2017

A Framework for Centralizing Inventory in Pharmaceutical Supply Chains

Kerry Kathleen Ward
Wright State University

A FRAMEWORK FOR CENTRALIZING INVENTORY IN PHARMACEUTICAL SUPPLY CHAINS

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Industrial and Human Factors Engineering

By

KERRY KATHLEEN WARD
B.S., Kent State University 2013

2017
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

April 26, 2017

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Kerry Kathleen Ward ENTITLED A Framework for Centralizing Inventory in Pharmaceutical Supply Chains BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Industrial and Human Factors Engineering.

Frank W. Ciarallo, Ph.D., Thesis Director

Jaime E. Ramirez-Vick, Ph.D., Chair
Department of Biomedical, Industrial
and Human Factors Engineering

Committee on
Final Examination

Pratik Parikh, Ph.D.

George Polak, Ph.D.

Frank W. Ciarallo, Ph.D.

Robert E.W. Fyffe, Ph.D.
Vice President for Research and
Dean of the Graduate School

ABSTRACT

Ward, Kerry Kathleen. M.S.I.H.E. Department of Biomedical, Industrial and Human Factors Engineering. Wright State University, 2017. A Framework for Centralizing Inventory in Pharmaceutical Supply Chains.

Effectively choosing the location of inventory in a pharmaceutical supply chain is central to the mission of pharmaceutical distributors. These supply chains are a critical backbone to the healthcare delivery process: the distributors must deliver products to pharmacies, hospitals and retailers at low cost and with a high level of reliability. This thesis provides a method for analyzing demand data to gain insight on the value of centralizing the inventory into either a hub-and-spoke or national network for a pharmaceutical distribution company. Demand data is analyzed using an off-the-shelf analysis tool as well as a new tool developed in SQL Server, finding expected daily demand, inventory targets for a specific customer service level, and expected inventory. Annual transportation costs for centralized locations are calculated using an overnight delivery cost rating tool in SQL Server and net savings are calculated for each product in each network. The results from this analysis are used in an Excel tool to select the optimal group of products for centralization to maximize savings while keeping transportation costs at or below a given budget. The results show that using a hub-and-spoke network can save over $10 million.

Contents

List of Figures

List of Tables

1 Introduction

In distribution networks, inventory may be held in multiple distribution centers across a large geographic area. While a company makes money selling products and supporting their retail customers, any inventory that remains in a distribution center incurs a holding cost related to the value of the product. The demand for some products is very sporadic, so when held in multiple distribution centers, the holding costs can accumulate very quickly. To help reduce these costs, a company can pool the inventory for products into a centralized network, holding inventory in a small fraction of the distribution centers, or even in one central location. By pooling the inventory, the overall inventory holding costs are reduced, though additional transportation costs are incurred.

For this thesis, the network of a large pharmaceutical distribution company is analyzed to evaluate the value of utilizing a centralized network. The company has a hub-and-spoke network developed for a project studying the inbound shipment aspect of their company and is looking to take advantage of the hubs by pooling the inventory from individual distribution centers for more products, instead of only the products that are delivered to the hubs to meet order minimums. The company's hub-and-spoke network is a fixed network, so this thesis will not study the hub location problem and will only study the use of the network for the outbound aspect of the business. The fixed hub-and-spoke network designates three of the twenty distribution centers in the company's network as hubs and the demand from each of the remaining distribution centers is assigned to one of the hubs. The three hubs are spread across the country, with one on both the East and West coast of the United States and the third in a central location. The company is also interested in exploring a fully centralized, or national, network for products that move slowly through the network and can be efficiently distributed to all customers from one location.

The analysis in this thesis makes several assumptions. First, the inbound side of the business is not considered, only the outbound side and general inventory policies. It is also assumed that the locations that supply the studied locations have a responsive and effective inventory control system. By studying demand and improving inventory management at the studied locations, the job of the inventory system elsewhere in the network becomes easier, leading to a smoother, more regular process. The effect that increasing or reducing inventory at a location has on labor and facility costs is also not considered in the analysis, but may be in future work. The company also gives an estimated inventory holding cost of 10%, though typical inventory holding costs range from 25-55% (Richardson, 1995). Finally, the analysis assumes that the company uses an order-up-to policy for inventory, meaning that the amount ordered is determined by the current inventory and the desired level of inventory, and that there is no fixed cost for placing an order. In the current system, there is a dedicated fleet that delivers to the studied locations each day.

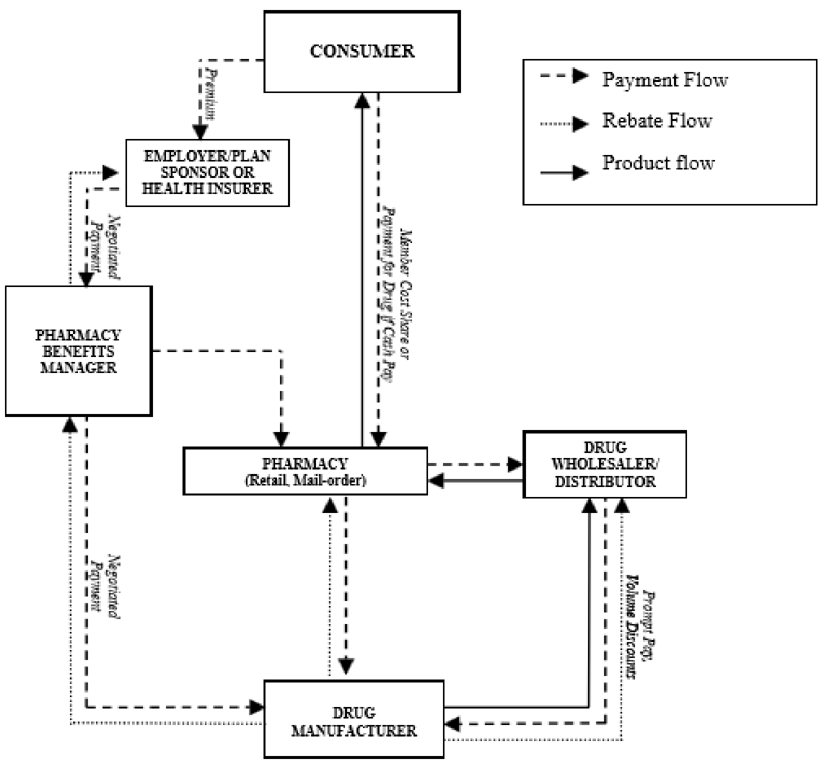

Figure 1: Flow of Products and Payments in a Pharmaceutical Supply Chain (Kaiser Family Foundation, 2005)

In a general supply chain, products and payments flow in opposing directions in a relatively straightforward way through a network of manufacturers, wholesalers, and retailers. Figure 1 shows that for a pharmaceutical supply chain, the flow is much more complex. In a pharmaceutical supply chain, the flow of products is similar to other types of supply chains, while the flow of payments follows a much more complicated pathway. A pharmaceutical supply chain includes the influence of (1) health insurance companies, as they are responsible for large portions of payments for products, as well as (2) pharmacy benefits managers, who manage mail order pharmaceuticals, and processing/paying prescription drug claims for insurance companies. With a complex flow of payments and products, it is important for each member of the supply chain to be as cost efficient as possible.

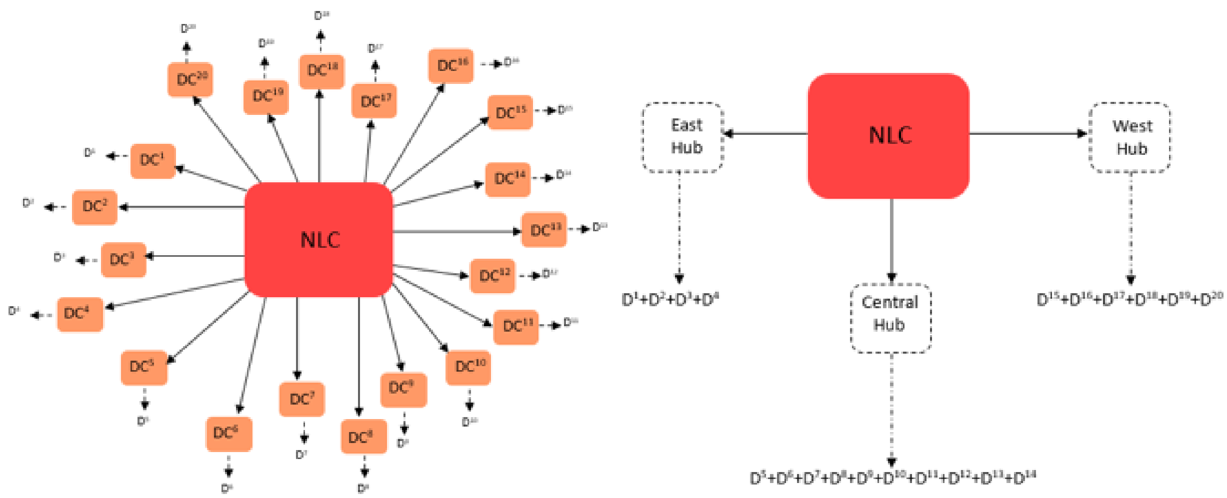

Figure 2: Decentralized vs. Hub and Spoke Network

Pooling inventory in a pharmaceutical supply chain network not only saves money by reducing inventory, it can help avoid excessive waste due to expiring products. In a pharmaceutical supply chain, many products are small in size and highly valuable, so the cost of transporting them across the network is relatively low, while the reduction in inventory holding costs is relatively high. When evaluating potential savings for the company, the analysis in this thesis showed a potential for multi-million dollar savings, even when centralizing a small fraction of the products in the network. Figure 2 demonstrates how pooling the inventory for the company combines the demands from the distribution centers included in each of the hubs. By combining the demands, variability in both the demand size and average demand intervals is reduced (see Section 3.5 for more details).

The process mapped out in Figure 3 shows the steps taken in this thesis to evaluate the potential savings from using a centralized network for the company. The resulting analysis is complex, featuring many steps and is supported by a large amount of data. The company provided real demand data for the entire network, information about the products, including unit cost and unit size (in cubic inches), and the layout of the predefined hub-and-spoke network. The company also designated a predetermined national hub, a budget for transportation costs, limits on space available in each of the hubs, and an overnight shipment rating tool for estimating the cost of shipments. Many of the steps in each portion of the overall analysis can be completed simultaneously across items and others, such as the overnight shipment cost rating tool are only completed once throughout the entire process.

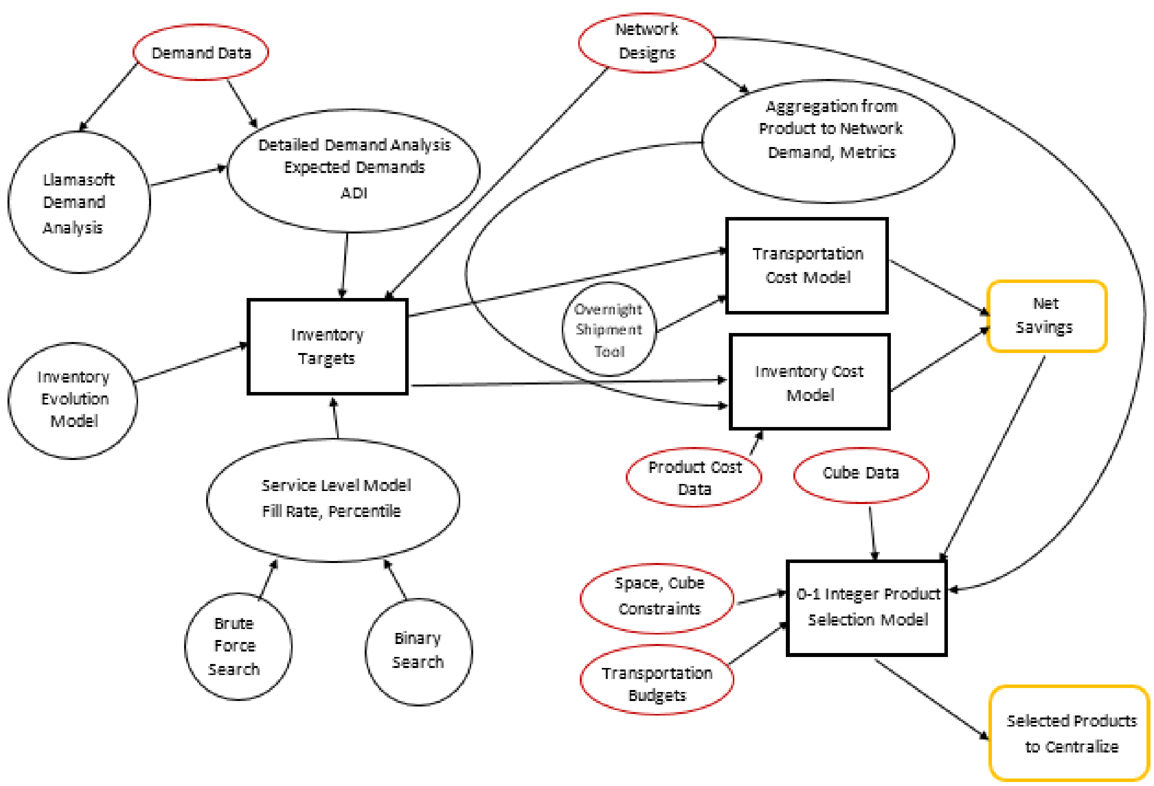

Figure 3: Process Diagram for the Demand and Inventory Analysis

First, the demand patterns are categorized using the Supply Chain Guru® software from Llamasoft. The classifications are a helpful tool when selecting the products that will be considered for centralization. A detailed demand analysis is then completed using SQL Server, computing the expected demand for non-zero demand, $E[D_t]$, for all products in all locations as follows:

$$E[D_t] = E[D_t | D_t > 0] \Pr\{D_t > 0\}$$

The expected demand for non-zero demand for a product considers the number of days with zero demand versus the number of days with non-zero demand, which helps give a more accurate idea of what a non-zero demand will look like. With this calculation the average demand interval is also computed to estimate the frequency of orders and further evaluate the benefits of centralizing the inventory. The value for expected demand for non-zero demand is then used in the inventory analysis, in which inventory targets, average inventory, and inventory costs are computed.

The cost of inventory includes many parts, including inventory holding costs, order costs, storage space costs, and many others. The analysis completed in this thesis focuses solely on the holding cost of inventory, as the goal is to reduce total inventory of a product in the network. To calculate the average inventory, I_t, for a given product in a given location, an inventory target for each product at each location must be set. The inventory for a product in each period, t, is calculated as follows (where D_t is the demand in period t):

$$I_t = I^{target} - D_t$$

The inventory target, I^{target}, is the required number of units of a product to be kept on-hand and is determined by the demand pattern and the desired customer service level, or fill rate, set by the company. The fill rate is a complicated value to find, so a binary search is used to efficiently set the correct values for each product in each location. As the binary search runs, the backorder level, B_t, is updated each time the inventory target is changed to calculate the fill rate as follows:

$$Fill\ Rate = 1 - \frac{\sum_{t=StartDate}^{EndDate} B_t}{\sum_{t=StartDate}^{EndDate} D_t}$$

For more detail on the calculation of the fill rate, inventory targets, and average inventory, see Section 3.3.3.

After the inventory analysis has been completed and the expected savings in inventory holding costs are calculated, the transportation costs associated with utilizing a hub-and-spoke network are computed. These costs are based on the actual shipments made by the company in the analysis time frame and are computed using an overnight shipment cost rating tool in SQL. The analysis is completed for a 6 ½ month time frame of actual demand data for over 30,000 products. The transportation costs are annualized to create a fair comparison with annual inventory costs and thus calculate the expected net savings for centralizing a product into either the hub-and-spoke or national network.

The net savings for the national aggregation for each product, p, is calculated as follows:

$$NetSavings_{Nat,p} = InventorySavings_{Nat,p} - C_{Nat,p}^{Trans}$$

The results of the initial analysis of net savings are used to define a set of products to be considered for centralization into either the hub-and-spoke or national network. Some products show a potential for large inventory holding cost savings, such as the one shown in Figure 4 with expected inventory savings of 67% when using the hub-and-spoke network and expected net savings of 82.5% when using the national network. The centralization of this product also reduces the average demand interval from just over 4 days to 1.24 days in the hub-and-spoke network and to about 1 day in the national network. This product shows that pooling inventory can dramatically reduce the total inventory of a product throughout the network, as well as reduce the average demand interval, leading to a more regular and predictable process.

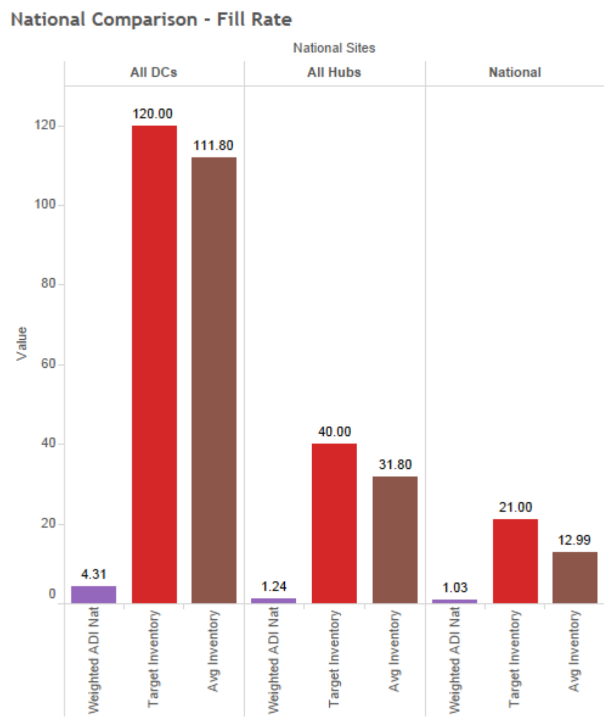

Figure 4: Sample Results from the Centralization of a Single Product

Once a reduced set of candidates for centralization into either the hub-and-spoke or national network is found, a 0-1 integer product selection model is developed and implemented in Excel, as seen in Section 4.2. This model investigates the trade-off between the potential savings for each product and the annual cost of transportation for the product in either network to recommend which products should be centralized into either network. When a product is being considered for centralization, the company must also account for the total space required to hold the required inventory in each of the hubs and the limit on how much money is available to spend on the additional transportation. For example, as shown in Section 4.4, consider a transportation budget set at \$1.5 million, with space constraints set low compared to the total space required for centralizing all 50 of the products in a sample data set. In this scenario, 29 items were recommended to be centralized nationally, 7 items were recommended to be centralized in the hub-and-spoke network, and the remaining 14 items were not recommended for centralization. This scenario had a total of over \$9 million in expected savings, with expected transportation costs of approximately \$1.5 million.

The remainder of the thesis is organized as follows. Section 2 presents a review of relevant literature in the fields of demand categorization, supply chain and inventory management, and pharmaceutical supply chains. Section 3 describes the methods used to categorize demand and investigate the value of utilizing a centralized network for the products in the company's network. Section 4 introduces a tool for selecting the products to be centralized. Finally, Section 5 presents conclusions and recommendations for further research.

2 Background

2.1 Demand Categorization

In any supply chain, it is important to have an idea of what future demand will look like to help set production and inventory goals. Most forecasting methods, however, are not necessarily suitable for all types of demand, so it is also important to categorize the demand of a product prior to forecasting its future demands. The paper by Syntetos, Boylan, and Croston (2005) compares two older, more traditional categorization schemes with a newer alternative approach. Demand categorization schemes are typically based on how often demand occurs (intermittence) and the variability of the demand size (lumpiness). In the two older methods compared, the cutoff points for each category were chosen based on the data used, so these methods may not be reasonable for other demand data sets. The newly proposed method in this paper uses the mean inter-demand interval (p) and the squared coefficient of variation of demand sizes (CV^2) to categorize the demand patterns. The cutoff points for different categories are based on an analysis of the performance of forecasting methods at different values of both variables. Figure 5 shows the framework for categorizing demand based on the method proposed by Syntetos, Boylan, and Croston (2005), along with the forecasting method appropriate for each category.

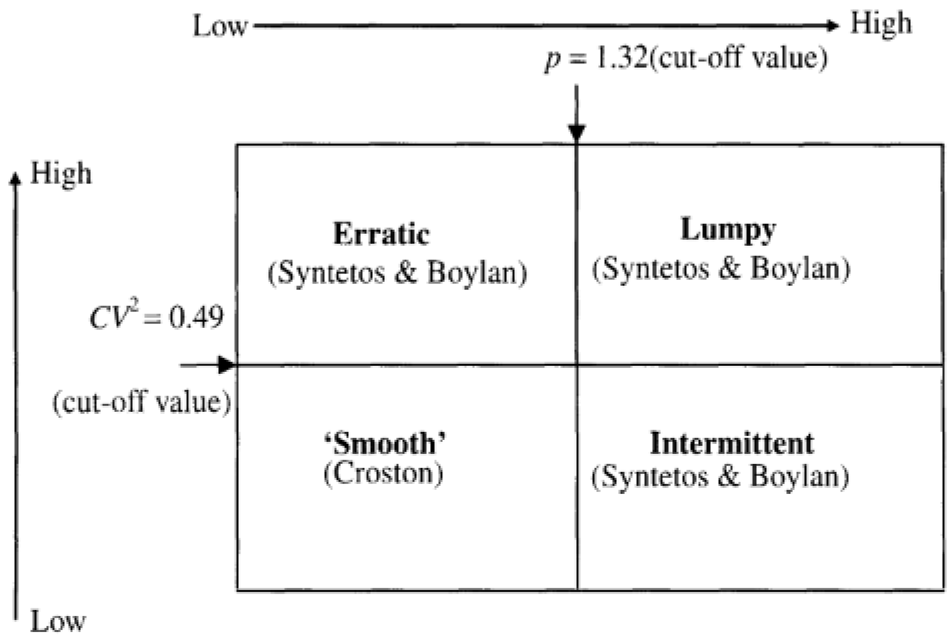

Figure 5: Categorization scheme as proposed by Boylan, Syntetos, and Karakostas (2008)

This framework is used in different demand categorization and analysis software, though cut-off values (names for categories may vary slightly in different sources). The Supply Chain Guru® software uses the cut-off values shown in Figure 6 for the classification of demand. For more information on Supply Chain Guru®, see Section 3.2.

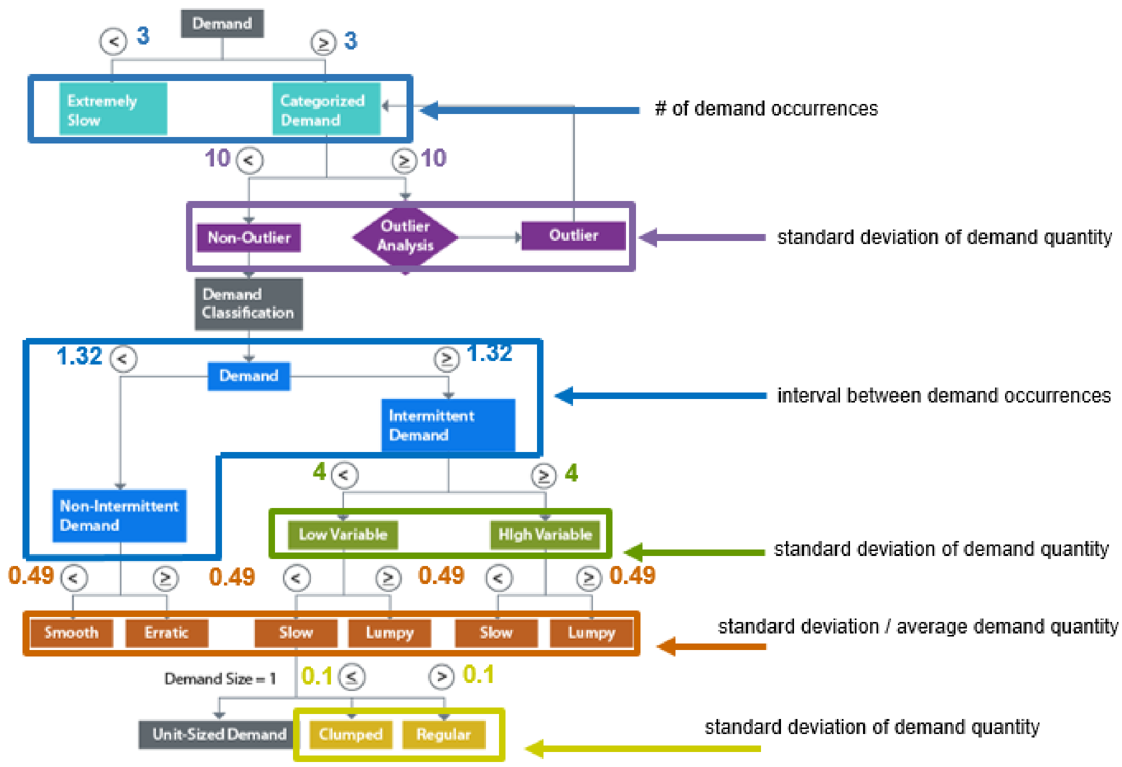

Figure 6: Supply Chain Guru® Classification Flowchart

The distinction between ‘smooth’ and ‘erratic’ demand is small, with a specific cut-off point for demand variability defining the difference. Figure 7 shows this distinction, with the size of demand (height of the bars) varying considerably more in the ‘erratic’ pattern than in the ‘smooth’ pattern. Both of these patterns have very few non-zero demand periods, thus the demand has a low inter-demand interval mean and is considered to be non-intermittent.

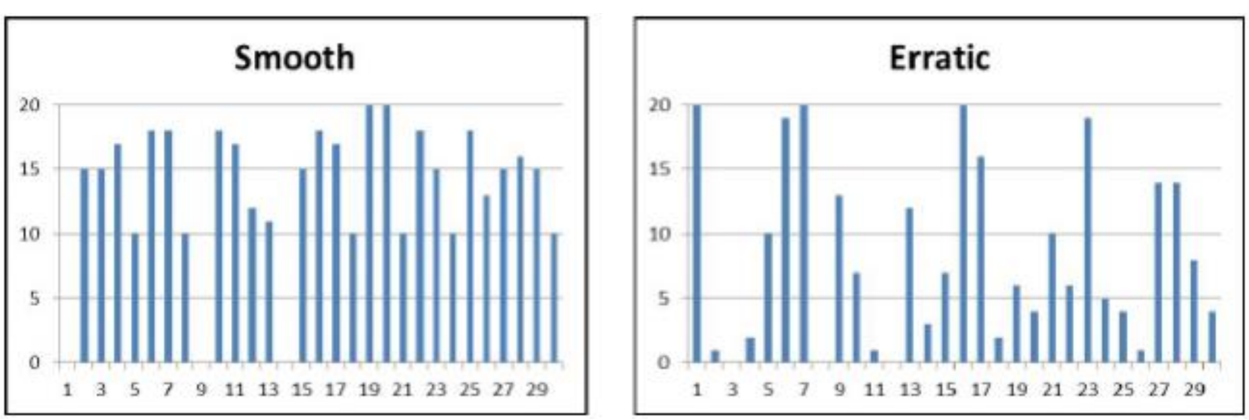

Figure 7: Graphical representation of smooth vs erratic demand from the Supply Chain Guru® documentation

The distinction between 'slow' and 'lumpy' demand is more difficult to discern, as both have high inter-demand interval means and the variability in demand sizes can be very close. Figure 8 shows this distinction, with the demand size variability clearly higher in the 'lumpy' pattern than in the 'slow' pattern. In the Supply Chain Guru® documentation, the distinction between these two shows that 'lumpy' demand is always the highest variability for intermittent demands, so 'slow-highly variable' demand is still overall less variable than 'lumpy' demand.

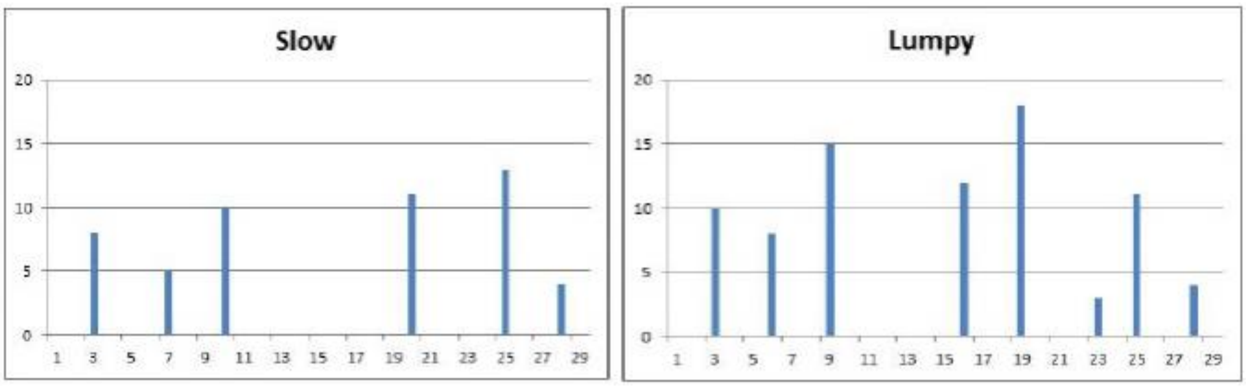

Figure 8: Graphical representation of slow vs lumpy demand from the Supply Chain Guru® documentation

2.2 General Supply Chain

In a traditional supply chain, products start at the manufacturer and move through a wholesale distributor to retailers for consumers to purchase. For a wholesale distributor, it is very important to set up the network in such a way that costs are at a minimum while the customer service level remains as high as possible. The wholesale network consists of distribution centers, a transportation fleet, and customers. A wholesale distributor must make many important decisions, including location and number of distribution centers, which distribution center will service each customer, how much of each product to keep at each distribution center, and how frequently products can be delivered. To decide how much of each product to keep at each distribution center, the wholesaler should first set their desired customer service level, which is the "fraction of orders filled on or before their due dates" (Hopp & Spearman, 2008). The service level can be computed using one of two methods: either as the probability that a given demand will be satisfied, or as the fill rate, which is the "fraction of demands that are met out of stock" (Hopp & Spearman, 2008). The result of the customer service level computation chosen will then be used to set a safety stock level, which finally leads to the total amount of inventory that will be kept on hand at any given time.

In many traditional supply chains, customers are serviced by only one distribution center for all orders. Companies with many distribution centers may choose to centralize the inventory of products into one or two locations to save money on inventory holding costs. It is most beneficial to a company to centralize, or pool, inventory when the demand for a product at individual distribution centers is highly variable, as the variability in demand at the centralized location will be reduced significantly (Berman, Krass, & Tajbakhsh, 2011). The centralization of inventory in a distribution network benefits from an effect known as risk pooling.

2.2.1 Inventory Pooling

Companies use inventory pooling as a way to reduce the variability of demand patterns and save money on inventory holding costs. Pooling inventory reduces the total number of units that need to be held in the system in order to fulfill the demands in the network at a given service level.

For example, consider a supply chain with n=4 locations stocking a single item, such as the sample network in Figure 9. Assume that each demand $D_1 \ldots D_4$ is normally distributed with mean θ and standard deviation σ. The base stock level at each location for a 97.7% fill rate for the item is $I_i = \theta + 2\sigma$.

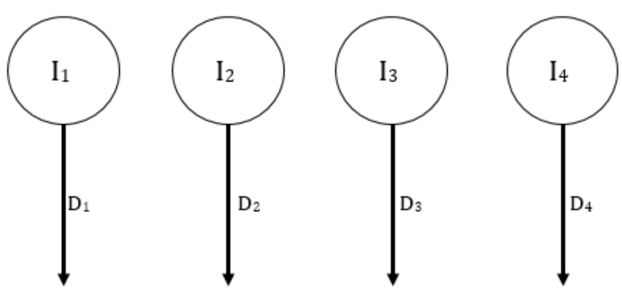

Figure 9: Supply Chain with 4 Locations

The cycle stock in each location (inventory for average demand) is θ and the safety stock (inventory to account for variance) is 2σ. Since the demand is identically distributed across all 4 locations, the total safety stock in the network is $2n\sigma$, or 8σ. When the inventory for this product is pooled into one location, as shown in Figure 10, the total demand is normally distributed with mean $n\theta = 4\theta$ and standard deviation $\sqrt{n}\sigma = 2\sigma$.

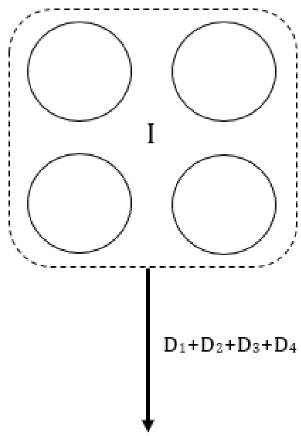

Figure 10: Supply Chain with Pooled Inventory

In order to meet the same 97.7% fill rate, the base stock level for the pooled inventory is now:

$$I = n\theta + 2\sqrt{n}\sigma = 4\theta + 4\sigma$$

The total safety stock for the pooled inventory is 4σ instead of 8σ in the original layout, reducing the safety stock by 50% for this product. In general, for this type of scenario, the ratio of safety stock in the decentralized network to the safety stock in the pooled network is $1/\sqrt{n}$, where *n* is the number of locations with inventory being pooled. Thus, safety stock is reduced by a factor of $\sqrt{n}$.

The paper by Berman, Krass, and Tajbakhsh (2011) studies the benefits of pooling inventory in a network with multiple locations versus the traditional non-pooled network. The paper demonstrates the idea with a newsvendor model with *n* locations with independent and identically distributed demands and studies the cost savings associated with inventory pooling, as well as the relationship between demand variability and benefits of inventory pooling. The paper finds that with very high variability in demand at individual locations, centralization of that item is necessary to maintain normal operations of the system.

When pooling inventory, the tradeoff between demand variance and cost variance is considered. The paper by Schmitt, Sun, Snyder, and Shen (2015) compares a centralized network with a decentralized network where the network has some level of supply uncertainty. The centralized network takes advantage of the *risk pooling effect*, where the variance of demand across multiple locations is combined, which results in lower inventories and lower expected costs. The decentralized network takes advantage of the *risk diversification effect*, reducing the negative impact of disruptions in the supply chain, which lowers the variance of the cost. In a network with multiple locations and both demand uncertainty and supply disruptions, the risk diversification effect is dominant when there are more frequent, longer, and costly disruptions. In the same system, the risk pooling effect is intensified as the number of locations increases, and this effect is dominant with larger demand uncertainty.

Companies may also choose to "virtually" centralize their inventory, meaning that customers may not receive all of their products from their primary assigned location, but also from secondary locations that are used when the primary location is not capable of filling the order. The paper by Ballou and Burnetas (2003) compares a traditional method of filling demand with a method in which an order can be filled from any number of locations. In a traditional system, each customer has an assigned location from which all orders are fulfilled and any demand that cannot be immediately satisfied is either put on backorder or is lost. In this method, the target fill rate is set at less than 100% in order to avoid holding excess inventory, based on demand forecasts. In a cross filling system, customers are assigned primary and secondary locations from which orders are filled. If an order is not completely filled from the primary location, the remaining demand is filled from the secondary location. This method allows each location to hold a lower inventory while maintaining the same or even higher fill rate. For example, stocking an item at an 80% fill rate level in four locations returns a probability of fulfilling the order using the cross filling system of 99.8% (1-0.2*0.2*0.2*0.2). In order to serve customers at that service level in a traditional system, inventory levels would be much higher and much more costly. The paper finds that not all items in a system are better served in the cross filling system, such as items that are consistently stocked for a very high fill rate. Items that are typically stocked at a lower fill rate benefit much more, with overall inventory decreasing in a cross filling system. Items that should be included in the cross-filling system will show a "favorable tradeoff of regular stock with safety stock" and all of the costs associated with the system will be vital to the decision. The paper does not consider the additional transportation costs that are required when fulfilling orders from different locations.

2.2.2 Hub and Spoke Network Design

When considering inventory pooling, it is also necessary to consider the layout of the network in which the inventory will move. Usually, a company will set up a hub-and-spoke network for this purpose, where a select few locations will serve as a central hub and distribute goods through the remaining locations (the spokes) in the network. The benefit of a single or very few hubs is that manufacturers need to make deliveries to a single location. This allows manufacturers to focus on production rather than complex logistics. Hub and spoke designs are also often used in transportation networks (such as trains, planes, and buses) or telecommunications networks. The use of a hub-and-spoke network can help a company to reduce costs, centralize processing, and allow transportation providers to take advantage of economies of scale by reducing the number of locations visited (Skipper, Cunningham, Boone, & Hill, 2016).

The paper by Correia, Nickel, and Saldanha-da-Gama (2011) uses a mixed-integer linear programming model to study the design of a hub-and-spoke network with variable capacity and balancing requirements. The objective of the model is to minimize total costs, including collecting, transferring, and distributing the flow, as well as the initial cost of establishing a hub. With larger quantities of product flowing in the sub-network defined by the hubs, the use of a hub-and-spoke network allows the company to take advantage of economies of scale related to transfer costs. In using balancing requirements, the network is set up such that each hub services a similar number of spokes in the network. This prevents the overloading of an individual hub. The paper uses two different capacity levels at each site, with the input value set at the beginning of the model, which describes the volume each site is capable of handling.

The paper by Yoon and Current (2008) studies the hub location and network design problem with a mixed zero-one integer programming formulation. The model is used to determine the number and locations of hubs, which network arcs should be included, and the routes to be used in order to minimize the total network cost. The total network cost includes the fixed cost of establishing a hub, the fixed cost of including an arc in the network, and the costs associated with traffic on the arcs. A network arc is a transportation route between two hubs or between a hub and a customer.

The paper by Zäpfel and Wasner (2002) uses a mathematical model to study the line haul problem, where the cost of transportation is minimized, as well as a model for the combined line haul and pickup/delivery problem, where the sum of the total transportation costs is minimized. In the design of a hub-and-spoke network, the strategic decisions include the selection of suitable locations, the assignment of customers to depots, the routing of the line haul, and the kind of transportation used. The operational decisions include the daily disposition of the number of trucks and the planning of pick-up/delivery routes. In a case study example of a real parcel service provider, the model showed an average of about 10% reduction in costs, so long as shipments can be made to the receiving site without being forced to move products through the hub.

The paper by Skipper, Cunningham, Boone, and Hill (2016) uses a multiple objective linear programming model to analyze optimal hub locations in a military-based example. The objectives of the model are to minimize cost while also minimizing the time needed to meet demands. The model is used to study an example network, and choose a single hub location.

2.3 Pharmaceutical Supply Chain

Pharmaceutical supply chains (PSCs) have many of the same general features as a standard supply chain, with manufacturers, distributors, and customers. PSCs have a number of unique traits which require specific attention.

2.3.1 Difficulties in the PSC

Since a PSC is responsible for distributing drugs from the manufacturers to the patients who need them, it is important for those products to be carefully tracked throughout their time in the system. As a result, the PSC is very complex with many difficulties at each step. In more recent years, regulations governing the research and development of new drugs has expanded, causing this phase to last as long as 10 to 15 years. The extended research and development phase has shortened the period of time for a company to have exclusive rights over a new product, which in turn causes drug prices to increase (Pedroso & Nakano, 2009). Another key struggle in the early stages of designing and operating the supply chain are that demand for new and existing products is difficult to forecast, with health care needs and competition constantly changing.

The paper by Pedroso and Nakano (2009) studies the flow of information in a PSC through interviews. It presents a qualitative and exploratory study of real pharmaceutical companies. Since the final customer (the patient) does not have much of a role in the product choice decision, it is important to keep the physicians and medical practices well informed about current drugs and drugs that are in production to ensure the creation of demand. Many pharmaceutical companies will begin advertising new products once clinical trials have reached the final stage so that physicians and other prescribers will begin generating demand when the product is ready for distribution.

As demand for pharmaceuticals continues to increase in the modern, dynamic healthcare market, pharmaceutical companies are starting to rely on outsourcing production as a way to keep prices lower, often from foreign suppliers. When products come from a foreign supplier, the government is not able to maintain the same regulatory control over the manufacturing of these products, so there are growing quality-related risks. Pharmaceuticals of subpar quality can lead to serious illness or even death in consumers (Nagurney et al., 2013). For example, the article by Payne (2008) studied an incident in which contaminated heparin made by a Chinese manufacturer led to many adverse reactions and deaths among users. Although outsourced firms pose these quality-related risks, pharmaceutical companies continue to rely on them as consumers increase the pressure to maintain or reduce prices of drugs.

The paper by Nagurney, Li, and Nagurney (2013) uses a game theory model to study the selection of outsourced contractors and the determination of the optimal product flow using price and quality competition. A numerical example shows that when there is no pressure for improvements in quality, prices are higher for lower quality products and as demand for a product increases, the quality of that product tends to decrease.

In a PSC, it is very important to maintain a customer service level (CSL, for example, fill rate) of as close to 100% as possible, since failure to do so has a negative impact on the health and safety of patients. The paper by Uthayakumar and Priyan (2013) develops a method for minimizing cost by finding optimal solutions for inventory, lot size, lead time, and the number of deliveries from the pharmaceutical company to the hospital in order to achieve the hospital's desired CSL. In order to achieve this target, many companies will keep a large inventory on hand, which is not always cost effective with perishable products. With perishable products, which are common in PSCs, large inventories that may be in stock for a long period of time are not an ideal strategy. The paper uses a numerical example to demonstrate the validity of the model as a decision support tool for operations, health policies, and PSC strategies. The model can be used to achieve the target CSL at a minimum cost or to maintain inventory without overstocking.

2.3.2 Product Flow in the PSC

In a PSC, products typically go through two manufacturing sites for full production. The primary manufacturing site produces the active ingredient, which is used in both branded and generic products. The secondary manufacturing site completes production, adding inactive ingredients, as well as any further processing and packaging for the specific brand (Shah, 2004). Completed products are then distributed to wholesalers, who then distribute products to pharmacies, hospitals, and physicians for final distribution to patients. Wholesale distributors may also set up drug buy-back programs that allow pharmacies, hospitals, and physicians to return expired or excess product to the manufacturer (Kaiser Family Foundation, 2005).

The paper by Amaro and Barbosa-Póvoa (2008) studies the planning and scheduling of industrial supply chains with reverse flows using a mixed-integer linear programming (MILP) model. The planning portion of the model involves macro-entities (equipment, warehouses, vehicles, etc.) and describes the grouping of different events in the supply chain over an extended period of time. The results of the planning portion are then used to solve the scheduling problem, which provides a more detailed representation of the supply chain events for a shorter time frame. The planning and scheduling model is then used in a case study of a pharmaceutical supply chain (PSC). The planning model is used for a period of 3 months with a time interval of one week, with scheduling run for a 5-day week with a time interval of two hours. The results of the study show that when product recovery (returns of unused product) is considered, profits are at their highest, due in part to the ability to remanufacture or recycle products instead of transporting them elsewhere to be destroyed.

The paper by Sousa et al. (2011) uses an MILP model to make a series of decisions in a PSC. These decisions include the allocation of primary and secondary manufacturing, management of available resources during the time horizon, production amounts and inventory levels for each manufacturing site, and the establishment of product flows between sites. The MILP model aims to maximize the net profit of the company, which is the total revenue minus the costs of production, transportation, inventory handling, products allocation, unmet demand, and taxes.

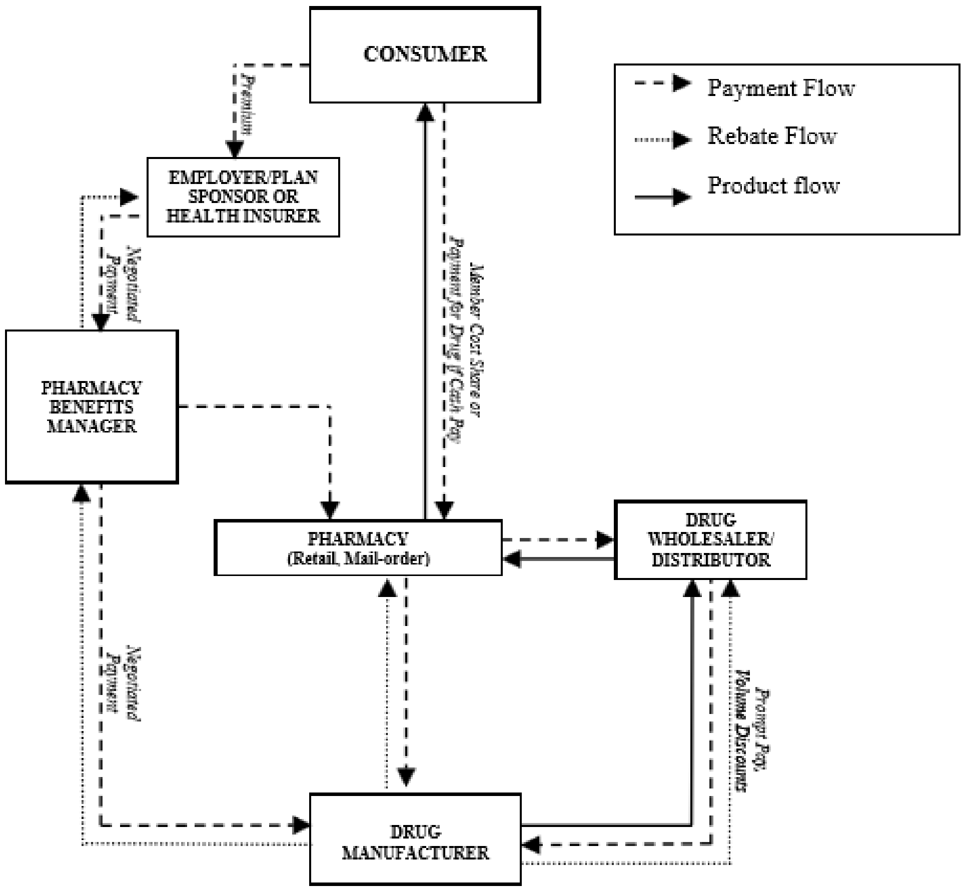

Figure 11: Flow of Goods and Payment in the PSC (Kaiser Family Foundation, 2005)

2.3.3 Payment Flow in the PSC

As shown in Figure 11, the flow of payment in a PSC is much more complex than the flow of payment in a standard supply chain. Like any other supply chain, the manufacturer has the greatest influence over the price of a pharmaceutical product. Manufacturers study the expected demand of a new product, assess the future competition, and project costs associated with marketing a new product to set the wholesale acquisition cost (WAC). The WAC is the lowest price at which a wholesaler can purchase the product. The final cost to the wholesaler will take into account discounts and rebates, the volume of the order, the promptness of payment, and the market share of the wholesaler. Pharmaceutical wholesalers then sell products to pharmacies and hospitals at the WAC plus a negotiated percentage. As of 2005, the top three wholesalers account for nearly 90% of the market. In 2015, the top three wholesalers accounted for $378.8 billion in revenue, which is approximately 85% of all revenues from drug distribution (MDM Market Leaders, 2016). These wholesalers have some influence over the pricing of generic products, which largely depends on the ability of the wholesaler to increase sales for a product. For branded products, however, pricing is typically left to the control of the manufacturer (Kaiser Family Foundation, 2005).

At the other end of the PSC, the end consumer (the patient) pays a premium to their health insurance provider, who will then cover some or all of the cost of a product, with any remaining costs paid by the consumer. The paper by Rossetti, Handfield, and Dooley (2011) studies the PSC and the changes it is going through as healthcare policies change. For example, in the U.S., the reimbursement policies for Medicare and Medicaid have altered the compensation for different members of the PSC and newer members of the PSC, such as third-party logistics providers, are taking on roles typically held by wholesalers. Another change in the flow of payment in a PSC is that pharmacy benefit managers (PBMs) are now processing nearly two-thirds of prescriptions in the U.S. PBMs achieve savings for their customers through the negotiation of discounts with manufacturers or wholesalers (Kaiser Family Foundation, 2005). The paper by Kouvelis, Xiao, and Yang (2015) models the competition between PBMs using an equilibrium analysis model. A PBM is an intermediary between pharmaceutical manufacturers and their clients, which includes employers, insurers, and programs such as Medicaid. PBMs provide administrative services, such as the processing of prescriptions, they help to negotiate better wholesale prices with the manufacturers, and they also set up the copay amounts for different drugs. The model uses flat rate rebates and pre-negotiated wholesale prices for each PBM to model the competition between PBMs when a client is choosing a provided PBM plan and the specific drug for the enrollee (branded, preferred branded, or generic).

3 Demand and Network Analysis

3.1 Data

For this analysis, demand data for a period of 6 ½ months for 43,000 products from a large pharmaceutical distributor are used. The demand data shows the quantity of a product that was shipped from one of the distribution centers to a specific customer on each day between September 2015 and mid-March 2016. Other information available for the analysis includes details about the individual products, such as their unit cost, total size (in cubic inches), the storage type required, and other identifiers for various drug types.

3.2 Demand Categorization

The first step in the demand and network analysis is to categorize the demand. For this thesis, demand was categorized using the Supply Chain Guru® software from Llamasoft©. Supply Chain Guru® is software that allows users to simulate and optimize the various aspects of a supply chain, including network layout, product flow, transportation, and acquisitions. Demand is categorized based on the following parameters:

- Non-zero demand mean
- Non-zero demand standard deviation
- Demand interval mean
- Coefficient of variation of non-zero demand

The breakpoints for each category are preset in the software and the categories and their descriptions are found in Table 1.

Category	Description
Smooth	Non-intermittent, low variable demand
Erratic	Non-intermittent, variable demand
Slow – Low Variable	Intermittent, low variable demand
Slow – Highly Variable	Intermittent, highly variable demand
Lumpy	Intermittent, variable demand (more variable than slow)
Extremely Variable	High coefficient of variation
Extremely Slow	Very large inter-demand interval mean

Table 1: Demand Categories as defined in the Supply Chain Guru® documentation

For the analysis, demand data was analyzed for all of the products with demand in the given timeframe at each of the distribution centers in the network. The analysis assumes a 6-day work week and the demand has a daily aggregation period. At the end of the analysis for each distribution center, each product has values for the parameters as shown in Figure 12.

Product Name	Total Demand Qty	NonZero Demand Qty Mean	NonZero Demand Qty StdDev	Demand Qty Mean	Demand Qty StdDev	Demand Interval Mean	COV	Intermittency	Variability	DemandClass	Demand Class Details
1	2	1.000000	NULL	0.007722	0.087704	86.000000	NULL	NULL	NULL	ExtremelySlow	ExtremelySlow
2	8	1.333333	0.516398	0.030888	0.213434	36.857143	0.150000	NULL	NULL	ExtremelyVariable	ExtremelyVariable
3	12	2.400000	1.140175	0.046332	0.360039	43.000000	0.225694	NULL	NULL	ExtremelyVariable	ExtremelyVariable
4	81	2.076923	1.402716	0.312741	0.918517	6.450000	0.456140	Intermittent	LowVariable	Slow	Slow-LowVariable
5	15	1.363636	0.504525	0.057915	0.292882	21.500000	0.136889	NULL	NULL	ExtremelyVariable	ExtremelyVariable
6	9	1.125000	0.353553	0.034749	0.203528	28.666667	0.098765	NULL	NULL	ExtremelyVariable	ExtremelyVariable
7	11	1.375000	1.060660	0.042471	0.295528	28.666667	0.595041	NULL	NULL	ExtremelyVariable	ExtremelyVariable
8	18	2.250000	1.669046	0.069498	0.477189	28.666667	0.550265	NULL	NULL	ExtremelyVariable	ExtremelyVariable
9	33	1.434783	0.662371	0.127413	0.452364	10.750000	0.213123	Intermittent	LowVariable	Slow	Slow-LowVariable
10	64	1.777778	1.806514	0.247104	0.906876	6.972973	1.032589	Intermittent	LowVariable	Lumpy	Lumpy

Figure 12: Sample output from Demand Analysis

Following the analysis in Supply Chain Guru®, the values for the means (non-zero demand, all demand, and demand interval) and standard deviations (non-zero demand and all demand) were calculated using a SQL query for verification purposes.

3.3 SQL Demand Analysis

Once the demand patterns for each product in each distribution center were categorized, further analysis was completed in SQL to obtain a deeper level of understanding of the demand. In this second analysis, the expected demand, average demand interval, inventory target, and average inventory for each item are calculated.

3.3.1 Expected Demand

The expected demand for a product on any given day can be calculated using the expected demand given that demand is non-zero. The expected demand is as follows:

$$E[D_t] = E[D_t|D_t > 0] \Pr\{D_t > 0\} \tag{1}$$

where $\Pr\{D_t > 0\}$ is the probability that demand is greater than zero.

The first term in (1), or expected non-zero demand quantity from Figure 12, is estimated from the sample average of the non-zero demands in the time frame used for the analysis as follows:

$$E[D_t\widehat{|D_t} > 0] = \frac{\sum_{t=StartDate}^{EndDate} D_t}{NumDemand} \tag{2}$$

In (2), $NumDemand$ is the number of days with non-zero demand, which can be estimated from the data as follows:

$$N[\widehat{D > } 0] = \sum_{t=StartDate}^{EndDate} NonzeroD_t \tag{3}$$

where $NonzeroD_t$ is a binary variable for day t as follows:

$$NonzeroD_t = \begin{cases} 1 \; if \; D_t > 0 \\ 0 \; if \; D_t = 0 \end{cases} \tag{4}$$

Thus, the expected value for non-zero demands is as follows:

$$E[D_t\widehat{|D_t} > 0] = \frac{\sum_{t=StartDate}^{EndDate} D_t}{\sum_{t=StartDate}^{EndDate} NonzeroD_t} \tag{5}$$

The second term in (1) is also estimated from the shipment data. First, the number of days in the time frame where demand is possible must be calculated. In this case, there are no Sunday deliveries, so shipments are only possible 6 out of 7 days of the week, so $NumShipDays$ can be estimated as follows:

$$Num\widehat{Ship}Days = \left\lfloor \left({}^{6}/_{7}\right) Days[EndDate, StartDate] \right\rfloor \tag{6}$$

where $Days[EndDate, StartDate]$ gives the number of days between $StartDate$ and $EndDate$, to include both the first and last days. The estimate is then rounded down to the nearest integer. Once this value is obtained, $\widehat{\Pr\{D_t > 0\}}$ can be estimated from the data as follows:

$$\widehat{\Pr\{D_t > 0\}} = \frac{N\widehat{[D > 0]}}{Num\widehat{ShipDays}} = \frac{\sum_{t=StartDate}^{EndDate} NonzeroD_t}{\left\lfloor \left(\frac{6}{7}\right) Days[EndDate, StartDate] \right\rfloor} \tag{7}$$

Using (1), (5), and (7), the expected demand is calculated as follows:

$$E[D_t] = \frac{\sum_{t=StartDate}^{EndDate} D_t}{\sum_{t=StartDate}^{EndDate} NonzeroD_t} * \frac{\sum_{t=StartDate}^{EndDate} NonzeroD_t}{\left\lfloor \left(\frac{6}{7}\right) Days[EndDate, StartDate] \right\rfloor} \tag{8}$$

3.3.2 Average Demand Interval

The average demand interval (ADI) is an important factor in the analysis of demand for a product. It can be calculated as follows:

$$ADI = \frac{1}{\Pr\{D_t > 0\}} \tag{9}$$

For example, if $\Pr\{D_t > 0\} = 0.5$, then $ADI = \frac{1}{0.5} = 2$, which means that on average, non-zero demand occurs every two days.

3.3.3 Inventory Target and Average Inventory

In a single demand period t, inventory evolution between periods is described with the following equation:

$$I_t = I_{t-1} + O_t - D_t$$

At the beginning of period t, inventory is I_{t-1}, and an order of quantity O_t is placed to bring the inventory level to a target level. Thus the target inventory can be described as:

$$I^{target} = I_{t-1} + O_t$$

We assume that orders are always successful and delivered before demand occurs in period t. A demand of D_t then occurs, lowering the total inventory for period t to I_t. This dynamic then repeats each period. In the event of a shortage, I_t is negative, which causes a customer to either wait until the inventory of that product to be replenished or to cancel the order. Either of these options cause the company to lose money or customer goodwill, so the desired customer service level is set very high to lower the possibility of a shortage. For this paper, the company uses a goal of 99.2% customer service level, meaning that 99.2% of orders can be filled directly from on hand inventory. This leads to a greatly reduced risk of a shortage occurring. To set the inventory target, I^{target}, two different approaches are studied: percentile and fill rate.

3.3.3.1 Percentile Based Inventory Target

In the percentile approach, the inventory target, I^{target}, is the p^{th} percentile, $D^{p\%}$, of the distribution of non-zero demands. The p^{th} percentile is the value such that p% of the data is at or below this value and (100-p)% are above it. In the case of demand data, the p^{th} percentile is such that p% of the demands in the given time frame are at or below this value. For the analysis in this thesis, the 99.2 percentile is being used.

To calculate the 99.2 percentile of the demand for a given product at a single location, the demands are ranked in ascending order and the value that is in the 99.2% index location is the final target. This value is calculated in SQL, using:

```sql
SELECT DISTINCT SiteName as SiteName, ProductName as ProductName
,CEILING(PERCENTILE_CONT(0.992) WITHIN GROUP(ORDER BY Quantity)
OVER (PARTITION BY SiteName, ProductName))
```

In the query, the partitioning ensures that the percentile is calculated for each individual product at each of the 20 distribution centers. The inventory target does not need to be a value within the set of demands, so the calculation assumes a continuous distribution for the demand data (Quantity). The value for I^{target} must be an integer, so the value for the 99.2 percentile is rounded up to the nearest integer to ensure a service level of at least 99.2%.

Using this inventory target, the end of day inventory, I_t, can be calculated as follows:

$$I_t = I^{target} - D_t \tag{10}$$

Thus, the expected end of day inventory is as follows:

$$E[I_t] = I^{target} - E[D_t] \tag{11}$$

3.3.3.2 Fill Rate Based Inventory Target

In the fill rate approach, the inventory target is more complicated to obtain. The fill rate for a product is the fraction of all demands that are met from the inventory on hand (Hopp & Spearman, 2008). The fill rate for a particular inventory target can be computed from the following formula:

$$Fill\ Rate = 1 - \frac{\sum_{t=StartDate}^{EndDate} B_t}{\sum_{t=StartDate}^{EndDate} D_t}$$

Where B_t is the backorder in period t resulting from a particular inventory target and D_t is the demand in period t. This calculation assumes the same inventory target is used in each period. The inventory target based on Fill Rate is set by finding the minimum inventory level required to achieve the desired service level, which in this case is 99.2%. To obtain this target value, a search over possible values is needed and two different approaches were used to complete this computation.

3.3.3.3 Brute Force Search Approach

In the brute force search, the inventory target for a fill rate of 99.2% is found by looping through possible values for I^{target} until the desired service level is achieved. This process starts with the inventory target set equal to the maximum value for demand. This is a reasonable starting point since there are no larger demands, and therefore this provides a fill rate of 100%. The target is then decreased by one in each step. Since the 99.2 percentile of demand is likely to be "near" the maximum demand value in many cases this strategy may only have to search through a limited number of values. To accomplish this, the following steps are taken:

1. Initialize: Set the initial value for I^{target} and fill rate for each product at each location

   $$I^{target} = max[t \in [StartDate, EndDate]: D_t]$$

   Initialize Inventories and Backlogs for each t starting from period 1:

$$I_t = I^{target} - D_t$$

$$B_t = \begin{cases} 0 \; if \; I_t \geq 0 \\ -I_t \; if \; I_t < 0 \end{cases}$$

Initialize Fill Rate:

$$Fill\ Rate = 1 - \frac{\sum_{t=StartDate}^{EndDate} B_t}{\sum_{t=StartDate}^{EndDate} D_t}$$

2. Loop through possible targets from largest to smallest until the fill rate drops just below the desired level

$WHILE\ Fill\ Rate > 0.992$

$$I^{target} = I^{target} - 1$$

Update Inventories and Backlogs for each t starting from period 1:

$$I_t = I^{target} - D_t$$

$$B_t = \begin{cases} 0 \; if \; I_t \geq 0 \\ -I_t \; if \; I_t < 0 \end{cases}$$

Update Fill Rate:

$$Fill\ Rate = 1 - \frac{\sum_{t=StartDate}^{EndDate} B_t}{\sum_{t=StartDate}^{EndDate} D_t}$$

$End\ WHILE$

3. Finalize the targets and fill rate to ensure 99.2% fill rate:

$$I^{target} = I^{target} + 1$$

$$I_t = I^{target} - D_t$$

$$B_t = \begin{cases} 0 \; if \; I_t \geq 0 \\ -I_t \; if \; I_t < 0 \end{cases}$$

$$Fill\ Rate = 1 - \frac{\sum_{t=StartDate}^{EndDate} B_t}{\sum_{t=StartDate}^{EndDate} D_t}$$

The final step of this loop increases the target by 1, which then gives the minimum inventory level required to achieve the fill rate of 99.2%. This is guaranteed to give the minimum target that achieves the fill rate because the fill rate is non-decreasing in I^{target}. This approach is not very efficient computationally (it is $O(n)$). To complete the calculations more efficiently, a second approach was developed to decrease the amount of time spent looping through values for I^{target}.

3.3.3.4 Binary Search Approach

The brute force search approach worked well for products with smaller demand values, but for products with much larger and highly variable demands, it was a very time consuming process. Thus, a new approach, the binary search approach, was developed. For this approach, the following steps are taken for each product in each location:

1. Set an initial upper bound (ub) and lower bound (lb) for I^{target} and $Fill\ Rate$

   $$target_{lb} = 0, fillrate_{lb} = 0$$
   $$target_{ub} = MAX(Demand), fillrate_{ub} = 1$$

   In the above, $fillrate_{ub}$ is 1 because when the inventory target is the maximum demand, all demands are able to be filled from the inventory on hand. Similarly, $fillrate_{lb}$ is 0 when the inventory target is 0, because no demands are filled when there is no on-hand inventory.

2. Iteratively bring the I_{lb}^{target} and I_{ub}^{target} closer together by bisecting their difference until they differ by 1 or less.

   $$WHILE(I_{ub}^{target} - I_{lb}^{target} > 1)$$
   $$I_{current}^{target} = \left\lfloor \frac{I_{ub}^{target} + I_{lb}^{target}}{2} \right\rfloor$$
   $$I_{t,current} = I_{current}^{target} - D_t$$
   $$B_{t,current} = \begin{cases} 0 & if\ I_{t,current} \geq 0 \\ -I_{t,current} & if\ I_{t,current} < 0 \end{cases}$$
   $$fillrate_{current} = 1 - \frac{\sum_{t=StartDate}^{EndDate} B_{t,current}}{\sum_{t=StartDate}^{EndDate} D_t}$$
   $$IF(fillrate_{current} > 0.992)\ \{$$
   $$I_{ub}^{target} = I_{current}^{target}, fillrate_{ub} = fillrate_{current}\}$$
   $$ELSE$$
   $$I_{lb}^{target} = I_{current}^{target}, fillrate_{lb} = fillrate_{current}$$

$End\ IF$

$End\ WHILE$

This is implemented in SQL. Once the loop has completed for each item in each location, the $fillrate_{current}$ will be at least 99.2% and the final value for $I^{target} = I^{target}_{current}$. Using either method, the average end of period inventory can be calculated from the inventory targets and the demands using (11). The binary search is $O(\log n)$.

3.4 Aggregating Demand

The ultimate goal of this analysis is to evaluate the value of aggregating inventory into a hub-and-spoke network or a fully centralized network. To assess the current performance, the demand analysis was carried out on the demands from each individual distribution center to obtain inventory targets, period by period inventories and average inventory for each DC. To assess the performance of each of the proposed networks, first, the demands for the DCs were aggregated based on either the hub-and-spoke design, or the fully centralized design. Then the demand analysis was performed for the demands in the proposed new networks. The demand aggregation requires summing the demand for each period (daily) across all the DCs in a hub.

3.4.1 Network Options

For this analysis, three distribution networks are available for use: the original decentralized network, a hub-and-spoke network, and a fully centralized national network. In the decentralized network, the 20 individual distribution centers are responsible for fulfilling the demands for all of their customers. Figure 13 shows an example of this type of network, with 20 distribution centers. In Figure 13, the NLC is a national logistics center, where manufacturers make all of their deliveries. In Figure 13, demand D^i is the demand associated with DC^i in the original network. The demand patterns vary between distribution centers, allowing for high variability in demand for a product across the network.

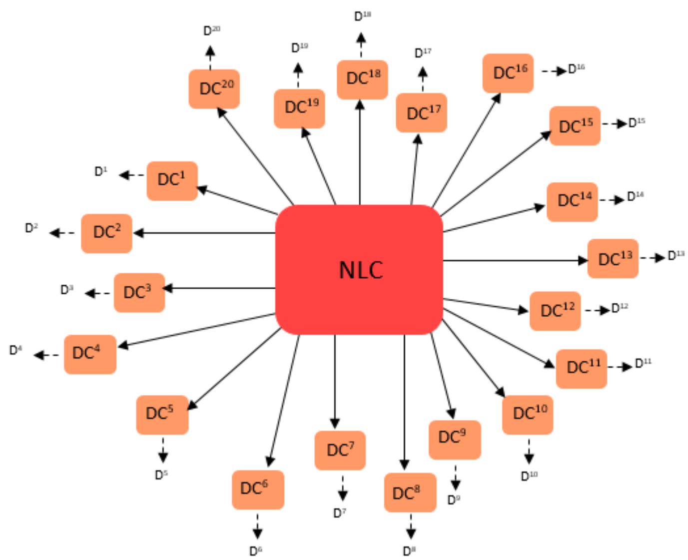

Figure 13: Decentralized Network Layout

In the hub-and-spoke network, the inventory from each distribution center is aggregated into one of three hubs, based on a predefined network from the company. Each of the hubs is an existing location within the network. Figure 14 shows the layout for this network. Each of the three hubs is responsible for fulfilling the demands from the customers for the individual distribution centers assigned to that hub. The demand patterns vary between the hubs.

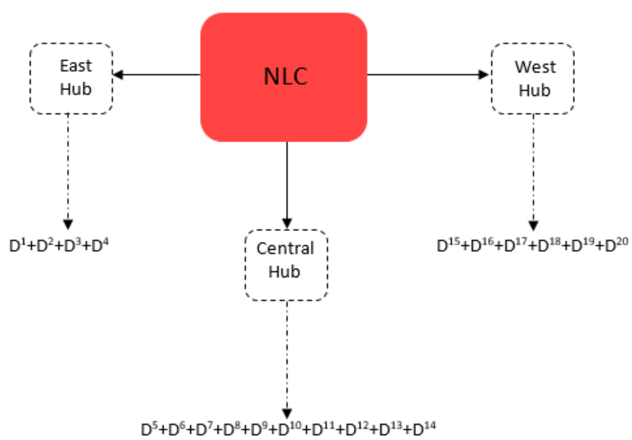

Figure 14: Hub and Spoke Network Layout

In the national network, the inventory from all of the individual distribution centers is aggregated into a single central location that is part of the existing network. Figure 15 shows how this network is laid out. The national hub is responsible for fulfilling all customer demand for the centralized products throughout the network.

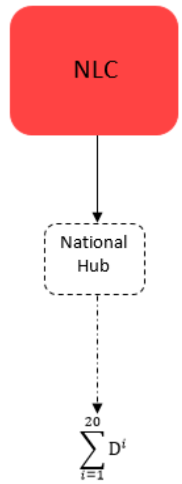

Figure 15: National Network Layout

Once the demand has been aggregated into the new networks, the analysis is carried out for each location in each network for mean non-zero demand, ADI, inventory target, and average inventory.

3.5 Comparing the Networks

Once the analysis has been completed for the decentralized, hub-and-spoke, and fully centralized networks, the results are compared to evaluate the value of using each option. To compare the decentralized network with the hub-and-spoke network, the target inventory, average inventory, and ADI are compared. For the average inventory, if distribution center *i* has an average inventory of $E[I_t^i]$, then the total average inventory across the distribution centers that are aggregated in Hub *j* is $\sum_{i \in Hub\ j} E[I_t^i]$. This is repeated for each Hub. It is more difficult to accurately represent the overall ADI when several distribution centers are aggregated into a hub. If the ADI at each of the distribution centers is averaged across the distribution centers included in a hub, then the demand from each distribution center is equally weighted. For a more fair representation of the overall ADI in the current network, a weighted average of the ADI from each distribution center is used. The weight is computed by the relative fraction of distribution center *i* demand to the total demand that will be aggregated in Hub *j*:

$$TotalD^{i} = \sum_{t=StartDate}^{EndDate} D_{t}^{i} \tag{12}$$

Using (12), the aggregated current ADI for the distribution centers being aggregated in Hub *j* are computed from the weighted sum:

$$ADI_{aggregate} = \frac{\sum_{i \in Hub\ j} TotalD^{i}\, ADI^{i}}{\sum_{i \in Hub\ j} TotalD^{i}}$$

As an example, the analysis was completed for a single product that is expensive (approximately $10,000 unit cost) but very small in size. The comparison of results for this item between the hub-and-spoke network (Figure 14) to the decentralized network (Figure 13) is shown in Figure 16. The ADI is given in days and the inventory targets and average inventory are in product units. In this example, it is clear that aggregating the demand for this product can save the company a reasonable amount of money just by utilizing the hub-and-spoke network. In the fill rate example, the average inventory is reduced by 67%, potentially saving the company approximately $80,000 annually in inventory holding costs on this single product. The ADI decreases from 2.67 days, 4.55 days, and 5.82 days in the individual locations to 1.27 days, 1.18 days, and 1.31 days, respectively, in the hubs. This decrease means that demand becomes more regular, with demands occurring almost every day in each hub instead of once or twice in a week at the original locations.

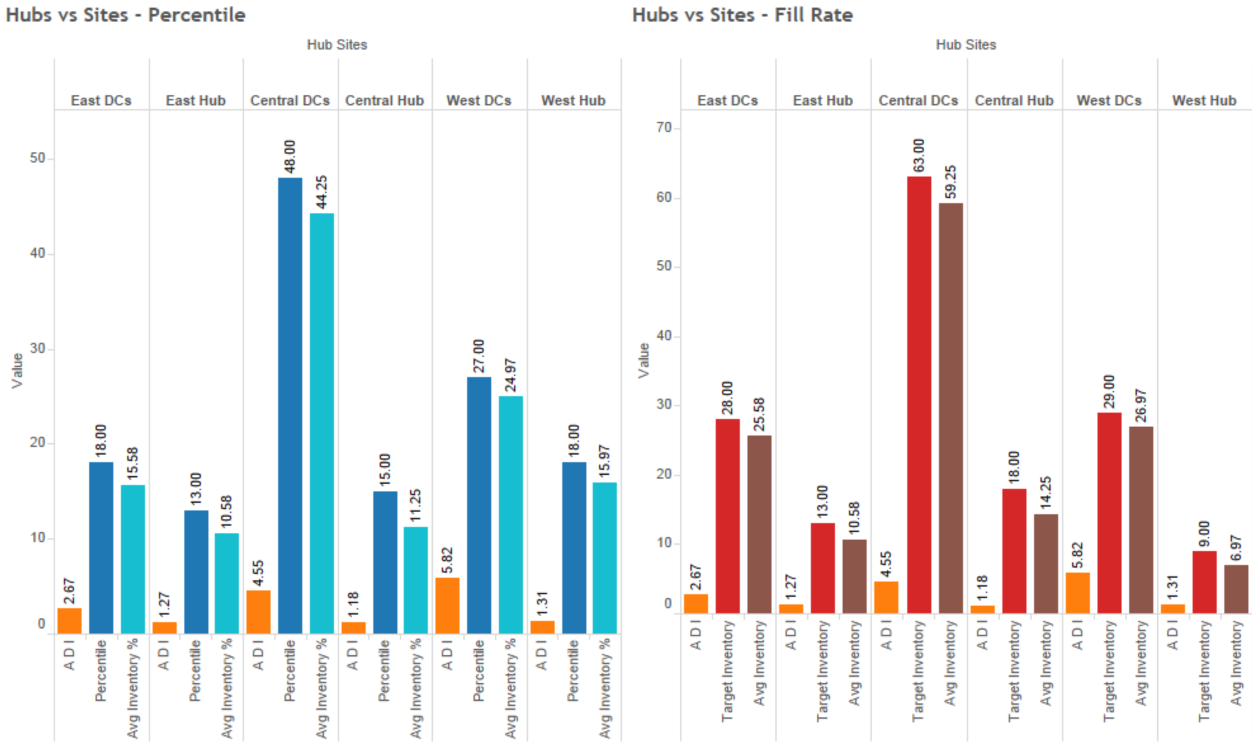

Figure 16: Comparison of the decentralized and hub-and-spoke networks for an item. The left graph shows results for the percentile based inventory targets. The right graph shows results for the fill-rate based inventory targets.

Figure 17 shows the comparison of each of the three network options for this example product. Once again, it is very clear that this product is worth considering for aggregating: In the fill rate analysis further aggregating the demand into a single national hub reduces total average inventory by about 100 units, which reduces inventory investment by $1,000,000 and saves the company about $100,000 in annual inventory holding costs (at a 10% annual holding cost rate). In both cases (national and hub-and-spoke), the ADI is reduced from around 4 (a non-zero demand approximately every 4 days) to about 1, (a non-zero demand almost every day). This leads to a more regular and predictable process.

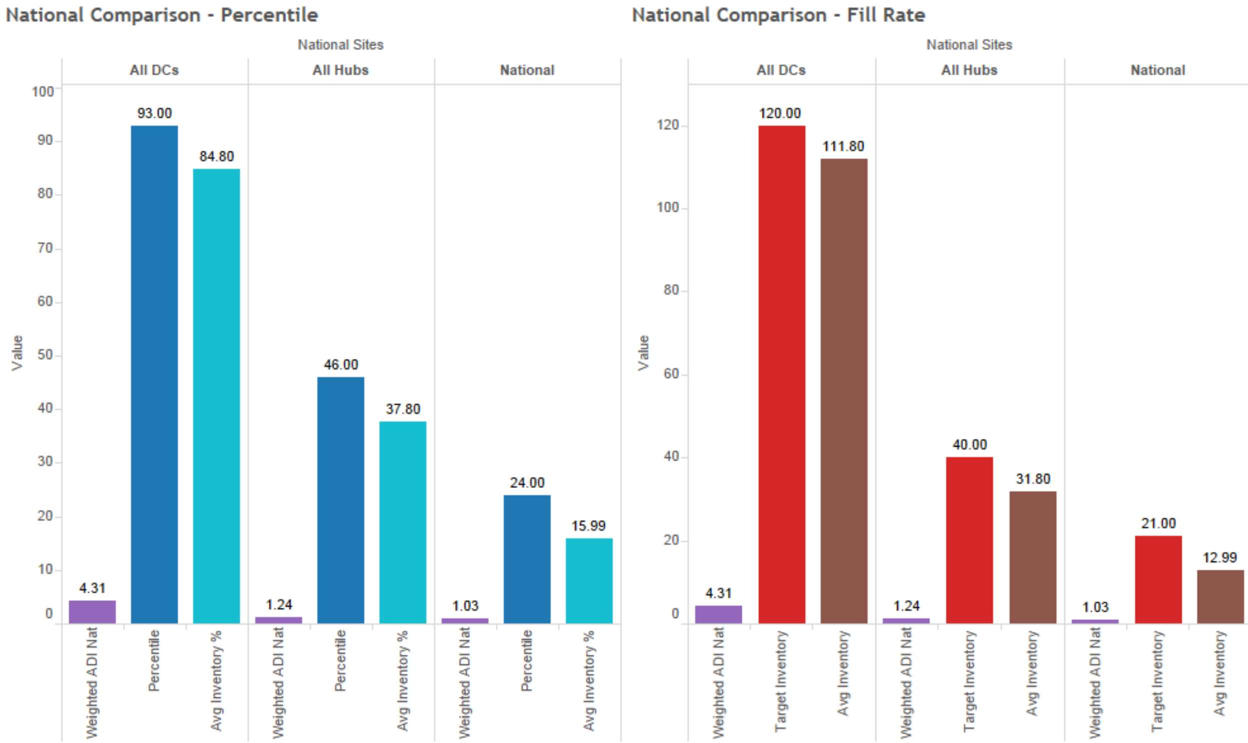

Figure 17: Comparison of all three network options

With the positive outcome from evaluating a single item, the inventory analysis was completed for each product in the network. From this point forward, only results using the fill rate approach for the inventory target will be presented, as this is the company's preferred method for setting the inventory target.

3.6 Evaluating Transportation Costs

When demand is aggregated into hubs instead of being fulfilled from all individual DCs, total average inventory may be reduced, but there are now extra costs associated with transporting products from the hubs to the customers. Because these shipments are outside of the normal shipping routes handled by the company's dedicated network, it is necessary to take extra measures to ensure timely deliveries and maintain current service standards. To do this, it is assumed that all demand for products that have been aggregated will be shipped via an overnight shipment, which is a relatively expensive but quick option available to the company outside of the normal shipping routes. We assume that products will not be shipped together with other products, even if both the origin and final destination are the same. Using this option for the analysis gives the 'worst case' for the transportation costs, so the actual costs may be quite a bit less based on the shipping options chosen by the customer or the ability to combine shipments for a single customer. To estimate the costs of shipping products from the hubs using overnight service, the following steps are taken:

1. Calculate the shipment cube (size in cubic inches) from the quantity being shipped and the size of the individual product
2. Determine the number and size of boxes required for a shipment, based on available box sizes and product cube
3. Calculate dimensional factor for the shipment to determine dimensional weight
4. Determine the number of zip code zones traveled
5. Determine the cost of the shipment, $C_{k,p,loc,t}^{shipment}$, based on the source and destination zip code zones and dimensional weight (for customer *k*, product *p*, in location *loc*, in period *t*). These costs were determined using an overnight shipping cost rating tool.

The shipping cost for a location is determined for each product by considering the individual customer demands within a day that make up the total demand in that day. Thus there will be a shipment from each hub in a network to individual customers with orders in that day, for each day in the analysis time frame of the demand data.

3.7 Calculating Net Savings

Once the average inventory and transportation costs have been calculated for all products in all sites, the final step in evaluating the hub-and-spoke and fully centralized networks is to calculate the net savings. The first step is to calculate the inventory holding costs for all three networks. In this calculation we use the unit value of each product V_p, and the company's given value of a 10% annual holding cost:

$$C_{network,p}^{Holding} = 0.1V_p * \sum_{loc \in network} \sum_{t=StartDate}^{EndDate} E\left[I_{loc,p,t}\right]$$

This is computed for three networks,

(1) decentralized distribution centers (subscript Sites as shown in Figure 13)

(2) three hubs (subscript Hubs as shown in Figure 14)

(3) the national network (subscript Nat as shown in Figure 15)

Once these are calculated, the total inventory savings from the decentralized network are calculated for the hub-and-spoke network for each product *p*:

$$InventorySavings_p^{Hubs} = C_{Sites,p}^{Holding} - C_{Hubs,p}^{Holding}$$

and the fully centralized network:

$$InventorySavings_p^{Nat} = C_{Sites,p}^{Holding} - C_{Nat,p}^{Holding}$$

The next step is to calculate the total transportation costs for the year. First, transportation costs $C_{loc,p,t}^{Trans}$ for day *t* are calculated for each location in the given network and for product *p*. Since the demand for product *p* at a location in day *t* includes shipments to multiple customers, $C_{loc,p,t}^{Trans}$ requires summing $C_{k,p,loc,t}^{shipment}$ for shipment *k* over all of these shipments:

$$C_{loc,p,t}^{Trans} = \sum_{k \in shipments} C_{k,loc,p,t}^{Shipment}$$

Since the demand data does not span the entire year, transportation costs must be annualized for a fair comparison with annual inventory costs. The annualized transportation costs for product *p* for the hub-and-spoke network are as follows:

$$C_{Hubs,p}^{Trans} = \frac{365}{Days[EndDate, StartDate]} * \sum_{loc \in Hubs} \sum_{t=StartDate}^{EndDate} C_{loc,p,t}^{Trans}$$

For the fully centralized National network, the annualized transportation costs are:

$$C_{Nat,p}^{Trans} = \frac{365}{Days[EndDate, StartDate]} * \sum_{t=StartDate}^{EndDate} C_{Nat,p,t}^{Trans}$$

With all of the components now calculated, the net savings for product p in the hub-and-spoke network are:

$$NetSavings_p^{Hubs} = InventorySavings_p^{Hubs} - C_{Hubs,p}^{Trans}$$
$$= C_{Sites,p}^{Holding} - C_{Hubs,p}^{Holding} - C_{Hubs,p}^{Trans} \tag{13}$$

and the fully centralized national network:

$$NetSavings_p^{Nat} = InventorySavings_p^{Nat} - C_{Nat,p}^{Trans}$$
$$= C_{Sites,p}^{Holding} - C_{Nat,p}^{Holding} - C_{Nat,p}^{Trans} \tag{14}$$

3.8 Analysis Results

Once all steps in the analysis have been completed and both (13) and (14) have been calculated, the results are reviewed to begin the process of product selection. Based on the analysis of inventory savings portion, a large number of products were candidates for centralization into either the hub-and-spoke or fully centralized network. For the hub-and-spoke network, 38,928 products showed positive inventory savings, while 40,183 showed positive inventory savings for the fully centralized network. For the Net Savings, including the transportation costs, the number of products to be considered for both networks is substantially reduced.

	National Network	Hub and Spoke Network
Number of items with $ Savings > 0	1,171	630
Net $ Savings	$18,126,000	$11,554,000
Total Transportation Costs	$10,180,000	$7,228,000
Inventory $ Reduction in Sites	$28,307,000	$18,782,000
Net Cube Savings	8,648,000	5,164,000
Total Cube in Hub	3,685,000	2,333,000
Total Cube Reduction in Sites	12,333,000	7,498,000

Table 2: Estimated annualized savings for All Products with Savings > 0 in the National and Hub and Spoke Networks

Table 2 shows an overview of the results from the analysis, including total savings, total transportation costs, and cube (space) savings. The number of products with positive Net Savings (>$0) for the hub-and-spoke network was 630 and for the fully centralized network was 1,171 products. The products eliminated from the pool of potential products for centralization typically either had only a small potential in inventory savings or were inexpensive items that are large and thus expensive to ship. Figure 18 shows a comparison of the potential savings using the hub-and-spoke network and the fully centralized network for all products with positive savings. The 630 products in the hub-and-spoke network show a potential for over $11 million in savings, while the 1,171 products in the fully centralized network show a potential for over $18 million in savings. These values do not account for the initial costs of centralizing the inventory of the products involved.

Since it is not feasible to centralize all of the products using either network, a smaller subset of products is also examined. Figure 19 shows a comparison of the potential savings using only the top 100 items in both network options. For the hub-and-spoke network, the company can potentially save over $10 million by centralizing the inventory of just 100 items. For the fully centralized network, the company can potentially save over $14 million by centralizing the inventory of just 100 items.

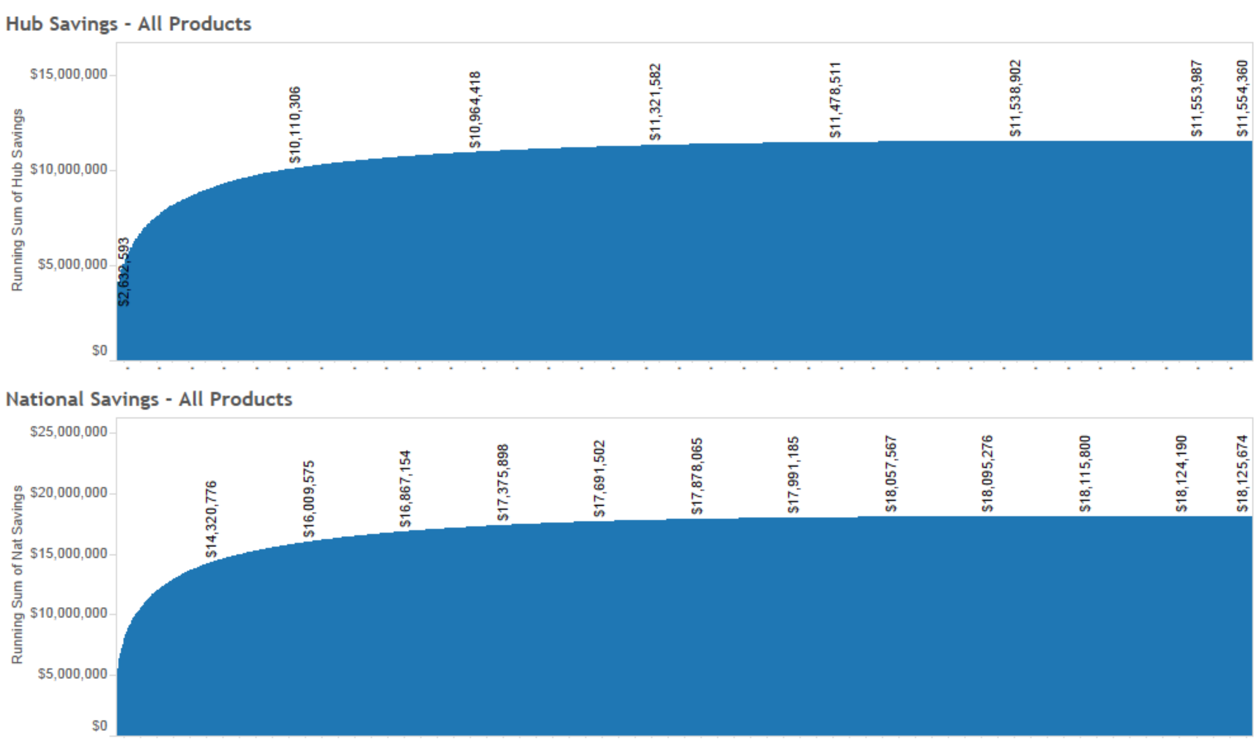

Figure 18: Comparison of results for all products with positive savings, ranked on the x-axis from the largest net savings to the smallest savings.

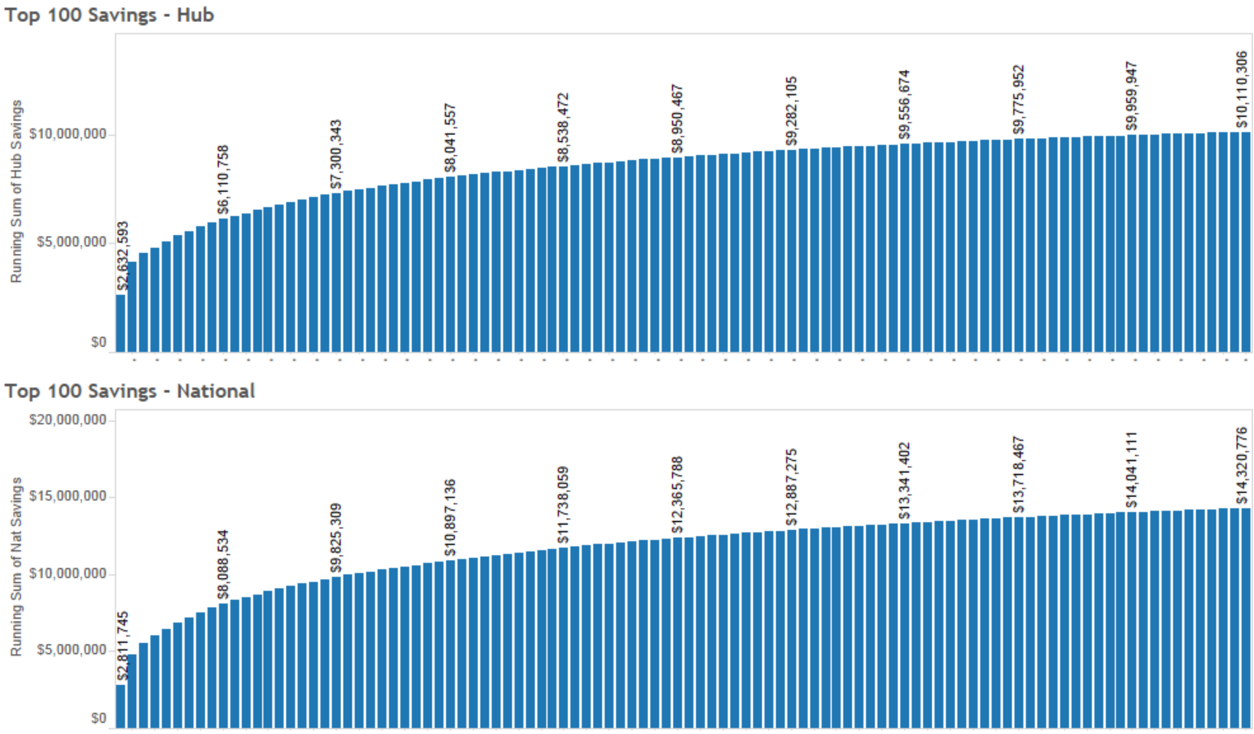

Figure 19: Comparison of potential savings for the top 100 items in both networks, ranked on the x-axis from the largest net savings to the smallest savings

Figure 20 shows the breakdown of the cost, cube, and demand categories for the top 100 items in the hub-and-spoke network. The x-axis for the Product Cost graph is displayed in thousands of dollars, with products grouped in bins of size $5,000. The x-axis for the Product Cube graph is displayed in cubic inches, with products grouped into bins of size 25. Approximately 80% of the items have a cost of $15,000 or less, with only about 35% with a cost less than $5,000, and almost all of the products are under 200 cubic inches in size. In general, it seems that small, expensive items tend to be good candidates for centralization. The third graph shows the percentage of distribution centers in which the demand for the products in the top 100 is categorized as either 'lumpy' or 'slow', which are the two categories that are more logical to centralize. The x-axis for this graph shows the fraction of locations in which a product was classified as 'lumpy' or 'slow', divided by the total number of distribution centers with demand for the product. About 90% of the products with savings in the top 100 were categorized as 'lumpy' or 'slow' in at least half of the distribution centers in which they had demand during the analysis time frame. Section 4 describes a method for choosing the best candidates for centralization in both networks.

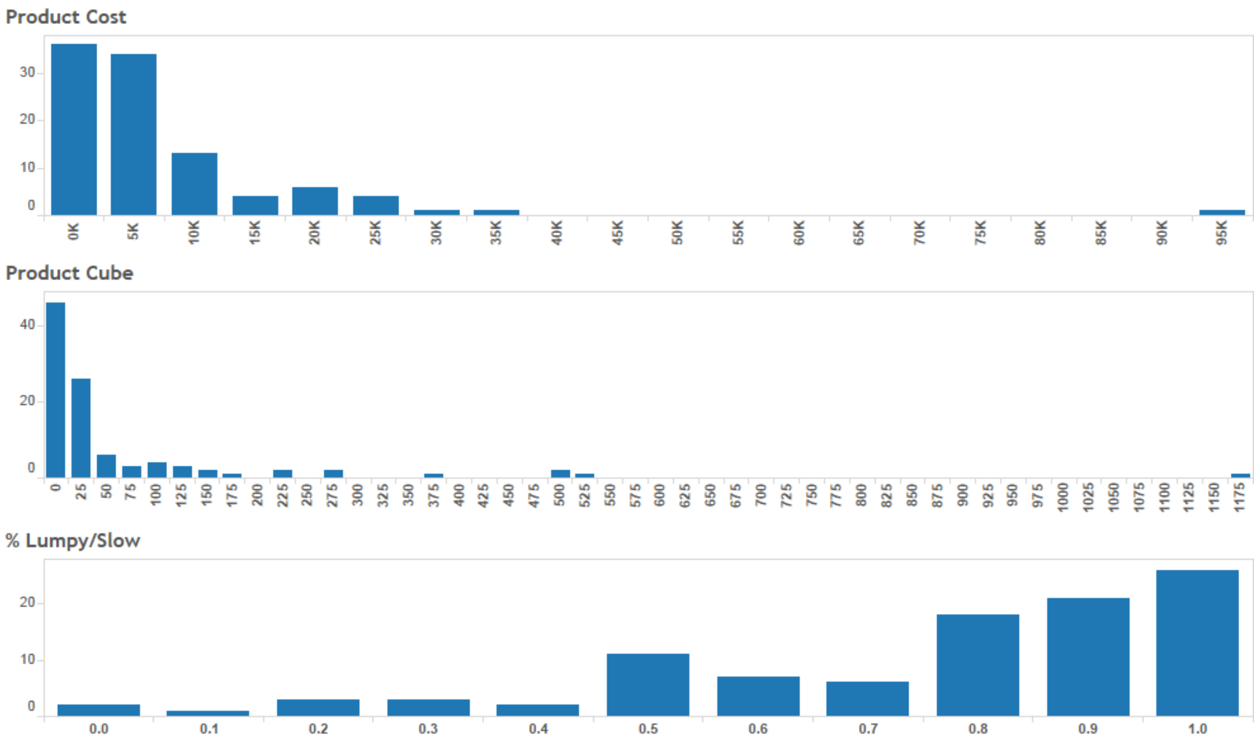

Figure 20: Breakdown of top 100 products for the hub-and-spoke network

4 Product Selection

Once the inventory and network analysis have been completed, the next step is to select the products to be centralized in either the hub-and-spoke or fully centralized network. Since it is not feasible to centralize all of the items in the network, products must be carefully selected in order to meet the company's goals for the centralization effort. This section details the product selection model recommended for use by the company.

4.1 Product Selection Model

The goal of the effort to centralize inventory into a hub-and-spoke or fully centralized network is to save the company money on inventory holding costs. In the previous section, estimated net savings for centralizing products into each network are calculated using (13) and (14). The final objective for the product selection model is to maximize the potential savings while centralizing products to either the hub-and-spoke or fully centralized network. A product can only be centralized into one of the two network options, otherwise it will remain in the current network. To solve the product selection problem, a 0-1 integer programming model is developed below.

4.1.1 Objective

The overall objective of the model is to maximize the potential savings with the use of one of two centralized networks for product inventories:

$$\sum_{i \in all\ products} NetSavings_i^{Hubs} * x_i^{Hubs} + \sum_{i \in all\ products} NetSavings_i^{Nat} * x_i^{Nat} \tag{15}$$

where x_i^{Hubs} and x_i^{Nat} are binary decision variables describing whether product *i* is to be centralized in the associated network:

$$x_i^{Hubs} = \begin{cases} 1 & if \text{ product } i\ is\ centralized\ in\ hub/spoke\ network \\ 0 & otherwise \end{cases}$$

$$x_i^{Nat} = \begin{cases} 1 & if \text{ product } i\ is\ centralized\ in\ national\ network \\ 0 & otherwise \end{cases} \tag{16}$$

The value $NetSavings_i^{Network}$ refers to the net savings for centralizing a product *i* into either the hub-and-spoke or national network, as calculated in Section 3.7. For this model, only the products for which net savings is positive are included for consideration for centralization into one of the new networks. This significantly cuts down the number of products being considered, as seen in Section 3.8.

4.1.2 Constraints

In selecting products for centralization, there are several factors that must be taken into consideration. A product can be centralized in no more than one of the available networks, so the following is applied:

$$x_i^{Hubs} + x_i^{Nat} \leq 1$$

(17)

Centralizing the inventory of a product means that the company will face additional transportation costs throughout the year. These costs are incurred because transportation from the centralized networks will be outside of the company's existing dedicated ground transport network. These additional costs come from the immediate cash flow, so they must be accounted for and are limited by a budget set by the company:

$$\sum_{i \in all\ products} C_{Hubs,i}^{Trans} * x_i^{Hubs} + \sum_{i \in all\ products} C_{Nat,i}^{Trans} * x_i^{Nat} \leq Trans\ Budget$$

(18)

where $i \in all\ products$ is the set of all products with positive net savings in the network being evaluated. The hubs being considered in the centralized networks are currently nodes (distribution centers) in the regular network, so the amount of space available for centralized inventory is limited. Thus, the space required for centralized inventory must also be accounted for. For each product, the amount of space required at the new location is determined by the maximum, or target, inventory at that hub. For the fully centralized network, the following is considered:

$$\sum_{i \in all\ products} Cube_i * I_i^{target,Nat} * x_i^{Nat} \leq National\ Cube\ Limit$$

(19)

where $Cube_i$ is the size of the product in cubic inches.

For the hub-and-spoke network, the space available in each of the three hubs in the network must be considered individually rather than as an overall space limitation for the network. The following constraint must be met at each hub location, indexed by $j = 1, 2, 3$:

$$\sum_{i \in all\ products} Cube_i * I_{i,j}^{target,Hubs} * x_i^{Hubs} \leq Hub_j\ Cube\ Limit$$

(20)

Depending on the setup of the networks, one of the hubs in the hub-and-spoke network may also be the national hub for the fully centralized network, as it is for the company in question. In this case, the space constraint for that particular location, Hub_j, is as follows:

$$\sum_{i \in allproducts} Cube_i * I_{i,j}^{target,Hubs} * x_i^{Hubs} + \sum_{i \in all\ products} Cube_i * I_i^{target,Nat} * x_i^{Nat} \leq Hub_j\ Cube\ Limit \tag{21}$$

4.2 Optimization Tool

To select products for centralization using the 0-1 integer program model described above, a prototype tool was designed using Solver for Microsoft Excel. This tool was chosen to demonstrate this function because it is readily available for the company to use, it is simple to put together the model, and it is very easy for the company to transfer the raw data from the SQL Server to the Excel document. A selection of sample data needed for the product selection model is loaded into the Excel tool, as shown in Figure 21. In this table, the TransCostHub and TransCostNat columns are the total annual cost for transportation of that product, or $C_{Hubs,i}^{Trans}$ and $C_{Nat,i}^{Trans}$. The HubSavings and NatSavings columns correspond to $NetSavings_i^{Hubs}$ and $NetSavings_i^{Nat}$, while ProductCube corresponds to $Cube_i$ in the product selection model. The values for the final four columns are the inventory targets for the product in each of the hubs, or $I_{i,j}^{target,Hubs}$ and $I_i^{target,Nat}$ in the model.

ProductName	TransCostHub	TransCostNat	HubSavings	NatSavings	ProductCube	HubETarg	HubCTarg	HubWTarg	NatTarg
1	$781,725.99	$667,555.79	$2,632,592.56	$2,811,744.75	3797.1	441	940	480	1671
2	$245,903.14	$217,432.85	$1,492,896.90	$2,011,528.70	31500	174	426	213	648
3	$121,077.50	$105,739.56	$408,122.54	$497,725.10	28000	38	114	57	180
4	$30,780.96	$31,588.52	$275,933.05	$383,034.63	6014	55	778	227	875
5	$179,625.54	$161,342.14	$271,068.07	$675,113.46	5404	191	883	659	1011
6	$336,139.27	$288,869.82	$258,940.27	$400,322.95	3797.1	93	494	302	635
7	$56,702.24	$51,863.07	$210,944.22	$251,431.38	2629.14	125	383	191	562
8	$51,687.83	$50,873.14	$206,601.10	$305,421.25	27773	17	65	15	61
9	$576,790.12	$492,412.38	$204,904.19	$434,620.48	3728.64	183	1013	502	1295
10	$13,614.87	$12,310.87	$148,754.73	$183,461.56	26188.64	17	11	7	22
11	$43,550.29	$39,728.76	$98,027.78	$290,913.92	16854.53	20	147	141	195
12	$87,859.76	$79,445.69	$74,620.65	$278,108.34	6422.15	46	342	426	507

Figure 21: Sample data for the Product Selection Model

Once the data is loaded, the solver tool is updated to accommodate the number of products included for consideration, in this case there are 12 items being considered in the sample data set. The solver tool for the sample data is shown in Figure 22.

DECISION VARIABLES				SUMS FOR CONSTRAINTS					
Product	Hub	Nat	Sum	Product Savings	Trans Costs	Cube Nat	Cube Hub E	Cube Hub C	Cube Hub W
1	0	0	0	$0.00	$0.00	0	0	0	0
2	0	0	0	$0.00	$0.00	0	0	0	0
3	0	0	0	$0.00	$0.00	0	0	0	0
4	0	0	0	$0.00	$0.00	0	0	0	0
5	0	0	0	$0.00	$0.00	0	0	0	0
6	0	0	0	$0.00	$0.00	0	0	0	0
7	0	0	0	$0.00	$0.00	0	0	0	0
8	0	0	0	$0.00	$0.00	0	0	0	0
9	0	0	0	$0.00	$0.00	0	0	0	0
10	0	0	0	$0.00	$0.00	0	0	0	0
11	0	0	0	$0.00	$0.00	0	0	0	0
12	0	0	0	$0.00	$0.00	0	0	0	0
13	0	0	0	$0.00	$0.00	0	0	0	0
14	0	0	0	$0.00	$0.00	0	0	0	0

OBJECTIVE	CONSTRAINTS
Total Savings	Trans Budget
$0.00	$3,500,000.00
National Items	Cube Limit Nat
0	350,000
Hub/Spoke Items	Cube Limit E
0	75,000
Trans Costs	Cube Limit C
$0.00	275,000
	Cube Limit W
	150,000

Figure 22: Sample Product Selection Model

On the left, the two columns with the decision variables, x_i^{Hubs} and x_i^{Nat}, are set as binary variables and change for each product based on whether that product is to be centralized in that network. In the middle portion of the model, totals for savings, transportation costs, and cube required are computed based on the values of the decision variables. Once the solver has finished running, the objective cell will display the total potential savings for centralizing products as shown in the decision variables columns. On the far right, the constraints for the transportation budget and space available at each hub are available to be set by the company. In Figure 22, the constraints for the transportation cost budget and space limitations are set at an arbitrary value for demonstration purposes only.

Using this model, the company can quickly evaluate which products should be considered for centralization into one of the available networks. The company may also decide that utilizing both networks is not an option, so the model can be easily adjusted to account for this decision. To use only the hub-and-spoke network, the constraints in the model are updated to include a constraint that sets all x_i^{Nat} zero. To use only the fully centralized network, all x_i^{Hubs} are set to zero.

4.3 Model Extensions

The model described in this section can easily be adapted to consider additional constraints that are deemed necessary by the company. One of these potential new constraints is the type of storage required for different products. For some pharmaceutical products, special storage requirements must be considered. For example, due to the nature of many of these products, refrigerated storage is necessary, but space in this type of storage is more limited than the standard storage considered in the original model. Other pharmaceutical products are heavily controlled by government agencies, so they require storage in a secure vault or cage. Space in these types of storage areas may be very limited, or it may be possible that government restrictions make it too difficult to consider centralizing this type of product. To consider these types of restrictions, a column is added to the original data that describes the storage type and any necessary constraints are added to the model.

Another possible constraint to consider is based on the average demand volume of an item. Some of this may already be covered in the transportation costs and space limitations. However some items with very high demand volume are still in the list of potential products. For an item with very high demand volume, it simply may not be feasible for the company to handle the number of individual shipments required to centralize the product.

4.4 Examples

In this section, three example scenarios for the product selection model are tested. The values for the transportation budget and cube limitations are picked using values within the range of the totals for the 50 sample products chosen.

4.4.1 Scenario 1

In the first scenario, the constraints for cube limits and the transportation budget are relatively low compared to the totals for the 50 products in the selected sample. Figure 23 shows the outcome of this scenario. Using these constraints, 29 of the 50 items were selected to be centralized in the national network, 7 items were selected to be centralized in the hub-and-spoke network, and the remaining 14 items were left out of the centralization all together, with a total of over $9 million in expected inventory savings. These savings are generated at an additional $1.5 million in annualized transportation cost.

OBJECTIVE	CONSTRAINTS
Total Savings	Trans Budget
$9,204,159	$1,500,000
National Items	Cube Limit Nat
29	250,000
Hub/Spoke Items	Cube Limit E
7	50,000
Trans Costs	Cube Limit C
$1,499,584	175,000
	Cube Limit W
	100,000

Figure 23: Output for Scenario 1

4.4.2 Scenario 2

In the second scenario, the limits for the available cube in each location remained the same and the annual transportation budget was increased to $5,000,000. Figure 24 shows the outcome of this scenario. Using these constraints, 42 items were selected to be centralized in the national network, 4 items were selected to be centralized in the hub-and-spoke network, and the 4 remaining items were not centralized in either network, with over $11 million in expected savings. These savings are generated at an expected $2.4 million in additional transportation costs. In both of these scenarios, most items were centralized into the national network. This network requires a much smaller inventory target for items, thus it returns much higher expected savings for all of the products. For Scenario 3, the national network is excluded from the selection process.

OBJECTIVE		**CONSTRAINTS**
Total Savings		**Trans Budget**
$11,036,067		$5,000,000
National Items		**Cube Limit Nat**
42		250,000
Hub/Spoke Items		**Cube Limit E**
4		50,000
Trans Costs		**Cube Limit C**
$2,337,684		175,000
		Cube Limit W
		100,000

Figure 24: Output for Scenario 2

4.4.3 Scenario 3

This scenario uses the same constraints as Scenario 1, with the only difference being that the national network is excluded from consideration. Figure 25 shows the outcome of this scenario. In this scenario, 41 out of the 50 products were selected to be centralized in the hub-and-spoke network and the remaining 9 were left out of the centralized network, with over $5.5 million in expected savings. These savings are generated at an expected $1.4 million in additional transportation costs. These 9 items had some of the highest expected net savings, but also have either high transportation costs, high cube, or fairly high inventory targets in each of the hubs.

OBJECTIVE		**CONSTRAINTS**
Total Savings		**Trans Budget**
$5,543,595		$1,500,000
National Items		**Cube Limit Nat**
0		250,000
Hub/Spoke Items		**Cube Limit E**
41		50,000
Trans Costs		**Cube Limit C**
$1,413,443		175,000
		Cube Limit W
		100,000

Figure 25: Output for Scenario 3

5 Conclusion

Centralizing the inventory in a pharmaceutical supply chain can potentially save a company millions of dollars. This thesis has detailed a method for analyzing demand patterns, reviewing the current network and inventory policies, and evaluating two new network options. A tool was then developed in Excel to help the company select the products to be centralized in one of the new networks. For each of the network options, old and new, demand patterns were analyzed to set an inventory target, which was then used to calculate expected inventory and annual inventory holding costs. The given demand data was used to estimate annual transportation costs for using the new networks, using an overnight shipment cost rating tool to evaluate the cost of shipping the actual orders and then annualizing those costs for each location. Using the expected annual inventory holding costs and expected annual transportation costs, the expected savings for centralizing products in each of the new networks were calculated. The set of products with positive expected net savings was then evaluated using the tool developed for product selection. The results of the analysis showed that by centralizing just 100 products into the hub-and-spoke network can save the company over $10 million, while centralizing just 100 products into the national network can save the company over $14 million.

The analysis completed in this thesis was limited to a small set of demand data provided by the company. To achieve better accuracy in the inventory policies and value of centralizing inventory, the analysis can be repeated with newer demand data from a longer period of time. Due to the time required to calculate the inventory targets and run the remainder of the analysis on approximately 43,000 products, the analysis can be efficiently repeated every 3 or 6 months, depending on how frequently the company would like to update their inventory policies.

The product selection tool developed in this thesis considers only the net savings, annual transportation costs, and size of the products being considered. For future work, the tool can be updated to include more specific constraints, such as the type of storage or the type of product. Some product types, especially those that are heavily regulated by the government, can be difficult to move around in a network, so they may be excluded from the centralization. In each hub, the available space for inventory in different types of storage, such as refrigerated storage or vault storage, may be more limited than standard storage in the warehouse. These additional space limitations can be easily added to the tool, as well as any other constraints the company wishes to add. The addition of constraints provides more accurate, in-depth results for the company.

The analysis completed in this thesis aimed to identify cost saving opportunities for the company. In future work, the company may also wish to revisit some of the assumptions made in the analysis completed in this thesis. One of those assumptions is the impact on labor costs when relocating inventory. When inventory is removed from one location, less labor is required in that location. Similarly, adding inventory to a location will increase the labor required in that location for picking customer orders and putting away incoming stock. Another assumption that can be reconsidered is the changes in facility costs, such as additional space required or additional empty space when relocating inventory. A more detailed analysis may be completed using these updated assumptions to give a clearer picture of the savings available when using a hub-and-spoke network.

References

2016 MDM market leaders: Top pharmaceutical distributors. (2016). Retrieved from https://www.mdm.com/2016-top-pharmaceuticals-distributors

Amaro, A., & Barbosa-Póvoa, A. (2008). Planning and scheduling of industrial supply chains with reverse flows: A real pharmaceutical case study. *Computers & Chemical Engineering, 32*(11), 2606-2625.

Bacchetti, A., & Saccani, N. (2012). Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice. *Omega, 40*(6), 722-737.

Ballou, R. H., & Burnetas, A. (2003). Planning multiple location inventories. *Journal of Business Logistics, 24*(2), 65-89.

Berman, O., Krass, D., & Tajbakhsh, M. (2011). On the benefits of risk pooling in inventory management. *Production and Operations Management, 20*(1), 57-71.

Boylan, J., Syntetos, A., & Karakostas, G. (2008). Classification for forecasting and stock control: A case study. *Journal of hte Operational Research Society, 59*(4), 473-481.

Choi, P., & Gellad, W. (2013). Prescription drug and pharmaceutical policy. In *An introduction to health policy* (pp. 101-118). Springer.

Çömez-Dolgan, N., & Tanyeri, B. (2015). Inventory performance with pooling: Evidence from mergers and acquisitions. *International Journal of Production Economics, 168*, 331-339.

Correia, I., Nickel, S., & Saldanha-da-Gama, F. (2011). Hub and spoke network design with single-assignment, capacity decisions and balancing requirements. *Applied Mathematical Modelling, 35*(10), 4841-4851.

Hopp, W. J., & Spearman, M. L. (2011). *Factory Physics.* Waveland Press.

Kaiser Family Foundation. (2005). *Follow the pill: Understanding the U.S. commercial pharmaceutical supply chain.* Retrieved January 24, 2017, from http://kff.org/other/report/follow-the-pill-understanding-the-u-s/

Kouvelis, P., Xiao, Y., & Yang, N. (2015). PBM competition in pharmaceutical supply chain: Formulary design and drug pricing. *Manufacturing & Service Operations Management, 17*(4), 511-526.

Llamasoft, I. (2015). *Supply Chain Guru 8.3 User Guide.* Retrieved 2017

Nagurney, A., Li, D., & Nagurney, L. S. (2014). Pharmaceutical supply chain networks with outsourcing under price and quality competition. *International Transactions in Operational Research, 20*(6), 859-888.

Narayana, S. A., Pati, R. K., & Vrat, P. (2014). Managerial research on the pharmaceutical supply chain - A critical review and some insights for future directions. *Journal of Purchasing and Supply Management, 20*(1), 18-40.

Payne, P. (2008). Is pharma's supply chain safe? *Pharmaceutical Technology Europe, 20*(7), pp. 15-16. Retrieved from http://ezproxy.libraries.wright.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=33240298&site=ehost-live

Pedroso, M. C., & Nakano, D. (2009). Knowledge and information flows in supply chains: A study on pharmaceutical companies. *International Journal of Production Economics, 122*(1), 376-384.

Richardson, H. (1995). Control your costs--then cut them. *Transportation & Distribution, 36*(12), 94.

Rossetti, C. L., Handfield, R., & Dooley, K. J. (2011). Forces, trends, and decisions in pharmaceutical supply chain management. *International Journal of Physical Distribution & Logistics Management, 41*(6), 601-622.

Schmitt, A. J., Sun, S. A., Snyder, L. V., & Shen, Z.-J. M. (2015). Centralization versus decentralization: Risk pooling, risk diversification, and supply chain disruptions. *Omega, 52*, 201-212.

Shah, N. (2004). Pharmaceutical supply chains: Key issues and strategies for optimisation. *Computers & Chemical Engineering, 28*(6), 929-941.

Skipper, J. B., Cunningham, W. A., Boone, C. A., & Hill, R. R. (2016). Managing hub and spoke networks: A military case comparing time and cost. *Journal of Global Business and Technology, 12*(1), 33-47.

Sousa, R. T., Liu, S., Papageorgiou, L. G., & Shah, N. (2011). Global supply chain planning for pharmaceuticals. *Chemical Engineering Research and Design, 89*(11), 2396-2409.

Syntetos, A. A., Boylan, J. E., & Croston, J. D. (2005). On the categorization of demand patterns. *Journal of the Operational Research Society, 56*(5), 495-503.

Uthayakumar, R., & Priyan, S. (2013). Pharmaceutical supply chain and inventory management strategies: Optimization for a pharmaceutical company and a hospital. *Operations Research for Health Care, 2*(3), 52-64.

Williams, B. D., & Tokar, T. (2008). A review of inventory management research in major logistics journals: Themes and future directions. *The International Journal of Logistics Management, 19*(2), 212-232.

Yoon, M.-G., & Current, J. (2008). The hub location and network design problem with fixed and variable arc costs: Formulation and dual-based solution heuristic. *Journal of the Operational Research Society, 59*(1), 80-89.

Zäpfel, G., & Wasner, M. (2002). Planning and optimization of hub-and-spoke transportation networks of cooperative third-party logistics providers. *International Journal of Production Economics, 78*(2), 207-220.