handling Evaluation Methods in the Grocery Retail Industry. *Packaging Technology and Science*, 21-29.

Saghiri, S. (2011). A structural approach to assessing postponement strategies: construct development and validation. *International Journal of Production Research*, 6427-6450.

Saghiri, S. S., & Barnes, S. J. (2016). Supplier flexibility and postponement implementation: An empirical analysis. *International Journal of Production Economics*, 170-183.

Samli, A. C., Pohlen, T. L., & Jacobs, L. (2005). Developments in retail logistics. *Journal of Marketing Channels*, 81-98.

Sampaioa, E. Q., & Sampaio, M. (2016). Managerial response to stockouts: the effect of remedies on consumer behavior. *Production*, 66-77.

Schlapp, J., & Fleischmann, M. (2018). Multiproduct Inventory Management Under Customer Substitution and Capacity Restrictions. *Operations Research*, 740-747.

Sezen, B. (2006). Changes in performance under various lengths of review periods in a periodic review inventory control system with lost sales: A simulation study. *International Journal of Physical Distribution & Logistics Management*, 360-373.

Shah, D., Kumar, V., & Zhao, Y. (2015). Diagnosing brand performance: accounting for the dynamic impact of product availability with aggregate data. *Journal of Marketing Research*, 147-165.

Shapiro, J. F. (1999). Bottom-up vs. top-down approaches to supply chain modeling. In *Quantitative models for supply chain management* (pp. 737-759). Boston: Springer.

Shin, H., Park, S., Lee, E., & W.C., B. (2015). A classification of the literature on the planning of substitutable products. *European Journal of Operational Research*, 1-14.

Shockley, J., & Turner, T. (2015). Linking inventory efficiency, productivity and responsiveness to retail firm outperformance: empirical insights from US retailing segments. *Production Planning & Control: The Management of Operations*, 393-406.

Simicevic, A. (2015). Measuring the impact of stock-keeping unit attributes on retail stock-out performance. *Operations management research: advancing practice through research*.

Sivakumar, B. (2008). Two-commodity inventory system with retrial demand. *European Journal of Operational Research*, 70-83.

Sloane, E. H. (1945). Reductionism. *Psychological Review*, 214-223.

Smith, S. A., & Agrawal, N. (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research*, 50-64.

Snyder, H., Witell, L., Gustafsson, A., Fombelle, P., & Kristensson, P. (2016). Identifying categories of service innovation: A review and synthesis of the literature. *Journal of Business Research*, 2401-2408.

Stangl, T., & Thonemann, U. W. (2017). Equivalent Invenotry Metrics: A Behavioral Perspective. *Manufacturing & Service Operations Management*, 472-488.

Stank, T., Esper, T., Crook, R., & Autry, C. (2012). Creating relevant value through demand and supply integration. *Journal of Business Logistics*, 167-172.

Stanley, T. (2016, 6 20). 5 Trends That Are Radically Reshaping Shopper Marketing. *Adweek*, pp. 22-25.

Staples CEO Outlines Shift Away from Store Footprint. (2013, 9 17). Retrieved from RIS: Retail Info Systems: https://risnews.com/staples-ceo-outlines-shift-away-store-footprint

Statacorp. (2015). Stata 14 Base Reference Manual. College Station, TX: Stata Press.

Sternbeck, M. G. (2015). A store-oriented approach to determine order packaging quantities in grocery retailing. *Journal of Business Economics*, 569-596.

Stevens, L., & Phillips, E. E. (2017, December 17). Amazon Puzzles Over the Perfect Fit--Boxes. *The Wall Street Journal*.

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 99-103.

Stuttgen, P., Boatwright, P., & Kadane, J. B. (2018). Stockouts and Restocking: Monitoring the Retailer from the Supplier's Perspective. *Journal of Business & Economic Statistics*, 471-482.

Tabucanon, M., & Farahani, M. (1985). Solution Techniques for Inventory Replenishment Policies with Linear Demand Trend. In H. Bullinger, & W. H.J., *Toward the Factory of the Future* (pp. 320-325). Berlin: Springer.

Tan, B., & Karabati, S. (2013). Retail inventory management with stock-out based dynamic demand substitution. *International Journal of Production Economics*, 78-87.

Tang, C. S., Rajaram, K., & Ou, J. (2002). Managing demand uncertainty for short life cycle products using advance booking discount programs. In J. Geunes, P. M. Pardalos, & H. E. Romeijn, *Supply Chain Management: Models, Applications, and Research Directions* (pp. 69-95). New York: Kluwer Academic Publishers.

Taylor, J. C., & Fawcett, S. E. (2001). Retail on-shelf performance of advertised items: an assessment of supply chain effectiveness at the point of purchase. *Journal of Business Logistics*, 73-89.

Teece, D. J. (1980). Economies of scope and the scope of the enterprise. *Journal of Economic Behavior & Organization*, 223-247.

Terblanche, N. (2017). Customer Involvement, Retail Mix Elements and Customer Loyalty in Two Diverse Retail Environments. *The Customer is NOT Always Right? Marketing Orientationsin a Dynamic Business World. Developments in Marketing Science: Proceedings of the Academy of Marketing Science.* (pp. 795-804). Springer.

Terekhov, D., & Beck, J. C. (2009). An extended queueing control model for facilities with front room and back room operations and mixed-skilled workers. *European Journal of Operational Research*, 223-231.

Teunter, R. H., Syntetos, A., & Babai, M. Z. (2017). Stock keeping unit fill rate specification. *European Journal of Operational Research*, 917-925.

Tiwari, S., Jaggi, C. K., Gupta, M., & Cardenas-Barron, L. E. (2018). Optimal pricing and lot-sizing policy for supply chain system withdeteriorating items under limited storage capacity. *International Journal of Production Economics*, 278-290.

Ton, Z., & Raman, A. (2010). The effect of producty variety and inventory levels on retail store sales: A longitudinal study. *Production and Operations Management*, 546-560.

Trautrims, A., Grant, D. B., Fernie, J., & Harrison, T. (2009). Optimizing on-shelf availability for customer service and profit. *Journal of Business Logistics*, 231-247.

Trauzettel, V. (2014). Optimal Stocking of Retail Outlets: the Case of Weekly Demand Pattern. *14th International Scientific Conference - Business Logistics in Modern Management*, (pp. 3-11). Osijek, Croatia.

Tse, Y., Tan, K., Ting, S., Choy, K., Ho, G., & Chung, S. (2012). Improving postponement operation in warehouse: an intelligent pick-and-pack decision-support system. *International Journal of Production Research*, 7181-7197.

Turk, J. I. (2012). *The Impact of Stockouts on Customer Loyalty to Lean Retailers.* Walden University.

Twede, D., Clarke, R. H., & Tait, J. A. (2000). Packaging postponement: a global packaging strategy. *Packaging Technology and Science*, 105-115.

University of Cambridge. (2018, 7 23). *FAQ/effect size*. Retrieved from University of Cambridge Cognition and Brain Sciences Unit: http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize

van Donselaar, K. H., & Broekmeulen, R. (2008). *Static versus dynamic safety stocks in a retail environment with weekly sales patterns.* Technische Universiteit Eindhoven.

van Donselaar, K., van Woensel, T., Broekmeulen, R., & Fransoo, J. (2005). Improvement opportunities in retail logistics. In G. I. Doukidis, & A. P. Vrechopoulos, *Consumer Driven Electronic Transformation: Applying New Technologies to Enthuse Consumers and Transform the Supply Chain* (pp. 9-21). Springer-Verlag Berlin Heidelberg.

van Hoek, R. I. (1998). Reconfiguring the Supply Chain to Implement Postponed Manufacturing. *The International Journal of Logistics Management*, 95-110.

van Zelst, S., van Donselaar, K., van Woensel, T., Broekmeulen, R., & Fransoo, J. (2009). Logistics drivers for shelf stacking in grocery retail stores: Potential for efficiency improvement. *International Journal of Production Economics*, 620-632.

Vernuccio, M., Cozzolino, A., & Michelini, L. (2010). An exploratory study of marketing, logistics and ethics in packaging innovation. *European Journal of Innovation Management*, 333-354.

Vulcano, G., Van Ryzin, G., & Ratliff, R. (2012). Estimating primary demand for substitutable products from sales transaction data. *Operations Research*, 313-334.

Wagner, J., Ettenson, R., & Parrish, J. (1989). Vendor Selection Among Retail Buyers: An Analysis by Merchandise Division. *Journal of Retailing*, 58-79.

Waller, M. A., Heintz Tangari, A., & Williams, B. D. (2008). Case pack quantity's effect on retail market share: An examination of the backroom logistics effect and the store-level fill rate effect. *International Journal of Physical Distribution & Logistics Management*, 436-451.

Waller, M. A., Williams, B. D., Heintz Tangari, A., & Burton, S. (2010). Marketing at the retail shelf: an examination of moderating effects of logistics on SKU market share. *Journal of the Academy of Marketing Science*, 105-117.

Walter, C., & Grabner, J. R. (1975). Stockout cost models: empirical tests in a retail situation. *Journal of Marketing*, 56-60.

Wan, X., Evers, P. T., & Dresner, M. E. (2012). Too much of a good thing: The impact of product variety on operations and sales performance. *Journal of Operations Management*, 316-324.

Wever, R. (2011). Design for volume optimization of packaging for durable goods. *Packaging Technology and Science*, 211-222.

Whitin, T. M., & Youngs, J. W. (1955). A method for calculating optimal inventory levels and delivery time. *Naval Research Logistics*, 157-173.

Williams, B. D., & Waller, M. A. (2011). Top-Down Versus Bottom-Up Demand Forecasts: The Value of Shared Point-of-Sale Data in the Retail Supply Chain. *Journal of Business Logistics*, 17-26.

Windrum, P., Fagiolo, G., & Moneta, A. (2007). Empirical Validation of Agent-Based Models: Alternatives and Prospects. *Journal of Artificial Societies and Social Simulation*, 8.

Wintle, F., & Patch, W. (2003, January 15). Slow-Moving Inventory: All Dressed Up and Nowhere to Go. *Inbound Logistics*.

Wu, T., Huang, S., Blackhurst, J., Zhang, X., & Wang, S. (2013). Supply Chain Risk Management: An Agent-Based Simulation to Study the Impact of Retail Stockouts. *IEEE Transactions on Engineering Management*, 676-686.

Wu, Z., Zhai, X., & Liu, Z. (2015). The inventory billboard effect on lead-time decision. *International Journal of Production Economics*, 45-53.

Xiao, S. H., & Nicholson, M. (2013). A multidisciplinary cognitive behavioural framework of impulse buying: a systematic review of the literature. *International Journal of Management Reviews*, 333-356.

Xue, W., Caliskan Demirag, O., Chen, F. Y., & Yang, Y. (2017). Managing Retail Shelf and Backroom Inventories When Demand Depends on the Shelf-Stock Level. *Production and Operations Management*, doi 10.1111/poms.12713.

Yang, B., Burns, N. D., & Backhouse, C. J. (2004). Postponement: a review and an integrated framework. *International Journal of Operations & Production Management*, 468-487.

Yang, B., Yang, Y., & Wijngaard, J. (2007). Postponement: an inter-organizational perspective. *International Journal of Production Research*, 971-988.

Yang, H., & Schrage, L. (2009). Conditions that cause risk pooling to increase inventory. *European Journal of Operational Research, 192*, 837-851.

Zentes, J., Morschett, D., & Schramm-Klein, H. (2007). *Startegic Retail Management.* Wiesbaden: Gabler - Springer Science+Business Media.

Zhang, C., & Tan, G. (2001). Classification of Postponement Strategies and Performance Metrics Framework. *PACIS 2001 Proceedings* (pp. 45-58). PACIS.

Zhang, J., Gou, Q., Zhang, J., & Liang, L. (2014). Supply chain pricing decisions with price reduction during the selling season. *International Journal of Production Research*, 165-187.

Zhou, W., & Piramuthu, S. (2015). Effect of ticket-switching on inventory and shelf-space allocation. *Decision Support Systems*, 31-39.

Zinn, W., & Bowersox, D. J. (1988). Planning physical distribution with the principle of postponement. *Journal of Business Logistics*, 117-136.

Zinn, W., & Liu, P. C. (2001). Consumer Response to Retail Stockouts. *Journal of Business Logistics*, 49-71.

Zinn, W., & Liu, P. C. (2008). A comparison of actual and intended consumer behavior in response to retail stockouts. *Journal of Business Logistics*, 141-159.

Zondag, M. M., & Ferrin, B. (2014). Finding the True Voice of the Customer in CPG Supply Chains: Shopper-Centric Supply Chain Management. *Journal of Business Logistics*, 268-274.