The Effects of Data Sharing on a Perishable Goods Supply Chain

by

Darren Wilson
B.Com. Strategic Management, University of South Africa, 2007

and

Gerald Yeung
B.S. Mechanical & Aerospace Engineering, Cornell University, 2008

Submitted to the Engineering Systems Division in Partial Fulfillment of the
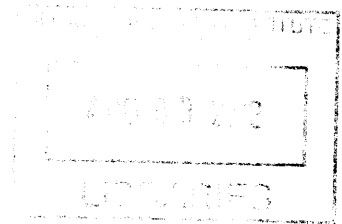Requirements for the Degree of

Master of Engineering in Logistics

at the

Massachusetts Institute of Technology

The authors hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this document in whole or in part.

Signature of
Author...
Master of Engineering in Logistics Program, Engineering Systems Division
May 6, 2012

Signature of
Author...
Master of Engineering in Logistics Program, Engineering Systems Division
May 6, 2012

Certified by...
Dr. Roberto Perez-Franco
Research Associate, Center for Transportation and Logistics, Thesis Supervisor

Accepted by...
Prof. Yossi Sheffi, Director, Center for Transportation and Logistics, Engineering Systems Division

The Effects of Data Sharing on a Perishable Goods Supply Chain

by

Darren Wilson

and

Gerald Yeung

Submitted to the Engineering Systems Division in Partial Fulfillment of the
Requirements for the Degree of

Master of Engineering in Logistics
at the
Massachusetts Institute of Technology

ABSTRACT

This research project explores the benefits of retail data sharing in a high-velocity perishable goods supply chain. While this technique has been largely effective in improving supply chain performance in different industries, its benefits are unproven in the perishable goods business. Specifically, due to the short shelf life of produce, it remains to be seen whether data sharing can generate actionable plans for retailers to reduce out-of-stock events and shrinkage due to spoilage. As a result, suppliers and retailers alike have been reluctant to invest in the technology and cultivate the business relationship required to enable data sharing. The findings of this thesis could help companies determine whether a business case can be built for suppliers to invest in the necessary technology, as well as for retailers to share operational data for the greater good of overall supply chain efficiency and profitability. Ultimately, our research indicates that without some fundamental changes to the retailers' ordering process, data sharing does not provide substantial operational benefits for the perishable goods supply chain.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The present thesis discusses the benefits of retailer data sharing, both for suppliers and retailers. In this section, we will outline the structure of this thesis and the research approach we used.

1.1 Topic Overview

As the business environment becomes ever more competitive, companies continue to look for ways to streamline their operations. To that end, they often push for efficiency and cost reductions within their internal operations, failing to realize that real benefits could be derived from co-operation with external players in their supply chain. According to Mentzer et. al (2001), using a strategic partnership to converge interfirm operational and strategic capabilities is one of the main philosophies of modern supply chain management. In recent years, data sharing has become an important element of interfirm cooperation, evident from the fact that many companies have succeeded in leveraging this new opportunity. This thesis will explore the benefits data sharing can bring to both suppliers and retailers of perishable goods.

1.2 Company Background

Our thesis sponsor is recognized as a leading packaged salad and fresh produce supplier, we will use an alias of SuperSalad for the purposes of this thesis. Despite SuperSalad's strong presence in the global fresh food market, the very short shelf life of its products creates significant challenges within its supply chain. Historically the average shrinkage rate has always been high, sometimes as high as 20%. While this is a problem common within most perishable good

businesses, it is nonetheless a massive source of waste that also adversely affects the profitability of both SuperSalad and its customers.

1.3 Research Question

As retailer data sharing is becoming increasingly prevalent, management of SuperSalad is hopeful that this technique can help improve their ongoing problem of product *shrinkage* (for the purpose of this thesis, *shrinkage* will be defined as loss of goods solely due to spoilage). However, skepticism remains about whether this process will be applicable to SuperSalad's high-velocity supply chain. To justify the capital investment required to implement this process – and to gain buy-in from retailers and internal stakeholders – SuperSalad need to see potential for significant improvements to the specific issues in their supply chain. To this end, they would like to see a qualitative and quantitative analysis of the impact data sharing could have on their and their retailer's businesses. For the purpose of this thesis, we decided to focus on whether data sharing can make a difference to the combined supply chain's biggest issue, shrinkage of fresh produce.

1.4 Motivation

In the fresh produce industry where stock turns are high and shelf lives are short, data sharing provides an opportunity to build a truly collaborative supply chain. The potential benefits, however, are difficult to quantify and as a result, many companies have had limited success in convincing their partners to collaborate in this fashion. Other obstacles include the risk of

exchanging competitive information and losing bargaining power, so a strong case needs to be made to convince companies to adopt this strategy for long-term sustainable cost reductions.

The findings from this thesis should provide valuable insight into how companies can make a business case for investing in data sharing. Ultimately, the insights derived from this thesis can be applied to other high velocity supply chains.

1.5 Approach Overview

For the first part of the thesis, a literature review has been performed to evaluate existing data sharing techniques and their respective effectiveness. To acquire real world perspectives, we have interviewed companies that have established successful data sharing relationships with retailers to understand the hurdles they have overcome.

For the second part of the thesis, based on our findings from the literature review and company interviews, we have selected quantifying techniques that are most relevant to SuperSalad's business model. We approached the thesis question by first quantifying the current shrinkage being experienced by retailers, followed by an analysis to quantify the benefits experienced by SuperSalad and other companies in the perishable goods market through data sharing.

1.6 Structure of Thesis

Section 1 defines the thesis problem statement and underlines its significance both for

SuperSalad and other perishable goods supply chains. This section also includes background information on SuperSalad and the motivation of the topic in a larger context. Section 2, the Literature Review, focuses on studies that have been performed on the subject of data sharing. We have discussed aspects of these studies that are relevant to SuperSalad's needs and commented on areas we identified for further research. Section 3 describes the research methods and approach of this thesis. Specifically, we present methods of interview and data collection and analysis.

Section 4 contains our data collection, in which findings from interviews and initial data analyses are summarized. Section 5 is devoted to data analysis and discussion based on results from the previous section. Some recommendations are made at the end of this section. Section 6 presents the thesis conclusion that summarizes our findings and recommendations. This section is followed by a list of references.

2. LITERATURE REVIEW

In this section, we present other research that has been done on the topic of data sharing and comment on how these studies can be applicable to SuperSalad's supply chain.

2.1 Overview of Benefits of Retailer Data Sharing

Both the supplier and retailer stand to benefit from retailer data sharing, as shown in a case study conducted by the Grocery Manufacturers Association (GMA) in 2009 on a company that piloted a retailer-direct data project in 2006. It is GMA's claim that "retailer-direct data sharing is new enough that most adopting suppliers gain significant benefits with unlimited opportunity and learning ahead," but little details had been provided in support of this claim. The specific benefits presented in the case study are as follows:

Supplier Benefits:

·   Improvement in retail in-stock positions

·   Reduction in lost sales

·   Exceeding customer service requirements

·   Decrease in raw, packaging and finished goods inventory

·   Increase in inventory turns

·   Reduction in demand variability

·   Enabling flexibility in physical distribution network design

·   Curtailing operating and transportation costs

·   Eliminating and automated many manual and administrative processes

Retailer Benefits:

- Lowering safety stocks at distribution centers

- Improvement in gross margin return-on-investment

- Increase in sales through higher in-stock positions

- Reduction in lead-time and improved order fill rates

- Enhancing co-managed or vendor-managed inventory capabilities

Because the above claims are presented with limited supporting evidence, further studies on some of the points above have been performed and the results are discussed below.

The rest of this literary review is structured as follows: Section 2.2 Benefits for Supply Chain Partners; Section 2.3 Benefits for Customers / Retailers; Section 2.4 Benefits for SuperSalad / Suppliers.

2.2 Benefits for Supply Chain Partners

In this subsection we will discuss the benefits of data sharing from a supply chain partnership point of view.

2.2.1 The Value of Effective Partnerships

Duffy and Fearne (2006) stated that successful companies are moving away from the traditional,

adversarial relationships that have been typical between suppliers and retailers. Researchers such as Lamming [1993] and Christopher [1998] stated that companies should realize that the transfer of costs between supply chain players does not improve the performance of the individual firms as all costs will ultimately be passed on to the marketplace. Therefore companies that recognize the benefits of long-term, co-operative relationships will enhance the performance of the entire supply chain and have a higher probability of being successful.

2.2.2. Maximizing Profit of the Whole Supply Chain

Li and Zhang (2008) argued that "double-marginalization," a concepts in economics in which each player in the supply chain maximizes its individual profit, is a known cause of supply chain inefficiency from a profit standpoint. In particular, the manufacturer tends to maximize profit by charging a high wholesale price and thereby hurting the overall surplus of the supply chain. As a remedy to the double-marginalization problem, the authors suggest the sharing of individual retailer demand data with the manufacturer. With this aggregate demand data, the manufacturer can then set the appropriate wholesale price. Overall, the authors argued - and proved through a simulation - that this data asymmetry within the supply chain will result in a lower equilibrium wholesale price.

However, the authors argued that for this model to work confidentiality must be maintained, i.e. the manufacturer must not disclose retailer information, both directly or indirectly, to other retailers. First, as proven in their simulation, double-marginalization could only be eliminated when retailers have no knowledge of their competitors' data. In addition, maintaining

14

confidentiality will encourage retailers to truthfully report data without having to worrying about giving away sensitive competitive information to their competitors.

While Li and Zhang (2008) provides valuable insight from an economist's point of view that is backed by running different scenarios in a three-stage game, there are some obvious shortfalls in the analysis. First, the authors admitted that there is an ongoing debate on whether information asymmetry has a positive or adverse effect on supply chain efficiency and profitability. Our thesis will attempt to address this question. Also, in this particular model, only wholesale price is used as a measurement of supply chain performance. This is certainly an area that calls for further research.

In addition, the analysis of Li and Zhang (2008) is relatively general and not product or industry specific. As a result, it does not account for some nuances of specific industries. Our thesis will further explore this topic to determine if this model can be applied to a perishable supply chain.

2.3 Benefits for Customers / Retailers

2.3.1 Better Customer Focus

In a case study on supply chain integration through information sharing, Grean and Shaw (2003) evaluated the merits of data-sharing on three fronts: transactional, operational and strategic. In particular, the strategic portion explained how information sharing aligned both firms' vision, allowing them to acquire a better understanding of each others' businesses and markets to better

serve their end customers.

With the development of a *data highway* that linked their database to Wal-mart's, P&G was able to combine real-time Wal-mart store data with their own market research to come up with actionable plans. This collaboration supported the development of a number of applications (e.g. joint business scorecards, EDI and category management) that enabled both companies to better serve their customers. In addition, data-sharing allowed both companies to reduce costs through lowering inventory and eliminating duplicate operations. These savings were in turn passed on to the customers and increased customer satisfaction.

Ultimately, the implementation of data-sharing successfully transformed the relationship between the two parties from "adversial" to "collaborative", a major breakthrough that cannot be easily quantified in monetary terms. In particular, the success of this partnership encouraged P&G to abandon their traditional short-term day-to-day approach and adopt a long-term partnership.

While this paper highlighted the many benefits Wal-mart and P&G reaped through data-sharing, it remains to be seen whether this model can be replicated between grocers and SuperSalad. First, P&G supplies hundreds of different products to Wal-mart, whereas SuperSalad provides less than 50 SKU's to any given customer. As such, it remains unclear whether SuperSalad's customers (i.e. retailers) can justify dedicating the time and resources needed to cultivate a similar partnership. In addition, since most grocers do not have the same scale and infrastructure as Wal-mart, it may be impossible for them to support data sharing without major upfront

investment.

2.3.2 Efficient Consumer Response

Duffy and Fearne (2006) also referred to an initiative known as Efficient Consumer Response (ECR), initially developed by Wal-mart and Procter and Gamble but later adopted as a trade initiative. Fiddis (1997) defined the European ECR initiative as "global movement in the grocery industry focusing on the total supply chain - suppliers, manufacturers, wholesalers and retailers, working closer together to fulfill the changing needs of the grocery consumer better, faster and at least cost." A primary enabler of this initiative is the sharing of data between firms, to create a holistic view of the supply chain which aids collaboration. Figure 1 illustrates the focus areas of ECR.
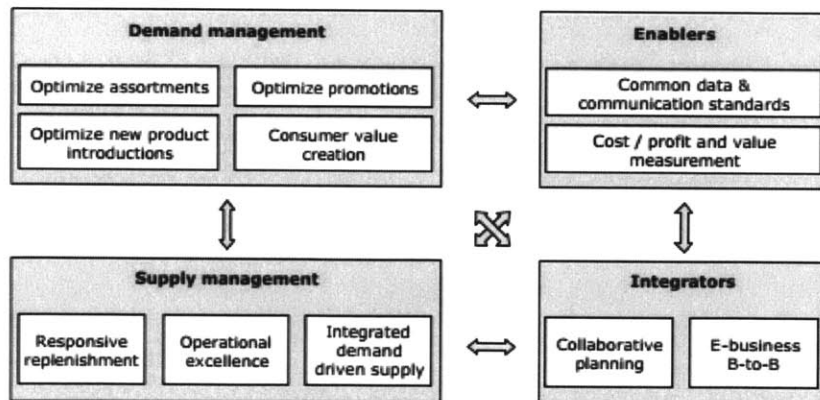
Figure 1. The Four Focus Areas of ECR (Sourced from ECR Europe, 2000)

The first ECR model was developed by the Food Marketing Institute in conjunction with Kurt Salmon Associates. This was a successful model that allowed supplier production to be

completely managed by the retailers' point of sale activities, streamlining production and inventory throughout the supply chain.

Although all of the above research illustrates a strong relationship between data sharing and supply chain improvement, it is inconclusive whether products with velocity as high as that of packaged salad can benefit sufficiently from the investment required to implement the process. We will attempt to answer that question in this thesis.

2.3.3 Reduce Out-of-stock and Lost Sales

Data sharing also allows for better inventory management, as demonstrated in a case study by Grant (2009) on Food Lion and Kimberly-Clark (K-C) supply chain cooperation. When Food Lion introduced the Vendor Pulse program in 2006 to share retail level data with suppliers, Kimberly-Clark saw it as an opportunity to strengthen their position in the personal care market. Despite some initial challenges, including K-C's large number of SKU's, both K-C and Food Lion eventually saw the benefits of this initiative. In particular, sharing of retail data greatly enhanced their reparation for promotional events, in which the out-of-stock rate could reach as high as 18%. Through real-time store level data broken down by SKU's, K-C were able to identify specific store locations and products that had the highest out-of-stock rates. This learning allowed them to better allocate inventory for the next event.

Furthermore, availability of retail data not only improved the responsiveness of their supply chain, it also enabled K-C to match their analysis with shipment data to identify shipping voids

when items were shipped to the wrong stores. In the end, the out-of-stock rate for promotional events was reduced to 10%. In addition, this result enhanced the relationship between K-C and Food Lion as the data enabled K-C to make inventory policy recommendations to Food Lion to help reduce their operating costs.

Although some of the results presented by Grant (2009) have been confirmed by Hudock and Vowell (2011) at a conference with Kimberly-Clark supply chain group, there is some skepticism on whether this model can be replicated for SuperSalad and its customers. First, retail stores carry more SKU's for K-C than for SuperSalad. As a result, retailers may be more reluctant to invest in a similar data sharing relationship with SuperSalad. In addition, K-C products are predominantly non-perishable goods that have much longer shelf lives than produce. With a much shorter response time, data sharing may not yield the same benefits achieved in the K-C case.

2.3.4 Reduce Shrinkage

In a paper entitled "Supermarket Loss Estimates for Fresh Fruits, Vegetables, Meat, Poultry and Seafood," USDA researchers quantified the percentages of fresh produce being discarded by retailers. For fresh lettuce, the average percentage of produce being lost was 13.9% which is one of the highest among fresh produce.

This is obviously an area with high potential for improvement, and if we can demonstrate that data-sharing does effectively reduce shrinkage, this will create a huge incentive for retailers to

invest in the technology.

2.4 Benefits for SuperSalad/ Supplier

2.4.1 Minimizing Disruption Risk

Retail data-sharing can minimize supply chain disruption risk. Disruption risk includes operational upsets, political instability, and issues arising from natural disasters. Wakolbinger and Cruz (2010) argued that data-sharing between supplier and retailers promotes joint problem solving in the event of supply chain disruption. By using a simulated supply chain network consisting of three tiers of decision-makers (supplier, retailer and market demand), the authors demonstrated how a mutual sharing of information among these three parties can lead them to make operational decisions that most effectively minimize disruption risk.

Through this simulation, the authors also pointed out that the beneficiaries of information-sharing activities are different for every supply chain, dependent upon the bargaining power of the players as well as the specifics of their risk-sharing contract.

The findings of Wakolbinger and Cruz (2010) are particularly useful for SuperSalad on two fronts. First, the fresh perishable goods business is highly susceptible to disruptions such as weather, natural disaster and labor law reforms. As demonstrated by Wakolbinger and Cruz (2010), data sharing can help both retailers and supplier reach decisions that will minimize overall disruption risk. Also, because of SuperSalad products' limited shelf life and SuperSalad's

high velocity supply chain, management needs to be able to react quickly in the event of supply chain disruption. The findings of Wakolbinger and Cruz (2010) seem to build a case for data-sharing from both retailers' as well as supplier's standpoint. In future research, the authors suggested validating their findings through real-life cases in industry.

2.4.2 Increase in forecast accuracy

In a paper titled "The Impact of Information Sharing on Supply Chain Performance" Zhao (2008) developed a number of mathematical models to quantify the benefits of data sharing. In conclusion, he states that "information sharing can always improve the manufacturer's forecasting accuracy." His models also prove that the value of information sharing increases as the delivery lead time decreases and the number of retailers increases.

This is particularly relevant to SuperSalad, as their products necessitate extremely short lead times and their products are sold in hundreds of retail locations. Zhao (2008), however, did not address the impact of product shelf life on data sharing benefits. As previously stated, unless retail data can provide substantial - and more importantly, immediate - benefits for SuperSalad and its customers, it will be difficult to justify the capital and time investment required to establish data sharing relationship.

2.4.3 Dampening of the bullwhip effect

One known benefit of data sharing is that it can effectively dampen the "bullwhip effect," a

common issue within industry supply chains in which the amplitude of order fluctuation increases as one moves further up the chain, as explained by Agarwal (2009). The term was first coined by Procter & Gamble to describe surprising oscillations observed in Pampers' production ordering when diaper consumption rate was supposed to be fairly constant. After further investigation, they found that order variation in production level can be attributed to lead-times at various stages of the value chain and to the fact that upstream suppliers generally have little knowledge of the actual demand pattern. In the case of SuperSalad where a significant portion of their revenue comes from promotion events, the high variability in demand, coupled with the bullwhip effect, will present even greater challenge in production and inventory management. As a result, data sharing has high potential benefit for SuperSalad's business.

To verify that data sharing does indeed mediate the bullwhip effect, Croson and Donohuf (2003) designed an experiment to understand how human ordering pattern changes when retail data becomes available. Participants in the experiment were given different levels of POS data and the goal of the experiment was to see if there was a correlation between availability of data and supply chain performance. They concluded that while access to POS data throughout the supply chain does not completely eliminate the bullwhip effect, it results in a marked reduction in order amplification within the supply chain. This benefit becomes more apparent upstream in the chain because it enables upstream suppliers to interpret internal orders and better understand the actual demand pattern in making order and inventory decisions. Ultimately, the authors argue that since the higher echelons have the most to gain from POS data, they have the strongest incentive to invest in data sharing technology.

Lee, So, and Tang (2000) used a simple two-stage supply chain model (one retailer and one manufacturer) to derive mathematically that demand distortion caused by the bullwhip effect can be ameliorated when retail level data is available to the manufacturer. Specifically, manufacturers see best results when variation of demand is high and when lead-time is long. While these two studies both arrived at the same conclusion - that data-sharing allows manufacturer to enjoy significant benefits in inventory reduction and cost savings - Lee et al. further argue that retailers stand to benefit from the arrangement as well, as visible cost savings for the manufacturer will help retailers to negotiate better pricing and lead time as an incentive for providing retail data. This argument by Lee et al (2000) would be all the more compelling if they had provided real life success stories in which retailers do enjoy visible benefits after providing store data to manufacturers.

3. METHODOLOGY

This section outlines the methodology through which we attempt to answer the main question of our thesis, specifically, the benefits of data-sharing on a perishable goods supply chain.

3.1 Methodology Overview

According to Dul and Hak (2008), theory-testing research follows the following steps:

1. Choosing research strategy - experiment, survey, or case study

2. Selecting specific methods based on strategy chosen

3. Formulating a hypothesis

4. Conducting qualitative and quantitative measurements

5. Conducting data analysis

To choose an appropriate research strategy, we elected to conduct preliminary qualitative research in the form of interviews to gain a background understanding of the problem. Since the focus of the project is to evaluate the benefits of data sharing, we began our research by interviewing subject matter experts - both SuperSalad employees and at SuperSalad's customers - to develop a thorough understanding of current industry landscape and existing data sharing practice. After sampling - determining appropriate interview subjects - we developed different protocols - sets of interview questions - based on interviewee job function and background.

Based on the results of our preliminary interviews, we formulated the hypothesis that data sharing would provide significant benefits to both SuperSalad and their customers. To prove this

hypothesis, we decided to use comparative case study as suggested by Dul and Hak (2008), which will be further explained later in this section. Specifically, to understand the effect of data sharing on SuperSalad's perishable supply chain, we collected retail data from a number of SuperSalad customers that have different level of data sharing policies and compared their performances.

After all qualitative and quantitative data has been collected, we proceeded to the final step, namely, data analysis as proposed by Dul and Hak (2008).We mapped out current buying process of SuperSalad products to identify possible system bottlenecks. Lastly, based on our quantitative data analysis, we draw a conclusion on whether data sharing would be beneficial to SuperSalad's high velocity supply chain. We also make recommendation on how to best use this data to produce best results.

3.2 Interviews

According to Weiss (1994), interviews provide access to the observations of others. In the context of business, they shed light on organizational goals and on challenges people confront on the job. To gain a thorough understanding of the fresh salad industry and the underlying issues - both supporting and opposing forces in data sharing implementation - behind the question of our thesis, we conducted interviews with experts from different areas of the business.

3.2.1 Sampling

In an attempt to understand the problem from all levels of the supply chain, we devised an interview plan to approach the issue from both SuperSalad's and retailers' perspectives. Using the *stratified purposeful sampling* method as proposed by Patton (2002), we chose our subjects based on their particular interest and job functions. Patton (2002) defines *stratified purposeful sampling* to be a sampling technique that "illustrates characteristics of particular subgroups of interest that allows for comparisons." We grouped our subjects into three main groups:

SuperSalad Internal Supply Chain Group

We chose to interview individuals from this group as they could help identify some of the drivers behind data sharing from SuperSalad's perspective.

SuperSalad Customer Service Managers

We chose to interview members from this group because their daily interaction with customers (retailers) could provide insight into the perceived value of data sharing from retailers' perspective. Specifically, we hoped to understand both the perceived benefits for the retailers and the challenges they must overcome to facilitate the exchange of retail data.

Produce Managers at various retailers

By interviewing the frontline employees at retailers, we could gain an understanding of their inventory management policy and the flow of SuperSalad's product in retailers' supply chain. This would also help us determine whether SuperSalad's assessment of their supply chain performance was aligned with the performance perceived by their customers.

3.2.2 Conduct Interviews

Through our interviews we hoped to achieve the following goals:

1. Refine our thesis question to ensure that we were tackling the topic of interest for SuperSalad

2. Cultivate relationships with key personnel

3. Identify opposing forces and aggregate our findings to develop a balanced, non-biased understanding of the issues

4. Identify areas of further research

5. Construct a framework for subsequent quantitative analysis

According to Rubin and Rubin (2005), qualitative interviewing involves three kinds of questions: main questions, follow-up questions and probes.

- Main questions typically evolve around the heart of the interview topic. Sample main questions relevant to our thesis topic included "Can you explain SuperSalad's current order fulfillment process" or "What is the shrinkage rate of SuperSalad's product."

- Follow-up questions are specific to the responses interview subjects have provided during the interview. An example of follow-up question include "Can you clarify your last point" or "How exactly is the term 'shrinkage' defined as you have mentioned." These questions are crucial in providing depth and details and can sometimes highlight nuances of the issue.

- Probes are lead-on questions that encourage interviewee to clarify on what has already been said and to expand on related topics. An example of probes would be asking for the subject's opinion on current process bottlenecks or on perceived benefits of data sharing.

To conduct interviews with a focus on our topic of interest and to provide subject the opportunity to share their experiences that may not be touched upon in our planned questions, it is important to incorporate all three types of questions in our interviews. As such, we had to develop specific sets of questions for each group, as shown below:

Main Questions for SuperSalad Internal Supply Chain Group

1. What are the key indicators of supply chain performance as perceived by SuperSalad and the retailers?

2. What are the main drivers for data-sharing? Do any of SuperSalad's current customers share store-level data?

3. Is there a seasonality in product demand? How does that affect in-store shrinkage rate?

4. What is shrinkage at SuperSalad's warehouse compared to that at retailers' stores?

5. What are the perceived benefits of data-sharing?

6. From SuperSalad's perspective, what are the infrastructure and managerial challenges the company must overcome to facilitate data-sharing?

Main Questions for SuperSalad Customer Service Managers

1. How is shrinkage perceived by retailers? How does that affect retailers' and SuperSalad's bottom line?

2. Have you observed different shrinkage rates between different store locations? Between different products? What are the main variables that affect shrinkage?

3. What short-term and long-term impact does shrinkage have on end consumer satisfaction?

4. What are some retailers' concerns over data-sharing?

5. From the retailers' perspective, what are the infrastructur and managerial challenges they must overcome to enable sharing of store data with SuperSalad?

Main Questions for Produce Managers at various retailers

1. How often are SuperSalad's product restocked?

. 2. What is the volume of the store?

3. How does the quality (shelf life, condition) of SuperSalad products compare with their competitors'?

4. At the retailer, who is in charge of the ordering process?

5. How is shrinkage and out-of-stock measured? What is the retailer's average shrinkage rate and out-of-stock rate?

6. What is the typical shelf life of SuperSalad products upon arrival at the store?

For the last group, we decided to conduct these interviews on the field, at the retailers' store locations. According to Lewis-Beck et al (2004), field research is conducted because it is difficult to create "authentic social conditions out of their actual context" in simulations and laboratory experiments. Humans, after all, are emotional and reflexive beings and observing

them in their natural everyday settings can provide insights that cannot be gleaned from simulated situations.

Because SuperSalad's supply chain and customer service staff are located throughout the US, most of the internal interviews would have to be performed over the phone. Interviews with Produce Managers would be conducted during a day trip to visit a number of grocers in the Boston, area. This field trip would also enhance our understanding of the daily operations at these locations. Specifically, it would provide insights on store employees' priorities and challenges on the job.

During these interviews it would be important to ask both direct, factual questions (e.g. specific chain shrinkage rate) and open-ended questions (e.g. what are the challenges...) to provide an opportunity for subjects to share personal experience and identify relevant issues that might not have been apparent to the rest of us initially.

3.2.3 Analysis of Interviews

For analysis, we followed the four processes of qualitative interview analysis as proposed by Weiss (1994).

The first tool, *coding*, is linking response of subject to concepts and categories relevant to our study (Weiss 1994 pp.154). In our case, example categories included "perceived benefits of data sharing", "limitations of data sharing technology" and "investment required."

Following coding, the subsequent steps as proposed by Weiss (1994) were performed in the following order (Weiss 1994 pp.158-160). After sorting our findings, we perform *local integration* in which we summarized what our subjects had spoken in one particular category and offered our interpretation on these results. Having performed this task for every category and concept, we finally proceeded to *inclusive integration* in which we connect all these categories into one single coherent story that was the theme of our thesis (Weiss 1994 pp.160-162).

In addition to the steps outlined by Weiss (1994), other factors and considerations must also be kept in mind. When analyzing the results of our interviews, we strived to see the issue from the perspective of the subject in a context specific to their job function and challenges they faced on the job. It was only through this effort that we could paint an accurate picture of the industry landscape and existing supporting and opposing forces in the area of data sharing. To determine reliability of results, we also checked for consistency of among responses. When substantial inconsistencies were identified, we went back to the interviewees for confirmation and clarification. When necessary, we performed more interviews to increase our sample size.

3.2.4 Develop Event Flow Network

Based on our learnings from these interviews, we could then develop an event flow network as introduced by Miles and Huberman (1994, pp.114) to map out the entire SuperSalad buying process from SuperSalad's warehouse to retailers' shelves. This diagram would identify all decision nodes and personnel participating at every stage of the ordering process. This exercise

could help identify system bottlenecks and determine the value that data-sharing brings to each

stage of the chain.

3.2.5 Identify Areas of Further Research

In areas where different interviewees offered opposing views, we would perform further research

to validate our findings. We would also explore other areas of interest identified during the

interviews.

3.3 Quantitative Data Collection and Analysis

To conduct quantitative data collection, we used *multiple case method* as recommended by Stake

(2006). We chose this method because it allows for comparison; specifically, we wanted to

compare supply chain performance between retailers that share data and those that don't.

Although Stake (2006) recommended no fewer than 4 cases and no more than 10 cases for the

best results, we were limited by the number of SuperSalad customers - three - who were willing

to share their data for this purpose.

Once we have identified subjects for our case study, we began by attempting to obtain store

operation data from a variety of retailers served by SuperSalad. Our main objective was to obtain

store level data for retailers on the opposite ends of the data sharing spectrum, as we would be

able to use this data to quantitatively analyze the differences data sharing makes to the perishable

goods supply chain. We found it challenging to obtain this data as most retailers either did not have the data we requested, or were unwilling to share sensitive retail data with us. However, by leveraging relationships SuperSalad management has with retailers, we were ultimately able to collect sufficient data for our purposes.

Three retailers, who will be referred to as Alpha, Beta and Gamma throughout this thesis, supplied us with the data we requested. Alpha's distribution center inventory is currently managed by SuperSalad using Vendor Managed Inventory, which is the name given to the process whereby the retailer shares inventory and sales data with the supplier, and effectively gives the supplier total responsibility for the inventory management process. Vendor managed inventory can be acknowledged as an advanced form of data sharing whereby the supplier not only has the data, but also has the ability to control the supply chain strategies for the downstream distribution of its products. Beta and Gamma manage their own inventory, and data is either not available or is not used by SuperSalad. We chose to use three retailers because Stake (2006) emphasized the importance of "triangulation", confirmation through multiple data points to provide the assurance that findings and interpretations are accurate and significant.

We requested data from the retailers in the hierarchy shown in Figure 2.
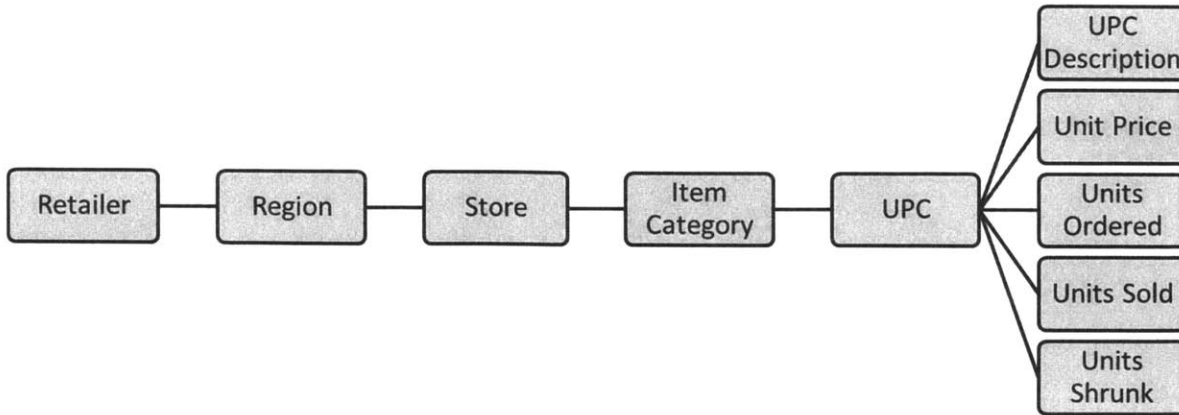
Figure 2. Requested Data Hierarchy

Our first set of analyses was an effort to determine the difference in shrinkage percentages between Alpha, whose inventory is managed by SuperSalad using VMI, and Beta and Gamma. We used statistical methods - specifically, average, standard deviation and coefficient of correlation - to establish whether significant differences in shrinkage existed between comparable stores in the VMI retailer's network and the non-sharing retailers' networks.

We then looked for correlations between shrinkage volume and other factors, such as Universal Product Code (UPC) volume, store volume, category volume and item cost. These analyses were done in an effort to determine if any of these factors are significantly related to shrinkage without the influence of data sharing.

3.4 Recommendations

Based on our findings from the interviews and quantitative analysis, we would provide our recommendations on data-sharing technology specifically for the perishable goods business. Our

recommendations would come in three tiers:

3.4.1 Does data-sharing improve overall supply chain performance?

We would summarize our findings to determine whether data-sharing would provide substantial performance improvement for the perishable goods supply chain. We would attempt to answer that question from both SuperSalad's and retailers' perspectives.

3.4.2 Do the benefits of data-sharing justify the investment?

We would make this recommendation by determining if the benefits - both financial and non-financial - justify the associated costs. This issue would be addressed from SuperSalad's perspective as well as from retailers'.

3.4.3 How can data-sharing enhance supplier negotiations?

We would provide recommendations on how SuperSalad should present this business case to their customers to convince them to become involved in this partnership. The focus here would be on the benefits retailers stand to gain from said collaboration.

4. DATA ANALYSIS

In this section we present the results attained through the methods outlined in the Methodology section. Interpretation of these results will be presented in the "Discussion" section.

4.1 Interviews - SuperSalad Internal Supply Chain Group

In this subsection, we will discuss some of the insights shared by SuperSalad's team on the topic of data sharing.

4.1.1 - Why Share Data?

Below we will explore the *perceived* benefits of data sharing, including building customer relationships, retaining customers, lowering out-of-stock and shrinkage, and reducing complexity for retailers.

4.1.1.1 Building Customer Relationships

Interviews with SuperSalad's customer-facing team have provided valuable insights on how they perceive data sharing can benefit retailers. They have met with other retailers who have been successful in using data sharing to build stronger business partnerships with vendors. In the modern business world where business relationships are cultivated largely on trust, they see data sharing as a way to ensure that retailers act in the best interest of the partnership.

In an interview with several category managers at SuperSalad, it was revealed that the vast majority of shrinkage takes place at the store, rather than at the distribution centers of SuperSalad or the retailers. For this reason, tackling the shrinkage rate at the store level will produce the biggest impact in reducing overall waste in the supply chain. In addition, they revealed that it is difficult for SuperSalad to understand the main drivers of shrinkage, as they have virtually no visibility of inventory at a store level. This is an area that requires further understanding, as identifying the source of the problem should not only help SuperSalad identify the bottleneck in their current operation, it should also help improve their relationship with retailers and their position in contract negotiation. Currently, store shrinkage rate affects SuperSalad's relationship with retailers, which in turn translates into more difficult price negotiations.

4.1.1.2 Customer Retention - Retailers

It is SuperSalad's belief that the more involved they are with retailers' ordering process, the harder it is for retailers to switch suppliers. In fact, they believe that the pain of switching supplier is so significant that offering effective inventory recommendations will give them an estimated 5% advantage in price, which is substantial in the grocery retail business where margins are often as low as 3%. In addition, in times of economic downturn when retailers are forced to manage costs closely, they are likely to see VMI as a strategic way of shifting risk back to the suppliers.

4.1.1.3 Customer Retention - End Consumers

Store-level shrinkage has a long-lasting effect on customers' perception of SuperSalad, and this type of damage to the brand typically takes a long time to recover from. For example, if a customer sees a poor quality or almost expired SuperSalad product on the shelf they will often turn to a competitor's product instead and, if satisfied by the product, are likely to continue to purchasing it in future. The cost of switching for a consumer is extremely low, hence this carries particular risk.

4.1.1.4 Lower Out-of-stock and Shrinkage

Ultimately, it is SuperSalad's belief that there is a tradeoff between shrinkage and lost sales. The average shrinkage rate in the fresh salad industry is believed to be around 14%. SuperSalad said that some of their retail customers were able to reduce shrinkage to 7% by ordering less, however, their view is that this contributed to an increase in lost sales. By enabling data sharing, SuperSalad believe that they will have greater visibility of store level activities, thus enabling them to make constructive recommendations to help customers reduce both shrinkage and lost sales.

4.1.1.5 Reduce Complexity for Retailers

Buyers at retail stores are responsible for a large number of SKU's and are often not able to develop in-depth expertise in every product group. This problem was clearly observed at one of SuperSalad's customer's stores, where SuperSalad's products experience the highest shrinkage

rates in June. Initially, SuperSalad found this counter-intuitive, as high demand in summer months typically increases inventory turn and lowers store-level shrinkage rate. Upon further investigation, SuperSalad found out that during the summer months, ordering managers at grocery stores typically have to handle a larger assortment of perishable products than in the winter. As a result, they are unable to devote as much time to ordering decisions for salads, which contributes to the higher shrinkage rate. SuperSalad believe that if they were involved in the inventory management and ordering process, they could reduce the burden on retailers and help manage an optimal inventory level.

4.1.2 - Why Not Share Data?

While most vendors and retailers are aware of the many perceived benefits of data sharing, some retailers remain reluctant to provide sales data to their vendors, we will define these retailers as non-sharers. The following section will detail some of SuperSalad's and retailer's concerns.

4.1.2.1 Confidentiality

First, most non-sharers are deterred by the idea of giving away confidential data that may disclose their operational and financial performance. In particular, some retailers are worried that they will lose their competitive advantage if sales data ends up in the hands of their competitors due to mishandling by the supplier. In the fresh produce business where margins are thin, rivalry is strong among competitors.

4.1.2.2 Lack of Resources and Funding

Because of the low margins in the fresh produce business, retailers must handle a particularly high volume and variety of products in order to be profitable. As such, most retailers do not have the budget, resources or time to set up the necessary processes for data sharing and its related functions. And even with strong management buy-in, most companies simply do not have the infrastructure to gather accurate internal data to share with vendors. In particular, shipment data is often difficult to obtain, especially when inventory is in transit. In addition, store level shrinkage rate - a key performance indicator in the perishable goods business - can only be obtained if items are scanned at every stage of the perishable supply chain, i.e. inbound, checkout and disposal. As a result, retailers must do a lot of work upfront to enable data collection and, subsequently, data sharing.

4.1.2.3 Skepticism

Ultimately, many retailers are still unconvinced that the perceived benefits of data sharing can outweigh the investment and effort required to implement such an effort. Some retailers are also unconvinced that suppliers would be able to glean better insights from retail information than the retailers themselves. Even when retail data is shared, suppliers may make recommendations for their own benefits rather than acting in the best interest of the whole supply chain. The inherent asymmetry of information at different stages of the supply chain were mentioned by Lee and Whang (1999), and for data sharing to be effective, incentives must be aligned. In summary, trust and a collaborative relationship are prerequisites for data sharing.

4.1.2.4 "Mind Your Own Business" Attitude

Lastly, in the highly competitive landscape that characterises the fresh produce industry, there is a prevalent "mind your own business" attitude, in which retailers are resistant to disclosing any data that may indicate their profitability, especially to their suppliers who are often viewed as competitors for profit rather than business partners. For example, vendors may be able to deduce competitive product margins from the data and price their own products accordingly. These insights on product margins may prevent retailers from obtaining competitive pricing and could jeopardize their relationships with other suppliers. This is further complicated by the fact that retailers work with hundreds of suppliers, therefore there is often little incentive for them to devote large amounts of time and effort to cultivating a business relationship with one particular supplier.

4.1.3 - Process Mapping - SuperSalad's Current Practice

SuperSalad currently has direct Electronic Data Interchange (EDI), a best practice of data sharing, links with Alpha, one of SuperSalad's biggest clients. This data is critical to the Vendor Managed Inventory (VMI) process that SuperSalad undertakes for Alpha. Using interview data we mapped out the VMI process, which is displayed graphically in Figure 3.

1. Data is transmitted daily from the retailer to SuperSalad, using EDI.

2. SuperSalad uploads this data into their order management system, which uses the store level ordering data and rolls it up to distribution center (DC) level.

3. The SuperSalad order management system uses the uploaded data to determine DC level inventory requirements and creates orders automatically.

4. The DCs receive orders daily or bi-daily, determined by the volume and capacity of each center. Orders are dispatched from SuperSalad's hubs to the retailer distribution centers and are received into stock.

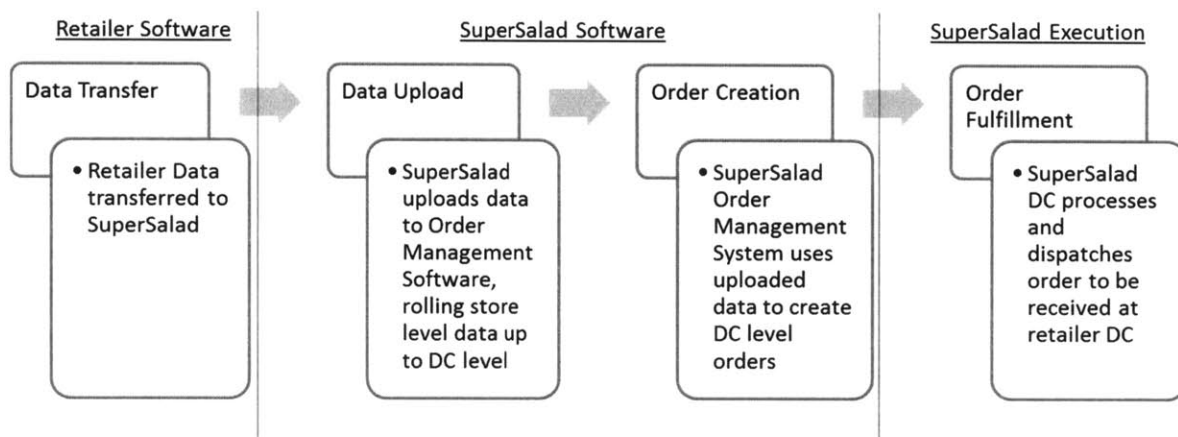| Retailer Software	SuperSalad Software		SuperSalad Execution
Data Transfer	**Data Upload**	**Order Creation**	**Order Fulfillment**
• Retailer Data transferred to SuperSalad	• SuperSalad uploads data to Order Management Software, rolling store level data up to DC level	• SuperSalad Order Management System uses uploaded data to create DC level orders	• SuperSalad DC processes and dispatches order to be received at retailer DC

Figure 3. Alpha – SuperSalad VMI Process

4.2 Retail Store Visits

To develop a first-hand understanding of the retailer replenishment process, our team visited a number of major grocery stores in Boston that carry SuperSalad's salad products. At certain stores within a large retailer that we will name Alpha, the ordering policy is facilitated by a

computer software system that manages in-house inventory. This system tracks inbound goods receipts, cashier activity and any other adjustments, and uses this data to generate recommendation for order quantities. Although store buyers have the option to override system's recommendations, they typically follow them due to the large number of SKUs they have to manage on a daily basis. At one Alpha store, SuperSalad products have an average remaining shelf life of 5 days and the shrinkage rate is reportedly approximately 20%. Shelves are stocked twice a day at this location. At a different Alpha store, which has a much higher volume, shelves are restocked 3-4 times per day. The average shelf life for SuperSalad products here is 8 days. The produce manager at this store claimed that this store experienced absolutely no shrinkage, however based on the quantitative data we were supplied this claim seemed highly unlikely.

At a separate retail chain, for which we will use the alias Delta, the average shelf life of SuperSalad products is about 7 days. The shrinkage rate for these products is reportedly about 20%, which is consistent with the rate Alpha had reported. Order quantities are determined by the employees that stock the shelves. Interestingly, the buyer indicated that out-of-stocks are never a problem at their store, even though it was clear that several products were out of stock at the time of our visit. This suggests that out-of-stock is perhaps not a performance indicator that store buyers keep close track of.

At another retailer, which we will give the alias Epsilon, SuperSalad products are not stocked, the retailer only sells salad products that are competitors to SuperSalad. The shrinkage rate for these products was also reportedly approximately 20%, also consistent with the other retailers. While the shelves were extremely well-stocked, the average shelf life for these products was less

than 2 days. Store employees replenished the shelves four times a day. The stocking employee we interviewed, who was also in charge of the ordering process, indicated that orders are made each day at 10am for next day delivery. Adding to the complexity of the order decision is the fact that an online grocery delivery service fulfills their orders from shelf inventory, which further increases the variability of demand. Similar to Alpha, Epsilon uses a computer system to keep track of store inventory. As a result, the ordering process is semi-automatic. In addition, the stocking employee noted that he had to manage thousands of items in the store and so did not have the luxury of time to develop ordering policies specific to each product group.

4.3 Quantitative Data Collection

Our next step was to collect retailer data from current SuperSalad customers. It was our intention that by comparing performances of retailers who share data and those who don't, we could accurately assess the benefits of data sharing, both for the retailers and for SuperSalad.

4.3.1 Beta Data

Beta is a grocery retail chain with over 180 stores across multiple states. Beta provided data for all 181 stores, including information such as number of units sold, total cases shipped, total units shipped, sell through percentage and shrinkage percentage; broken down by individual store and by SKU. Sell through percentage is defined to be total units sold divided by total units shipped. Shrinkage percentage is defined as (1 - sell_through %). This data was collected over 7 months.

SuperSalad does not currently have a data sharing partnership with Beta.

4.3.2 Gamma Data

The next chain we studied was Gamma, a grocery retailer and distributor with 516 stores. Gamma provided data for all 516 stores from its five subsidiaries, broken down to individual store location, category, and SKU. Information included units sold, sales revenue, store purchase quantities, and shrinkage units. This data was collected over 26 weeks.

SuperSalad does not currently have a data sharing partnership with Gamma.

4.3.2 Alpha Data

The last chain we studied was Alpha, which operates over 1,000 retail stores and is one of the largest grocery chains in the US. Each store offers approximately 2,900 products. Alpha provided us with data for 565 stores, information included units sold, sales revenue, store purchase quantities, and shrinkage units.

SuperSalad currently manages Alpha's distribution center orders using VMI.

Having collected all of the necessary information, we began to glean insights and came to several conclusions, which can be found in the following section.

5. DISCUSSION

In this section we provide interpretations of the results presented in the last section. We then draw conclusions from our findings that will eventually be the basis of our recommendations.

5.1 Interpretation of Data from SuperSalad customers

Using the quantitative data provided to us, we were able to draw direct comparisons between retailers. The main variable of interest here is product shrinkage. Specifically, we wanted to compare shrinkage rates between retailers that share data and those that don't. Also, we wanted to confirm that there is indeed a correlation between shrinkage and store volume, as it had been suggested by multiple SuperSalad employees in different interviews that higher SKU or store volume typically results in lower shrinkage rates. Table 1 contains a summary of the results that follow.

Summary of Shrinkage Rates	Beta	Gamma	Alpha
Data Sharing	No	No	Yes
Stores in Sample	181	516	565
Average Shrinkage Rate	15.27%	14.95%	12.77%
Store Volume – Shrinkage Correlation	-0.55	-0.24	-0.10
SKU Volume – Shrinkage Correlation	-0.73	-0.21	-0.22

Table 1. Summary of Shrinkage Rates and Correlations

5.1.1 Beta Data

Across all Beta stores, the average shrinkage rate by store is 15.27%, with a standard deviation

of 4.18%. Our findings from store visits indicated that shrinkage rate decreased as store volume increased, suggesting a negative correlation statistically. From Beta's store data, we found that the coefficient of correlation between store volume and shrinkage rate is -0.55, confirming that as store volume increases, shrinkage rate decreases. An absolute value of 0.55 confirms that there is some correlation between these two parameters.

The average shrinkage rate by SKU is also 15.27%, with a standard deviation of 15.59%. Interestingly, the coefficient of correlation between SKU volume and shrinkage rate is -0.76, showing a markedly higher correlation between these two parameters. This indicates that higher volume products are less likely to experience shrinkage.

5.1.2 Gamma Data

In the 516 Gamma stores we sampled, the average shrinkage rate is 14.95%, with a standard deviation of 6.2%. As is the case for Beta, a negative correlation between store volume and shrinkage rate exists, however at a lesser rate of -0.24.

The correlation between volume and shrinkage rate at a SKU level is similar to the correlation between volume and shrinkage rate at a store level, at -0.21.

It must be noted that several anomalies exist in the Gamma dataset, with some products even experiencing *positive shrinkage*, which is by nature impossible. This suggests that shrinkage data is not measured accurately by this retailer, as for positive shrinkage to exist other

adjustments have to have been interpreted as shrinkage. Gamma could benefit from measuring this data more accurately, as without realistic data it will be very difficult to understand and correct the issue.

5.1.3 Alpha Data

Across Alpha's sample of 565 stores, the average shrinkage rate by store is 12.77%, which is roughly consistent with the 20% reported to us in our interviews, with a standard deviation of 12.56%. From Alpha's store data, we found that the coefficient of correlation between store volume and shrinkage rate is -0.10.

The average shrinkage by SKU is 12.77%, with a standard deviation of 26.29%. Once again the correlation between SKU volume and shrinkage is higher than that of store volume and shrinkage, at -0.22.

5.1.4 Correlation between Volume and Shrinkage

The above results do not conclusively show a correlation between volume and shrinkage. In Beta's case, the correlation seems relatively strong; however in Alpha and Gammas' cases it shows a negligible relationship. This result does not allow us to confirm the suggestions from the interview process that indicated a definite correlation should exist.

Some of the strongest claims that the correlation exists were from produce managers at Alpha's

stores, however when we separated the top 10% of stores by volume for each chain, the effect was that Alpha's top 10% actually had a higher rate of shrinkage than the overall average for all Alpha stores. This was also true for Gamma. The only retailer that had a lower shrinkage rate for the top 10 stores when compared to the overall average was, as expected, Beta. These results can be seen in Table 2.

Shrinkage Rates of Top 10%	Beta	Gamma	Alpha
Data Sharing	No	No	Yes
Stores in Sample	181	516	565
Average Shrinkage Rate	15.27%	14.95%	12.77%
Shrinkage Rate of Top 10% Stores	11.23%	17.33%	15.10%
Shrinkage Rate of Top 10% SKUs	9.98%	14.16%	8.53%

Table 2. Shrinkage Rates of Top 10% of Stores and SKUs by Volume

To test whether the difference in average shrinkage rates between Alpha and Beta is statistically significant, we performed a statistical analysis by comparing the means of these two distributions using confidence interval testing. The formula and variables for the calculation is shown below.

$$\left[\mu_1 - \mu_2 - c\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \mu_1 - \mu_2 + c\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right]$$

(For confidence level $\beta = 98\%$, c = 2.326)

Equation 1. Comparison of the Mean of Two Distributions

	Alpha		Beta	
	By Store	By SKU	By Store	By SKU
Mean (μ)	12.8%	12.8%	15.3%	15.3%
Standard Deviation (s)	12.5%	26.2%	4.2%	15.6%
Sample Size (n)	338	39	181	43

Table 3. Variables for Confidence Interval Calculation for Entire Sample

Using the formula above, the 98% confidence interval for the difference between the two shrinkage distributions is [-4.3% , -0.7%] by store, and [-13.7% , 8.7%] by SKU.

The confidence interval test shows that when shrinkage data is sorted by stores, 98% of the time we see a reduction in shrinkage by anywhere between 0.7% to 4.3 % when data sharing is introduced. Assuming all else being equal, the result suggests that data sharing effectively reduces shrinkage 98% of the time.

The confidence interval test, however, does not allow us to draw the same conclusion when shrinkage is sorted by SKUs, as the upper bound of the confidence interval is a positive number, meaning that 98% of the time we see both a reduction and an increase in shrinkage when data sharing is introduced. There are two possible explanations. First, the inherent variation in shrinkage is higher when the data is sorted by SKU rather than by store. Second, there are more sample points by stores than by SKUs, and since the size of the confidence interval is inversely proportional to sample size, more sample points result in closer bounds on the confidence interval, meaning that we can more confidently identify the true effect of data sharing. Still, this

statistical analysis seems to suggest that data sharing has at least some effect on reducing shrinkage, especially on the store level.

It is interesting that measuring shrinkage by the Top 10% of stores and by the Top 10% of SKUs, in terms of volume, produced very different results. Our data was insufficient to fully understand the nature of these relationships; to do so would require an in depth analysis into the business and ordering processes of individual stores and chains that affect shrinkage, which was not in the scope of this thesis. Some potential factors include automated ordering systems, seasonal factors that were not accounted for in our data, geographical differences in demand and the fact that higher volume stores probably result in buyers having less time dedicated to individual salad products.

The relationship between SKU level shrinkage and volume seems more relevant, as all of the top 10 shrinkage percentages by SKU were lower than the overall averages, with Alpha and Beta being significantly lower. This shows that high volume products do in fact result in lower shrinkage rates, as these values were averaged out over all of the stores within our samples, thus removing the possibility of individual human errors, as well as variations in ordering processes and replenishment policies among the sample stores. This allows for a more direct comparison between volume and shrinkage.

When plotted graphically the correlations between store volumes and shrinkage rates become even more apparent, with Beta displaying a distinct downward trend in shrinkage in relation to volume, as shown in Figure 4. Alpha, however, shows that after a similar downward trend, the

rate of shrinkage actually begins to increase after a certain level of volume is reached, as shown in Figure 5. The blue lines in each figure display the actual data, with the red lines indicating the trends. The values on the horizontal axis are ranked by volume, from smallest to largest.
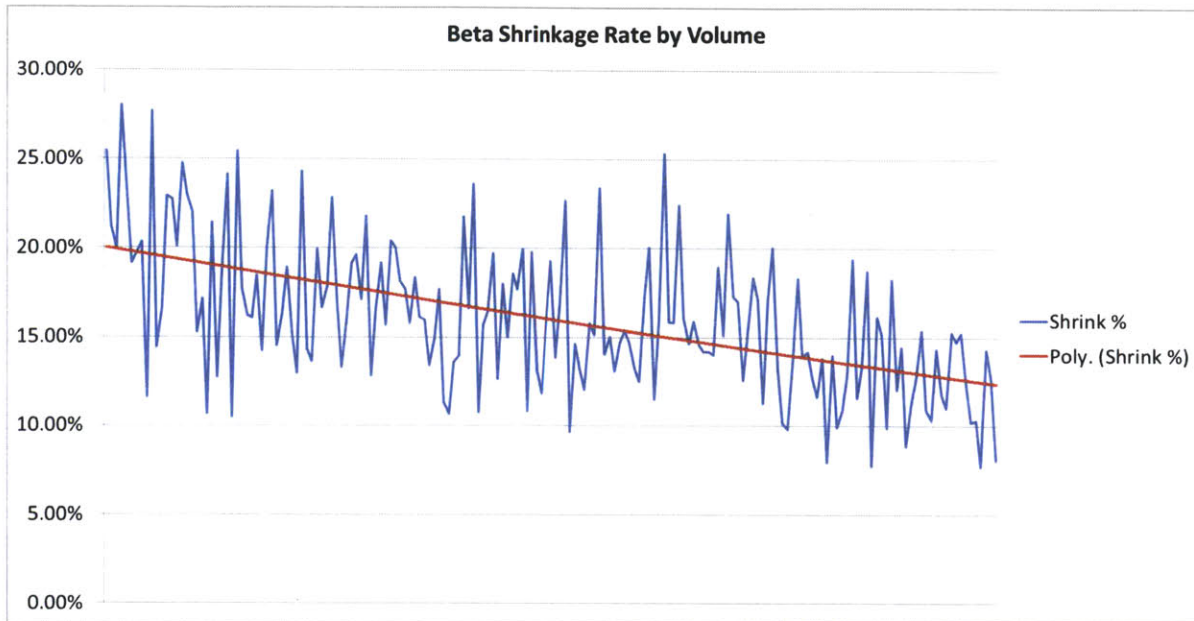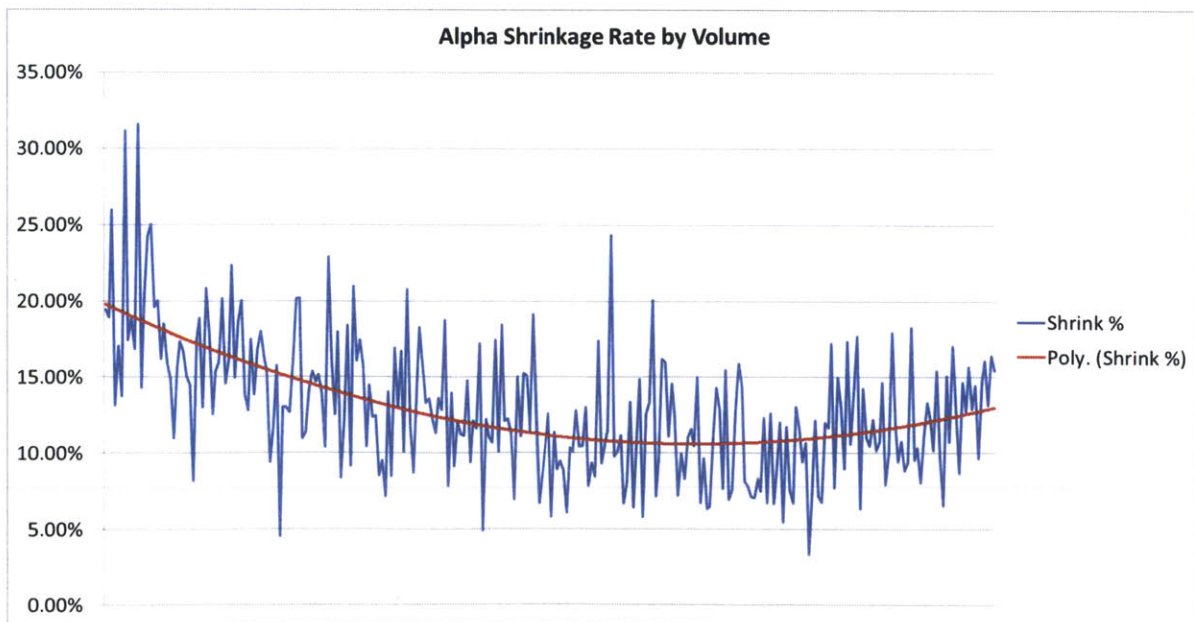

Figure 4. Beta Shrinkage Rate by Volume

Figure 5. Alpha Shrinkage Rate by Volume

It should also be noted that the largest Alpha stores in our sample processed significantly higher volumes than the largest Beta stores in our sample. This could indicate that once a store reaches a certain level of volume, the management of inventory becomes progressively more difficult. We attempted to test this theory by comparing the best thirty stores in terms of shrinkage in each sample, as this would remove any volume bias in the analysis. We also looked for the range of thirty consecutive shrinkage rate values in each sample that had the lowest average, which we will refer to as the *sweet spot*.

As expected, the resulting samples of Beta stores for both tests were primarily the highest volume stores in the full Beta sample. Alpha's samples were primarily concentrated around the minimum on the trend curve shown in Figure 5. Interestingly, when comparing the best thirty stores and sweet spots, Alpha had markedly lower shrinkage in both sets. This differs from the results when comparing the top thirty stores in terms of volume, as in those cases Beta's shrinkage rates were lower than Alpha's initially, however other than the first few values it is difficult to see a substantial difference between the two sets. These relationships are displayed graphically in Figures 6 , 7 and 8 below.
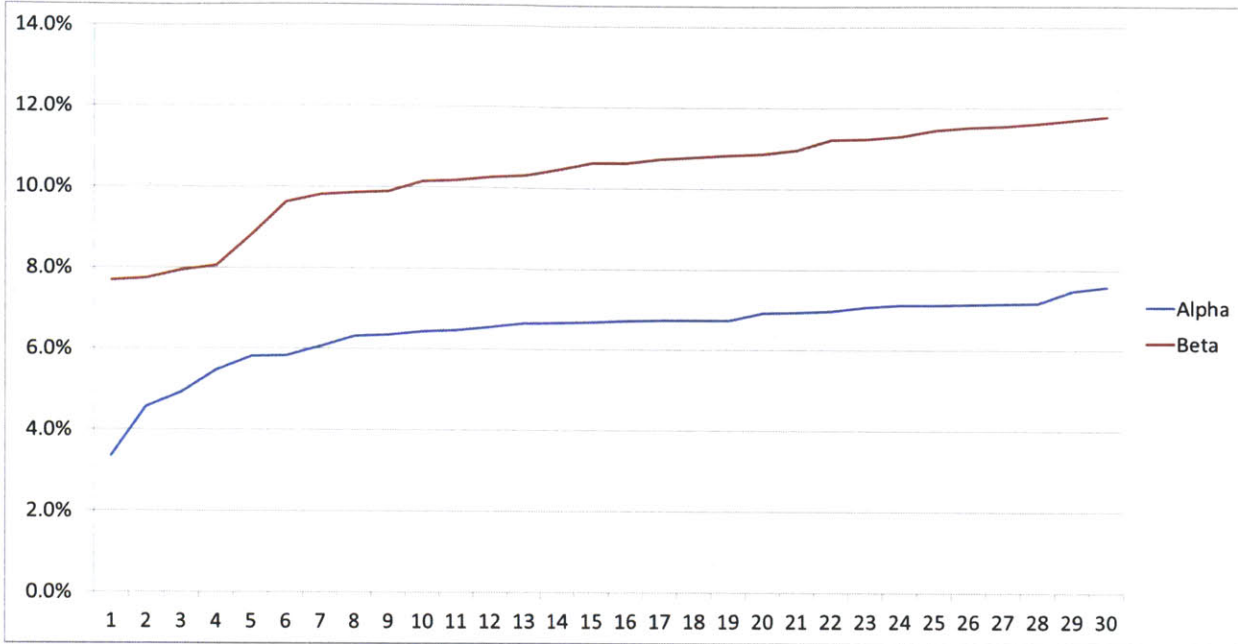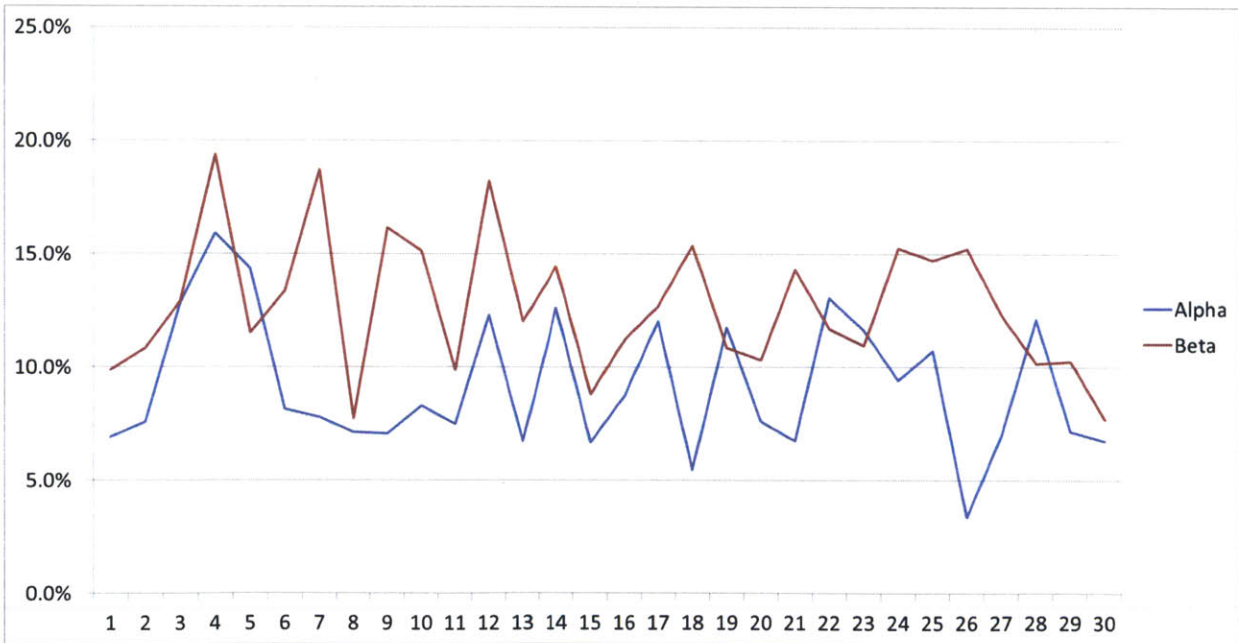
Figure 6. Shrinkage Rates of Best 30 of Alpha and Beta

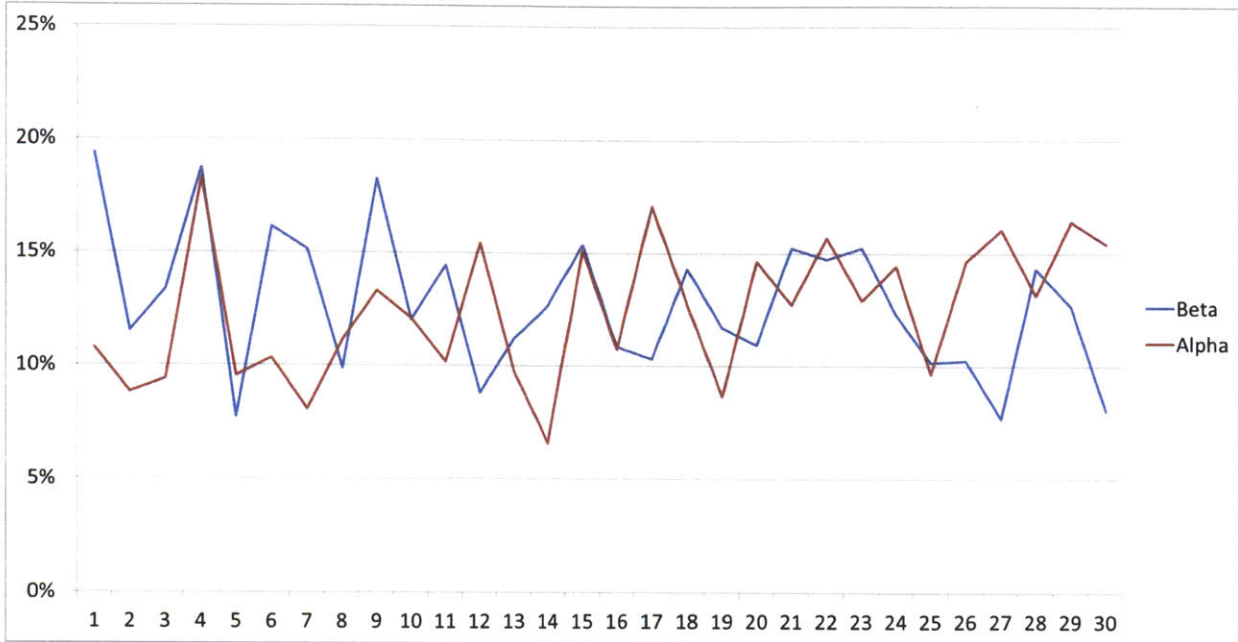*Figure 7. Shrinkage Rates of Sweet Spot 30 of Alpha and Beta*

Figure 8. Shrinkage Rates of Highest Volume 30 of Alpha and Beta

We also performed the same confidence interval analysis as in Table 3, for both the best stores and the *sweet spot* stores, to test for statistical significance. For the best stores, using the variables summarized in Table 4, the results show that at the 98% level, the confidence interval of difference in shrinkage is [-4.5% , -3.1%], suggesting that there is some correlation between data sharing and shrinkage. For the *sweet spot* stores, using the variable summarized in Table 5, the results show that at the 98% level, the confidence interval of difference in shrinkage is [-3.8% , -3.1%], also suggesting a correlation between data sharing and lower shrinkage around the *sweet spot*.

	Alpha	Beta
	By Store	By Store
Mean (μ)	6.5%	10.3%
Standard Deviation (s)	0.9%	1.2%
Sample Size (n)	30	30

Table 4. Variables for Confidence Interval Calculation for Best 30

	Alpha	Beta
	By Store	By Store
Mean (μ)	9.2%	12.7%
Standard Deviation (s)	3%	3.1%
Sample Size (n)	30	30

Table 5. Variables for Confidence Interval Calculation for Sweet Spot 30

This analysis shows that while the overall average shrink rates of Alpha and Beta are not significantly different, the best stores from Alpha, in terms of shrinkage, perform better than the best stores from Beta. It is also evident that the best stores of Alpha are not the highest volume stores, indicating that there may be a sweet spot in terms of store volume, beyond which shrinkage increases. This is also illustrated by the fact that the best stores from Alpha are actually relatively similar, in terms of sales volume, to those of Beta.

Thus, it appears that store level shrinkage can be attributed to a number of factors other than data sharing. It would appear that human error plays a role, as when the effect of this factor is reduced by aggregating the SKU level data, the results are improved. Store volume also plays a

role, with greater store volume resulting in reduced shrinkage up to a point, after which the volume begins to have a negative effect.

5.1.5 Data Sharing vs No Data Sharing

From the data presented above it appears that, although Alpha did perform better than Beta in some areas, there is no significant difference in overall shrinkage rates between data sharing retailers and non-sharers. This observation can be explained by either of the two following hypotheses: 1) data sharing provides no noticeable benefit in shrinkage for SuperSalad's high velocity supply chain, or 2) the existing data sharing processes are not properly used to reap optimal results, in terms of shrinkage. Our next step was to further investigate the latter.

To determine whether existing VMI data is being properly used, we evaluated the current ordering process.

The retailer distribution centers receive inventory as shown in the process outlined in section 4.1.3. Individual stores will then place orders from the distribution center inventory, the frequency of these order cycles vary by store volume. SuperSalad has no influence over this part of the process, they can only estimate what the stores may order using forecasting methods. Each store has buyers, who are responsible for a large number of products, to place orders manually. It was suggested that SuperSalad staff could have a better understanding of the demand of their products, however insufficient evidence exists to state that they would be able to manage the overall process better. Many factors can influence a store's sales, including product

placement, price changes, promotions on complementary products and other regional competitive factors, therefore it would be very difficult for centralized SuperSalad staff to keep track of such intricacies considering the number of stores their products are sold at.

Based on data collected from several interviews, it can be stated that the vast majority of product shrinkage takes place in the retailer stores, rather than at the distribution centers of SuperSalad or the retailers. The efficiency of distribution centers plays a significant role in this as the faster stock turns in those locations, the greater the remaining shelf life will be of products delivered to the stores. Based on the data gathered from a sample of stores, we conclude that, although exceptions exist, the distribution networks of our sample retailers are similar in efficiency, as indicated by the average remaining shelf life remaining upon store receipt.

Therefore, referring back to our second hypothesis above, our research indicates that the shrinkage issues are caused primarily by inefficiencies in the store ordering process, as indicated in Figure 9. SuperSalad's VMI process is efficient, however its effectiveness is limited by its lack of scope. Without having any influence in the processes downstream of the retailer DCs, SuperSalad's use of retailer data does not appear to make a significant difference in shrinkage.

Data from our interviews also revealed that several Alpha stores have implemented software that calculates order quantities and places orders for them automatically, and it was suggested that this has resulted in a noticeable reduction in shrinkage.
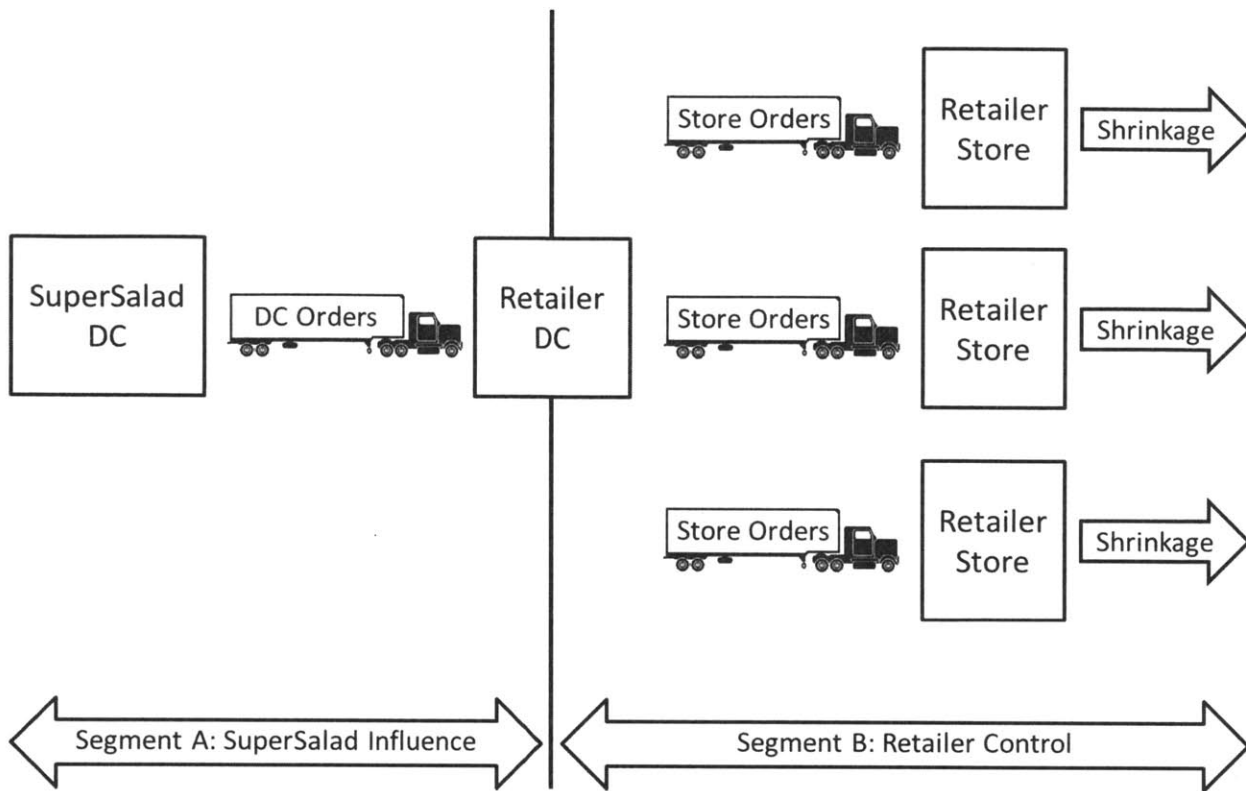
Figure 9. SuperSalad – Alpha Supply Chain Segments

Based on this data it appears that greater benefits exist in retailers improving their ordering

system at a store level, rather than suppliers investing in processes to improve ordering at the DC

level. The DC to store segment of the supply chain requires the most attention, as demand at an

individual store level is far more volatile than at the rolled up DC level. Improvements,

technology or otherwise, to store level ordering should result in a reduction in shrinkage and

reduce waste throughout the supply chain.

5.2 Data from Store Visits

As the data obtained from our store visits was partly subjective, it could only be used to validate

other data collected. These store visits, however, provided useful insight on actual ordering process, frequency and shelf replenishment practices. Despite the high perishability of salad, it turned out that most stores do not have specific inventory policy for this product group, which suggests that even if good inventory and ordering recommendations are made by SuperSalad, it is still unlikely that they will be adopted. In addition, even with data sharing and VMI, SuperSalad only has control over ordering at the DC level. SuperSalad has no control over ordering on store level, which is where most shrinkage takes place. Without control of the last step (from DC to store), SuperSalad cannot reduce store shrinkage even when retail data is available. Unless these two inadequacies in the current process are addressed, data sharing will not provide sufficient benefit to justify the investment.

We also found it surprising that most grocery store managers did not measure out-of-stock rates, and several also claimed that out-of-stock was consistently zero. These claims should be challenged, not only because of personal shopping experiences, but also because even as the store managers were making their claims we could see that some items were out of stock. This observation was telling in two ways. First, one of the benefits of data sharing is to reduce out-of-stocks. If store managers are unaware of shelves being stocked out, or if they do not have the physical manpower to monitor all shelves frequently enough, they are not going to realize any benefits brought about by data sharing. Even if data sharing allowed the stores to receive the correct amounts of inventory, the physical action of restocking the shelves regularly is still a critical component in the process. In addition, without the proper mechanism in place to measure current and future out-of-stock level, it would be difficult to quantify the improvement brought about by data sharing, in order to justify the investment of the technology in the first place.

It is easy to see why retailers find it difficult to measure out-of-stocks, compared to shrinkage. Most grocery stores nowadays have built-in mechanism to track shrinkage by scanning every inbound, sold and disposed item. Lost sales due to stock-outs, on the other hand, are much more difficult to quantify. When customers are unable to find their desired product on the store shelf, the obvious thing to do is to look for the next closest substitute. Rarely would they go out of their way to bring attention to out-of-stock items to store employees. Even when customers do request an item, the priority for the store employee at that point is to restock the item rather than to record the incident of stock-out. In addition, compared to other types of retail stores grocery stores usually have much fewer employees on the floor. Most on-the-floor employees at grocery stores are usually already engaged in other tasks such as restocking and cleaning, which diminishes their availability to help customers and hence decreases the likelihood of customers turning to them to report a stock-out. Also, not every shopper walks into grocery stores with a clear agenda. People tend to buy what appeals to them at that particular point in time, and, more importantly, what is available on the shelf on that day. This is particularly true in the fresh produce business where freshness and availability are top priorities for consumers.

6. CONCLUSION

With VMI, SuperSalad can help optimize ordering policy at DC level. However, because of the high volume of SuperSalad product at the DC level, there is more room for error in DC inventory management. At the store level, where most shrinkage occurs, SuperSalad does not have the power to influence ordering even when retail data is available. Hence, while our quantitative analysis demonstrates that data sharing improves performance in certain areas, it remains inconclusive whether it justifies the expense of implementing such a process.

Many arguments exist to justify data sharing in other ways, such as the benefits of closer relations with customers, promotional advantages, and being able to more accurately forecast demand for internal production. The challenge is that these points remain difficult to quantify, and make the business case of promoting data sharing in this environment more intuitive rather than financially motivated.

Data sharing is without a doubt a major component in the future of supply chain management, however the research presented in our thesis shows that, in certain cases, all functional barriers, as shown by the segmentation in Figure 9, must be removed for the process to be fully effective. In almost all modern supply chains, inventory is stored in multiple locations along the chain, and unless all of those locations are coordinated, the supply chain will never be optimal. This issue is particularly relevant in cases such as SuperSalad's, as the highly perishable nature of the products result in significantly greater risk throughout the supply chain.

CPG companies stand to benefit a great deal from data sharing, even if their influence only

reaches as far as the retailer DCs, as their products do not suffer from the same types of overage

costs, such as shrinkage, as fresh produce does. CPG companies are primarily concerned with

underage costs, as their products are highly competitive and any stock out could result in

consumers moving to competitor brands. The added complexity of products with very short shelf

life makes managing a fresh produce supply chain significantly different, as the risks associated

with underage must be balanced against the penalties incurred with overage. For these reasons,

the model used for CPG supply chains cannot be applied directly to fresh produce supply chains,

as the objectives are fundamentally different.

Therefore, if suppliers and retailers in a fresh produce supply chain are serious about eliminating

waste, a more coordinated effort will be necessary all the way through the chain. There should

not be functional barriers between supplier and retailer responsibilities, both parties should work

together to manage inventory all the way from production to final store locations. Financial

models, such as revenue or cost sharing, encourage this coordination, which will ultimately

eliminate a massive amount of waste from the overall supply chain.

Our final recommendations are as follows. If retailers are unwilling to allow suppliers to

influence the final movement of inventory from DCs to stores, data sharing will not fully solve

the issue of shrinkage. Significant barriers to retailers agreeing to such a process exist, such as

the decentralized nature of store ordering and the fact that stores carry inventory from hundreds

of suppliers in hundreds of categories, most with different complexities to those of fresh produce.

However, the unique nature of highly perishable goods requires additional consideration, and the

extra investment required to coordinate this process could be justifiable if offset against the

current shrinkage costs retailers are experiencing in this category. A more centralized, software

driven process used by retailers has been shown to decrease shrinkage; if this can be rolled out

further in addition to allowing suppliers to understand and influence the automated ordering

process, shrinkage should be significantly reduced.

7. REFERENCES

Agarwal, P. (2009), "Supply Chain Management - Cracking the bullwhip effect", pp.4-7

Buzby, J., Wells, H.F., Axtman, B., Mickey, J. (2009), "Supermarket Loss Estimates for Fresh Fruit, Vegetables, Meat, Poultry, and Seafood and Their Use in the ERS Loss-Adjusted Food Availability Data", Economic Information Bulletin No. 44, pp. 12

Croson, R. and Donohue, K. (2003), "Impact of POS Data Sharing on Supply Chain Management: An Experimental Study", Production and Operations Management, Vol. 12, No. 1, Spring 2003

Duffy, R. and Fearne, A. (2004), "Buyer-Supplier Relationships: An Investigation of Moderating Factors on the Development of Partnership Characteristics and Performance", International Food and Agribusiness Management Review, Volume 7, Issue 2, 2004

Dul, J. and Hak, T., *Case Study Methodology in Business Research* (Burlington, MA: Butterworth-Heinemann 2008), p.61-65

ECR Europe (2000), "ECR Community", www.ecr-all.org, accessed March 2012

Fiddis, C., *Manufacturer-Retailer Relationships In the Food and Drink Industry: Strategies and Tactics in the Battle for Power* (London, UK: Financial Times Retail and Consumer Publications

1997)

Grant, J. (2009), "Kimberly-Clark Leverages POS Data To Help Alpha Reduce Out of Stocks", CPGmatters, February 2009

Grean, M. and Shaw, M.J. (2003), "Supply-Chain Integration through Information Sharing: Channel Partnership between Wal-mart and Procter & Gamble", pp. 2-21

Hudock, L. and Vowell, Julie (2011), "Retail Data - How Kimberly-Clark Uses It," study presented at a meeting at SuperSalad office, California, November 2011

Lee, H. and Whang, S. (1999), "Decentralized Multi-Echelon Supply Chains: Incentives and Information", Management Science, Vol. 45, No. 5, May 1999, pp. 633-640

Lee, H.L., So, C.S., Tang, C.S. (2000), "The Value of Information Sharing in a Two-Level Supply Chain", Management Science, Vol. 46, No. 5, May 2000, pp. 626-643

Lewis-Beck, M. S., A. Bryman, and T. F. Liao., *The Sage Encyclopedia of Social Science Research Methods* (Thousand Oaks, CA: Sage Publications, 2004)

Li, L. and Zhang, H. (2008), "Confidentiality and Information Sharing in Supply Chain Coordination", Management Science, Vol. 54, No. 8, August 2008, pp. 1467-1481

Mentzer, J., DeWitt, W., Keebler, J. S., Min, S., Nix. N. W., Smith, C. D., and Zacharia, Z. G.

(2001), "Defining Supply Chain Management", Journal of Business Logistics, Vol. 22, No. 2, 2001, pp. 7

Miles, M. B., and Huberman, A. M., *Qualitative Data Analysis* (Thousand Oaks, CA: Sage Publications, 1994), p.114

Patton, M. Q., *Qualitative Research & Evaluation Methods* (Thousand Oaks, CA: Sage Publications, 2002), p.244

Rubin, H. J., and Rubin, I. S., *Qualitative Interviewing* (Thousand Oaks, CA: Sage Publications, 2002), p.134-151

Stake, R. E., *Multiple Case Study Analysis* (New York, NY: The Guilford Press, 2006), p.17-36

Vision Chain Inc, CROSSMARK, Teradata Corporation (2009), "Retailer-Direct Data Report", Grocery Manufacturers Association, pp. 1-27

Wakolbinger, T. and Cruz, J.M. (2010), "Supply Chain Distruption Risk Management Through Strategic Information Acquisition and Sharing and Risk-sharing Contracts", International Journal of Production Research, Vol. 49, No. 13, pp. 4063-4084

Weiss, R. S., *Learn from Strangers* (New York, NY: The Free Press, 1994), p.1-15, 153-159

Zhao, Y. (2002), "The Impact of Information Sharing on Supply Chain Performance" (PhD.

diss., Northwestern University, 2002), pp. 108-110