Modeling Order Guidelines to Improve Truckload Utilization

by

Jaya Banik
B.Sc. & M.Sc. Statistics, Delhi University, 2006

and

Kyle Rinehart
B.A. International Studies, The Ohio State University, 2006

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

June 2011

Signatures of Authors..
Master of Engineering in Logistics Program, Engineering Systems Division
May 6, 2011

Certified by..
Dr. Jarrod Goentzel
Executive Director, Master of Engineering in Logistics Program
Thesis Supervisor

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

Modeling Order Guidelines to Improve Truckload Utilization

by

Jaya Banik and J. Kyle Rinehart

Submitted to the Engineering Systems Division on May 6, 2011
in partial fulfillment of the requirements for the degree of
Master of Engineering in Logistics

ABSTRACT

Freight vehicle capacity, whether it be road, ocean or air transport, is highly under-utilized. This under-utilization presents an opportunity for companies to reduce their vehicular traffic and reduce their carbon footprint through greater supply chain integration. This thesis describes the impact of ordering guidelines on the transport efficiency of a large firm and how those guidelines and associated practices can be changed in order to gain better efficiency. To that end, we present three recommendations on improving the guidelines based on the shipment data analysis.

First, we discuss the redundancy of one of the company's fill metrics based on a scatterplot analysis and a chi-square independence test. Second, we explore the impact of using linear programming to allocate SKUs to different shipment, highlighting the reduction in the number of shipments through better truck mixing. Finally, we divide the SKUs into three groups: cube-constrained, neutral, and weight-constrained. Based on this segmentation, we present a basic model that mixes different SKUs and helps a shipment to achieve a much higher utilization rate. The application of the last two findings can be further explored to address under-utilization in freight carriers across different industries.

Thesis Supervisor: Dr. Jarrod Goentzel

Title: Executive Director, Master of Engineering in Logistics Program

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

According to a study by the World Economic Forum (2009), an astonishing 24% of all freight vehicles in the European Union (EU) run empty. The average load factor for the remaining vehicles is only about 57%, resulting in just 43% overall efficiency for European trucks. While the aggregate trucking statistics for the United States are not readily available, we assume that they are not much better than Europe. The implication, then, is that a great opportunity lies in optimally utilizing the trucking capacity available.

This thesis project seeks to improve truckload utilization in the United States for a large multi-national firm. This company will be referred to as "Shipper." The basis of the project stems from the sub-optimal vehicle loading statistics reported by Shipper. As such, Shipper wants to explore the various dimensions of how it can achieve better utilization for its trucking fleet and significantly reduce the amount of shipments required over a given time frame. Given the current pressure on companies to have not only innovative supply chains but also green ones, Shipper can obviously reduce its carbon footprint by decreasing vehicular traffic. Furthermore, considering its enormous scale of operations, even a marginal improvement in Shipper's operations, such as saving one truck per week, can lead to huge annual savings.

The scope of this thesis is primarily focused on trucks. However, other modes of goods transport, such as freight planes and shipping containers, can benefit from this research. Freight planes, for example, have volume and weight utilization metrics just as trucks do. In that sense, planes can benefit from a density based mix; but, of course, aircraft would have to take center of gravity, and other possible variables, into account in order to build on this paper's findings.

The different parameters used to measure truckload fill are: weight fill, cube fill, and floor fill. They are defined as follows:

$$Weight\ Fill = \frac{Net\ weight\ of\ the\ shipment\ (excluding\ pallets, fillers/racks)}{Maximum\ legal\ weight\ per\ each\ specfic\ lane\ \&\ vehicle\ type} \tag{1.1}$$

$$Cube\ Fill = \frac{Net\ volume\ of\ the\ shipment\ (excluding\ pallets, fillers/racks)}{Maximum\ volume\ per\ vehicle\ type} \tag{1.2}$$

$$Floor\ Fill = \frac{Actual\ number\ of\ pallets\ in\ the\ shipment}{Maximum\ number\ of\ pallets\ per\ region} \tag{1.3}$$

Shipper has set their benchmarks for the above three parameters based on the dimensions of a standard trailer of a certain length and a given region of operations. Presently, Shipper's overall statistics for cube fill and weight fill are below the optimal level. Its future vision is to achieve a much greater utilization rate for both parameters. It is worth mentioning that all three parameters are given equal weight in Shipper's ordering guidelines. To elaborate, Shipper declares a truck full if any one of these parameters is met. Our goal is to go beyond an emphasis on any one metric, and optimize all three simultaneously. While simultaneous optimization is the best solution, it is, of course, more complex and difficult to achieve.

To improve truckload utilization, three different research areas can be explored: fill optimization, volume efficiency, and transport, as shown in Figure 1. They can be thought of as loading, pre-loading, and post-loading, respectively. Fill optimization focuses on achieving better weight and volume fill for a truck, and has more direct impact on the truck loading decisions. Volume efficiency focuses on aspects of product design and packaging. These changes are more fundamental and, in most companies, are more related to marketing than operations. Lastly, transport includes various aspects of the actual operation of trucks, such as network optimization and drivers' efficiency. These measures are only applicable to post-loading and are more focused on obtaining better driving efficiency. Amongst these three, we have decided to focus on fill optimization because, more than the other two, it provides us the opportunity of discovering immediate, as opposed to long-term, improvements.

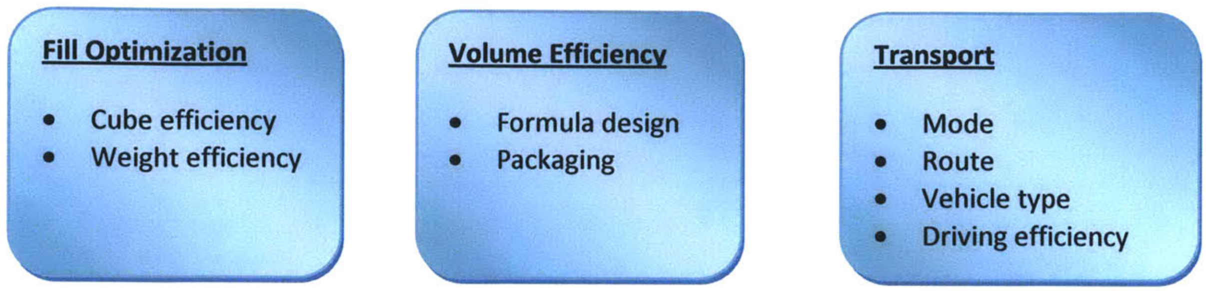

Figure 1: Areas that impact truckload efficiency

To understand fill optimization, one must understand current ordering guidelines and examine how truckload utilization can be improved by re-modeling those guidelines. Next, our literature review explores the importance of the order fulfillment process and its link to transportation operations. In our methods section, we discuss the various statistical and mathematical tools that we use to analyze the data set provided by Shipper. Our findings highlight the major insights gained through applying these tools to Shipper's problems. It also lists an algorithm that we develop in order to better mix SKUs to fill trucks more fully.

2. LITERATURE REVIEW

The literature concerning how companies model their ordering guidelines, in practice, is limited. Given the competitive nature of business, making such information public may give advantages to one's competitors. To overcome this barrier, we focus on literature that explains the foundations of order guidelines, including the order fulfillment process, its design, and a summary of academic research on how to formulate an ordering strategy. Lastly, it is important to note that not all of the models explained in this literature review can be implemented by Shipper due to some key assumptions and restrictions, which will be explained in the findings section.

2.1 The Order Fulfillment Process

While academic researchers disagree on the contemporary definition of the order fulfillment process (OFP), they all agree that its core goal should be how to best meet a given customer service level. For example, Kritchanchai and MacCarthy (1999), and Turner (2002), recognize OFP, and customer service level, as one of the core business processes in any organization. And while OFP varies based on unique business requirements, such as make-to-order or make-to-stock requirements, it always holds two common objectives:

1) To deliver products that meet customers' satisfaction in terms of price, quality, delivery time, and location. (Zhang, Jiao & Ma, 2010)
2) To obtain greater flexibility in managing uncertainties that originate from internal & external environments. (Christopher, 1992; Goldman, 1995)

The above steps describe the basics of OFP, but not its contemporary form. OFP has changed tremendously across different industries over the past few decades. Beginning as a simple process of receiving orders and fulfilling them, given a specified customer service level, OFP has now evolved into a highly specialized field. Indeed, new focus by supply chain experts has sprung up around the order fulfillment process seeking to integrate supply chains across multiple echelons and optimize the fulfillment of orders through collaborative channel planning. The bargaining power in the supply chain appears to have shifted from manufacturers' hands into retailers' hands. Davis-Sramek, Germain, and Stank (2010) summarize this shift as, "…the increased ability of retailers to influence consumer purchases, suggesting that manufacturers should understand not only consumer perceptions of delivery service, but also retailer perceptions." (np)

2.2 Designing the Order Fulfillment Process

Zhang, Jiao and Ma (2009) described the OFP process as, "…a process that starts with the receiving of customers' orders and ends with delivering final products," thus including all the activities involved between these two end points. An and Srethapakdi (2006), in their thesis report, describe order fulfillment, and the aforementioned activities between end points, as summarized in Figure 2.

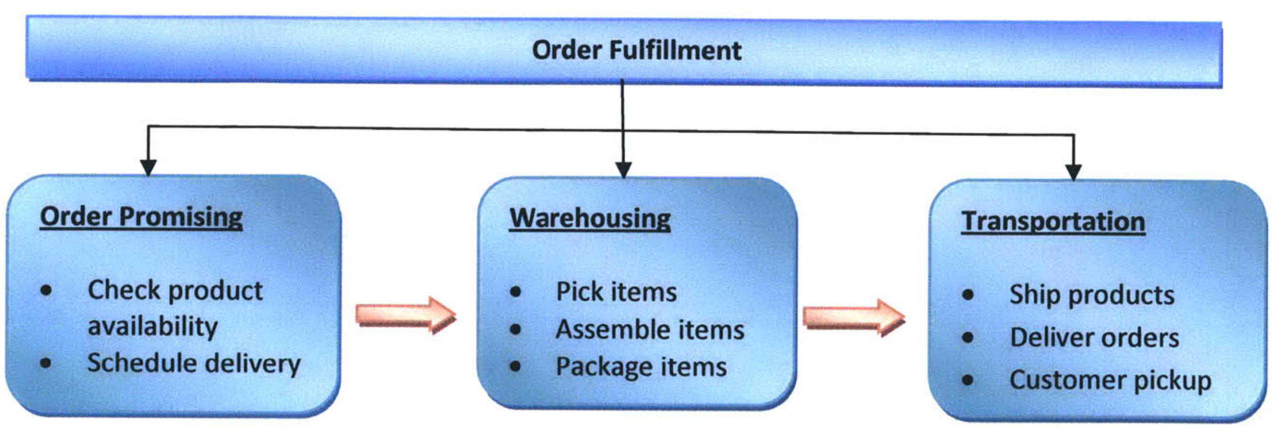

Figure 2: Basic Order Fulfillment Process (based on An and Srethapakdi, 2006)

Figure 2 demonstrates the basic design of the OFP. As one can see, the overall process influences other supply chain activities, such as warehouse management and transportation. Our thesis focuses on the beginning step in the order fulfillment process, order promising. Specifically, we are interested in OFP's influence on transportation, and how these order promising guidelines can be improved in order to better utilize trucks. In tackling this problem, we are fortunate to have Shipper, a large multi-national company, to work with due to their wide-range of products. This greater SKU variety gives us more potential combinations for mixing trucks, and subsequently, greater latitude in exploring different solutions for optimizing Shipper's shipping efficiency.

To achieve greater efficiency, we need to understand some common operating heuristics, since, as mentioned earlier, they influence OFP execution. Olhager (2010) introduces the term "customer order decoupling point" (COPD), which helps to illuminate operating heuristics. Olhager defines COPD as, "...the point in the value chain for a product, where the product is linked to a specific customer order" (p.1). The CODP seeks to define a demarcation line between the upstream and downstream processes in a supply chain. For example, upstream processes are generally make-to-stock assembly lines focused on low costs and predictable demand. Meanwhile, downstream processes are more likely to be make-to-order with higher safety stocks necessitated by unpredictable demand (Olhager, 2010). Every company has different operating heuristics depending on whether they choose to be agile or lean. This decision leads to different COPDs in the supply chain, as shown in Figure 3.

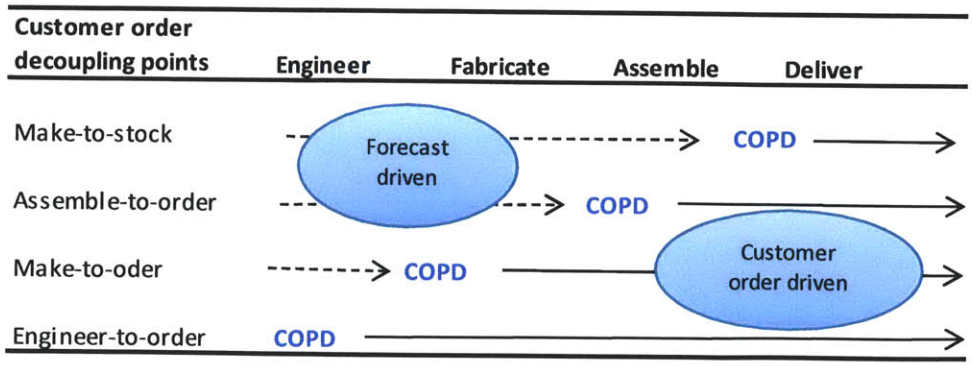

Figure 3: Different customer order decoupling points (based on Sharman, 1984)

For a flexible supply chain, initiatives should be focused on downstream operations whereas, for a company striving to be lean, more focus should be upstream of the CODP (Olhager, 2010). Basically, Olhager is emphasizing that while supply chain integration is important, it is also imperative to know where and how to draw the line between different supply chain functions and echelons. Shipper runs a menu of operations ranging from make-to stock to make-to-order. For its make-to-stock operations, it can achieve greater flexibility in its supply chain through better control of its downstream operations. This includes the important step of product delivery. In short, better control of its transportation activities can help Shipper become more flexible and strengthen its OFP.

2.3 Ordering Strategy

Apart from operating heuristics, supply chains also differ based on the products they order. Xiao, Jin, Chen, Shi, and Xie (2009) explore the ordering decision for a seasonal or perishable product in a three-stage supply chain, which consists of one retailer, one manufacturer, and one subcontractor. Through the use of a game theory model, they explore the varied impact of demand uncertainty on pricing, lead times, and order quantities. Their findings suggest that, "higher unit holding cost implies a lower optimal lead-time and order quantity while [increasing] higher unit-wholesale prices; [thus,] the basic demand uncertainty increases the optimal lead-time and order quantity while [it] decreases the unit-wholesale prices" (p. 840). Shipper can definitely incorporate the latter finding for meeting uncertain demand for its seasonal goods. Knowing Shipper can serve customers through longer lead-times, the firm can mix its orders better in order to achieve better truckload utilization.

Webster and Weng (2007) develop a model for ordering policy in a two-echelon fashion products supply chain. They develop two scenarios. One scenario describes a manufacturer-controlled chain while the other describes a distributor-controlled chain. The manufacturer-controlled scenario is where the manufacturer controls the supply chain stocking decisions and bears the risk of overstocking costs. The distributor-controlled scenario works in the opposite direction. Their paper suggests that the overall profit of the supply chain is maximized when the process is distributor-controlled, contrary to the popular notion that manufacturer-controlled is better.

We also reviewed three works concerning supply chain integration and cooperation and their impact on pricing strategy, rather than ordering strategy. The literature speaks to a well-known theme among supply chain professionals that cooperation and integration can lead to a profit sharing situation where both the upstream and downstream suppliers win (Yan & Wang, 2010). This idea involves game theory as it defies the conventional notion that suppliers are competing with each other for a finite portion of profits. In fact, the idea is that suppliers can push the boundaries of profitability for both through a collaborative relationship. Yan and Wang (2010) describe this theory for retailers, stating, "When the giant retailer adopts a noncooperative structure to maximize its own profit, then there exists optimal service level and pricing strategy for the giant retailer." (p. 63)

Finally, we last order fulfillment process piece of literature we reviewed is the bin packing problem. The bin packing problem is an important problem in operations research with extensive applications to logistics decisions. In logistics, the bin packing problem helps in the loading of freight by primarily packing a finite number of items with the least number of bins (Liu, Tan, Huang, Goh & Ho, 2007). The bin packing problem is based on the principle of running various possible combinations until the optimum is reached. There are often conflicting criteria in this optimization process. In the case of Shipper, these constraints are floor position, volume, and weight. In our findings, we employ a simple linear program example of the bin packing problem to optimize the weight of a truck, given various constraints.

Based on our literature review, there is a lack of academic research in practical utilization guidelines as set by various players in the shipping industry. The focus of our research, then, is to understand and analyze Shipper's ordering guidelines and benchmarks in order to help the company improve its fill optimization. To approach this problem, we base our analysis on developing a basic algorithm that can better mix SKUs and trucks.

3. METHODS

This thesis' data set consists of a sample shipment data of Shipper, which covers a certain time period, with extensive shipment details such as SKU information, origin-destination, weight, volume, etc. The data set is substantial, but it is not so large that we cannot analyze it using Excel. Specifically, we use Excel in conjunction with two main tools: statistical reasoning and mathematical programming. Statistical reasoning is a broad discipline. With concern to this thesis, statistical reasoning encompasses data sampling, scatterplots, and chi-square independence tests. For mathematical programming, we focus specifically on Excel's Solver add-in in order to formulate linear and integer programming models.

3.1 Statistical Reasoning

At the core of statistical reasoning, and at its most basic form, is summary statistics. We use summary statistics to obtain an overall picture of the data set before diving into specifics. Examples include variable averages, percentiles, and graphs. Looking at these numbers, in the beginning, helps us to fully understand the problem, and gain a sense of intuition about the data.

3.1.1 Scatterplot analysis

One basic technique we employ is the scatterplot. A scatterplot simply diagrams data points onto a two-axis graph containing one variable on the x-axis and the other on the y-axis. It is a useful way to display data in order to visually interpret relationships between two variables. Scatterplots are naturally associated with correlation. We use correlation in this work in order to examine quantitative relationships between different fill metrics. Correlation is performed on two variables and produces a rational number ranging between -1 and 1, a negative or positive correlation, respectively. A negative correlation means that as one variable increases, the other decreases. A positive correlation, then, means that both variables vary in the same direction simultaneously. No correlation, or numbers hovering around 0, asserts little to no relationship between variables. Variables range between 1 and -1 and are rarely ever perfectly correlated. Finally, we are able to display correlation in table form in order to look at collinearity among more than two variables. Collinearity is defined as two or more variables having similar high correlations.

3.1.2 Chi Square Independence Test

A chi-square independence test is another statistical tool for displaying a relationship between two variables. Unlike correlation, an independence test attempts to show whether two variables vary independently of one another. The input to this test is a contingency table. A contingency table has x rows and x columns. The test then calculates the expected values for each row and column based on equal probabilities. If the actual values match the expected values, the variables are dependent. If not, the variables are independent of one another, meaning they do not affect one another. The chi-square independence test is also very useful because it computes confidence levels for its output, which enables us to say with assigned confidence whether the test is conclusive or not.

3.2 Linear Programming

Mathematical programming is also used in this work. Specifically, we utilize linear and integer programming via Excel's Solver add-in. The Solver can handle up to 200 decision variables to be used in computing the aforementioned objective function. The Premium version of Solver extends the number of decision variable up to 1000, which can be used to apply our linear program to larger problems. Linear programming utilizes powerful algorithms in order to compute the optimal answer for a given problem. To solve a problem, one must first mathematically define an objective function. A linear objective function is one that needs to be maximized, minimized, or set to a specific value. The inputs to the objective function output are called decision variables. For linear programming to work, both the objective function and constraints have to be linear functions. A constraint is any bind placed on the decision variables or the objective function. For example, one constraint forces variables to be non-negative. Another constraint that has already been mentioned is requiring variables to assume integer values. These formulas are often quite simple. However, the computation required to complete all equations simultaneously is very difficult, given that so many constraints and decision variables figure into it.

Moving forward to our data analysis, we start with basic summary statistics. These statistics highlight the anomalies in the data set and help us to gauge the importance of each fill metric. Based on these initial findings, we then conduct an analysis for the floor position metric using a scatterplot and chi-square test of independence.

Based on the independence test, we explore excluding floor position as a measuring fill parameter and limiting its use to the exceptional cases or outliers. We develop a more strategic algorithm that can achieve better truck utilization. We base this strategic approach on two aspects: trucks and SKUs. First, we implement linear programming, which allocates an example SKU set to a particular truck, in order to better mix trucks at a tactical level. This example helps realize greater efficiency both in terms of fill statistics and a reduction in the number of trucks required for shipping the same SKU set. Secondly, for the efficient mixing of both we apply an algorithm based on the density of SKUs. This algorithm illuminates opportunities that can be further explored by applying it to a customer profile or geographic location.

4. FINDINGS

After analyzing Shipper's guidelines, we propose three solutions to how they can improve their logistics. First, we explore the redundancy of a particular measuring parameter. Second, through an example, we propose a linear programming model for mixing trucks and explain how that model can reduce Shipper's total number of shipments. Third, we explore how Shipper can mix different categories of SKUs, on the basis of cube and weight, in order to gain better utilization rates.

4.1 Metrics Evaluation

As previously mentioned, Shipper uses three metrics to determine whether a truck meets their minimum shipping requirements: volume, floor position, and weight. It is worth mentioning that Shipper uses a constant for measuring volume, with 0 defining a truck as completely empty, and 3750 defining a truck as completely full. Once a given metric is reached, the truck can be shipped. As an operating heuristic, each metric carries equal weight under this framework. After exploring the data, there is evidence that floor position is not as critical because it is largely incorporated in the volume metric. If a truck reaches 100% volume, then, obviously, all of that truck's floor positions are also full; and we observed that it is much less frequent to reach floor position capacity without reaching volume capacity. To analyze the relationship between volume and floor position metrics, we first perform a correlation test, build contingency tables, and then perform a chi square independence test.

4.1.1 Test for collinearity

Before running the chi-square independence test for any table, we first examined the data visually with a scatterplot, as shown in Figure 4. This scatterplot graphs volume against FP in order to examine the relationship between the two for every single shipment in the data set. One can see a clear upwards-diagonal trend, as well as clustering, in the data points. If the variables were completely independent of one another, there would be no clustering. Indeed, the correlation between variables is strongly positive, .69. This result is evidence of collinearity.

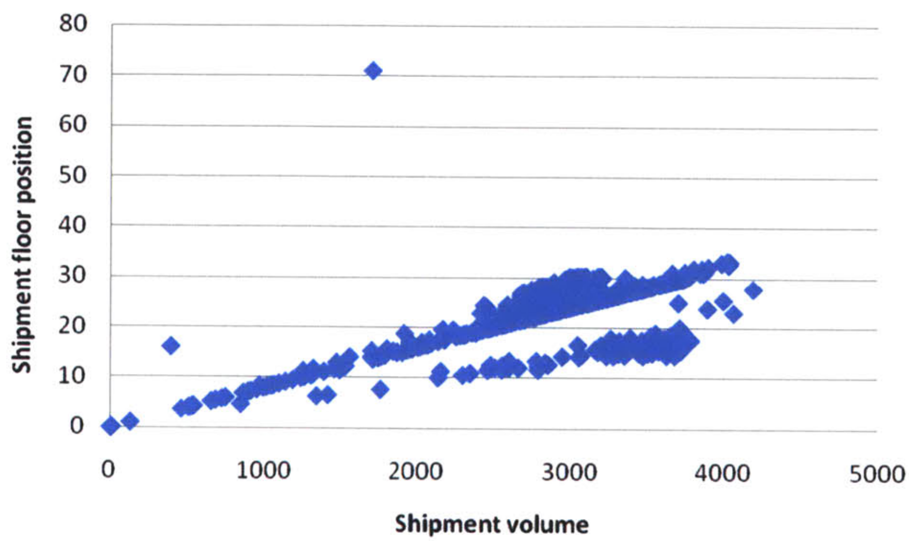

Figure 4: Scatterplot between volume and floor position

4.1.2 Contingency tables analysis

Next, a contingency table is required. In Figure 5, the variables are volume (Vol.) and floor position (FP). In each table, the values represent a count of shipments. In the 100% table, a value "hits Vol." if it is greater than or equal to 3750, the volume maximum. Likewise, a shipment "hits FP" if it is greater than 29.

Since the real world rarely achieves perfect metrics, we also test the numbers by relaxing the volume constraints in 1% increments from 100% to 97%. For example, 99% volume is defined as $3750 * .99 = 3712.5$. We hold FP constant at 100%, or greater than 29, because FP is the variable in question.

100%:

	Does not hit FP	Hits FP	Total
Does not hit Vol.	790	661	1451
Hits Vol.	6	521	527
Total	796	1182	1978

Parameters :

	Volume	FP
0	< 3750	<= 29
1	>= 3750	> 29

Within 99%:

	Does not hit FP	Hits FP	Total
Does not hit Vol.	779	224	1003
Hits Vol.	17	958	975
Total	796	1182	1978

Parameters :

	Volume	FP
0	< 3712.5	<= 29
1	>= 3712.5	> 29

Within 98%:

	Does not hit FP	Hits FP	Total
Does not hit Vol.	761	190	951
Hits Vol.	35	992	1027
Total	796	1182	1978

Parameters :

	Volume	FP
0	< 3675	<= 29
1	>= 3675	> 29

Within 97%:

	Does not hit FP	Hits FP	Total
Does not hit Vol.	742	178	232
Hits Vol.	54	1004	1058
Total	796	1182	1978

Parameters :

	Volume	FP
0	< 3637.5	<= 29
1	>= 3637.5	> 29

Figure 5: Contingency Tables

4.1.3 Chi square test of independence analysis

After seeing visual evidence for collinearity in Figure 4, we feel justified in conducting a chi-square test. Beginning with the 100% table, we can see that many points hit both metrics together, as well as missing both metrics simultaneously. For the dependence hypothesis to be true, these points have to be concentrated in the "both volume and FP" and "neither volume nor FP" corners of the contingency table. We used 1% increments in order to identify the best relaxation of the constraints. If one adds the bottom-left and top-right corners of each table, one sees that the "Within 98%" table is most in favor of the dependence hypothesis. It is important to note that the purpose in identifying this table is to relax the constraints just enough in order to see the most realistic, or truest, picture of volume versus floor position.

The bottom-left corner of each table is appropriately very low. It is intuitive that if volume is maximized at 100% full, then the pallets on in the floor positions of the truck should also be full. The six shipments in the 100% table of Figure 5 are anomalies. Each one of them is entirely composed of a particular type of product that is stacked, irregularly, in bags, instead of boxes. Because it is not comprised of boxes, every pallet varies. Since the pallets vary, they cannot be perfectly slotted into the neat geometrical patterns that Shipper's box-comprised pallets fit into. As such, the volume of the truck may meet Shipper's maximum, while the pallet numbers do not match up to a standard 100%.

Continuing with the 100% table, the majority of the values hit in the "neither" or "both" boxes. This is a good sign in favor of the dependence hypothesis. However, 661 shipments still comprise one third of the total 1,978 shipments. To begin the independence test, we compute the expected counts for each value:

Table 1: Expected counts table

Expected Counts	Does not hit FP	Hits FP
Does not hit Vol.	584	867
Hits Vol.	212	315

The expected count is computed as the row sum multiplied by the column sum, divided by the total number of shipments. 584, for example, is calculated as $\frac{(790+6)*(790+661)}{1978} = 584.$ What the independence test is doing is computing probabilities given the actual numbers. It then calculates the statistical distance of the expected numbers from the actual numbers. Statistical distance is computed as follows:

Table 2: Deviation from expected value table

Distance from Expected	Does not hit FP	Hits FP
Does not hit Vol.	73	49
Hits Vol.	200	135

This is the actual value minus the expected value squared, divided by the expected value. 73, for example, is calculated as $\frac{(790-584)^2}{584} = 73$. This calculation leads to the next step of the test, being the chi-square statistic. In simplistic terms, the chi-square test is testing the statistical distance of the expected values from the actual values. The farther the expected value is from the actual value, the less independent the variables are. It also calculates a p-value, which we use to conclude dependence or independence. A small p-value is evidence that the row and column variables are dependent. For the 100% table, the p-value is less than .0001. This means that we can say with greater than 99% confidence that the variables are not independent. Each of the other tables, from 99% to 96%, also produce the same less than .0001 p-value and 99% confidence level that the variables are not independent. Based on this evidence, we conclude that floor position should be further assessed as a critical metric for determining whether a truck should ship. Given the statistical evidence that it is not independent of volume, it may be possible to remove this metric and utilize an ordering policy based on only two capacity metrics, which would be simpler to implement and manage.

4.2 How to Mix Trucks Better

When deciding what products to put on a truck, one must account for several capacity constraints of weight, cube, and floor position. One must ensure that the products' aggregate weight, including pallets, is not higher than 45,000 pounds; and, that its aggregate volume does not exceed the volume capacity of the truck. One must also stack individual products into mixes that make good pallets. A useful tool for maximizing the ratio of volume to weight, while working with numerous variables and constraints, is linear programming.

4.2.1 Designing the Linear Program Problem

To utilize the resourcefulness of linear programming, we pick a particular customer that has enough shipments in the sample data to actually highlight the saving of this process. The chosen customer has had four shipments sent to them in the defined time period. To design the model, we first define the decision variables, constraints, and objective function below. We have considered a decreasing-weighted objective function to maximize the weight in each shipment, so when the program is run more weight is allocated to the first shipments and linearly decreasing weight to the following shipments:

Objective function: To maximize the weight of each truck in progressive sequence:

$$4(\sum_{Truck\ 1} X_i * weight) + 3(\sum_{Truck\ 2} X_i * weight) + 2(\sum_{Truck\ 3} X_i * weight) + (\sum_{Truck\ 4} X_i * weight) \quad (4.1)$$

Note: $X_i$are SKU decision variables and $X_i$is binary with 1 representing that a SKU has been selected for a shipment, or 0 representing that that SKU has not been selected for a shipment

Constraints:

Weight of each shipment: $\sum W_i * X_i \leq 45{,}000lbs$ (4.2)

Volume of each shipment: $\sum Cube_i * X_i \leq 3750$ (4.3)

Floor position of each shipment: $\sum FP_i * X_i \leq 30$ (4.4)

Each SKU is allocated to only one shipment: $\sum X_j = 1$ (4.5)

Note: j is the number of shipments (four in this example), and $Cube_i$ and FP_i are volume and floor position, respectively, for X_i

Using this model, the objective function is interchangeable. One can maximize weight, volume, or floor position as one desires. We discovered that it is best to maximize weight due to the fact that Shipper's products are, for the most part, volume heavy. Given our linear program's formulation, volume is also maximized in the process of maximizing weight. The spreadsheet model is listed in Table 3.

Table 3: Screenshot of the Model

SKU	SKU Weight	SKU Cube	SKU Floor Position	Truck 1	Truck 2	Truck 3	Truck 4
1	610.20	20.88	0.17	0	0	1	0
2	348.00	15.40	0.12	1	0	0	0
3	42.56	4.48	0.04	0	0	1	0
4	58.80	10.40	0.08	0	0	1	0
5	48.80	6.24	0.05	0	0	1	0
6	28.60	5.72	0.05	0	0	1	0
7	532.98	20.21	0.16	0	0	1	0
8	26.55	4.68	0.04	0	0	1	0
9	81.12	5.07	0.04	1	0	0	0
10	386.10	68.64	0.55	1	0	0	0

4.2.2 Findings of the Model

The results of the aforementioned model are listed below in Table 4. Shipper's original configuration consists of four shipments. The linear program, on the other hand, suggests a configuration of only three shipments. One can see that the volume parameter is also maximized. As a result of this model, the SKUs are more efficiently mixed and the fourth shipment is no longer required.

Table 4: Findings of the Model

Actual Configuration: Utilization of each parameter for the four shipments

	1	2	3	4
Weight	85%	35%	35%	99%
Volume	100%	99%	18%	81%
Floor position	100%	99%	18%	81%

Suggested Configuration: Utilization of each parameter for the four shipments

	1	2	3	4
Weight	100%	85%	69%	0%
Volume	98%	100%	100%	0%
Floor position	98%	100%	100%	0%

4.3 Mixing of SKUs

Based on our initial data analysis, by product category and customer, it is easy to conclude that most shipments are not optimally mixed. For example, many shipments are entirely dedicated to one volume-heavy product line, causing Shipper to reach "cube fill" while being very inefficient on "weight fill". To achieve better truck fill, we build a model that mixes SKUs through a comprehensive segmentation of weight and volume into different categories.

4.3.1 Categorization of data

The first step of this model is categorizing data. The pallet density, for every SKU listed in the shipment data sample, is calculated using the equation below:

$$Pallet\ Density\ of\ SKU = \frac{Pallet\ weight\ of\ SKU}{Pallet\ cube\ of\ SKU} \tag{4.3}$$

The maximum weight that can be loaded onto a truck, minus the sum of all pallets' weight, is 41,040 pounds. The maximum cube, or volume, is 3750. Thus, the optimal density is 10.944 lbs/cube. To gauge the spread of data around the optimal density, we examine the data visually as shown in Table 5 and Figure 6. As evident in Figure 6, the SKUs above the red optimal level line are weight-constrained, whereas the SKUs below the optimal level line are cube constrained. Based on this spread of data, we divide data into three different categories namely cube-constrained, neutral, and weight-constrained as shown in Table 6.

Table 5: SKU spread around the optimal density

Density bracket (in lbs/cube)	% of SKUs
<= 10.944	46%
> 10.944	54%

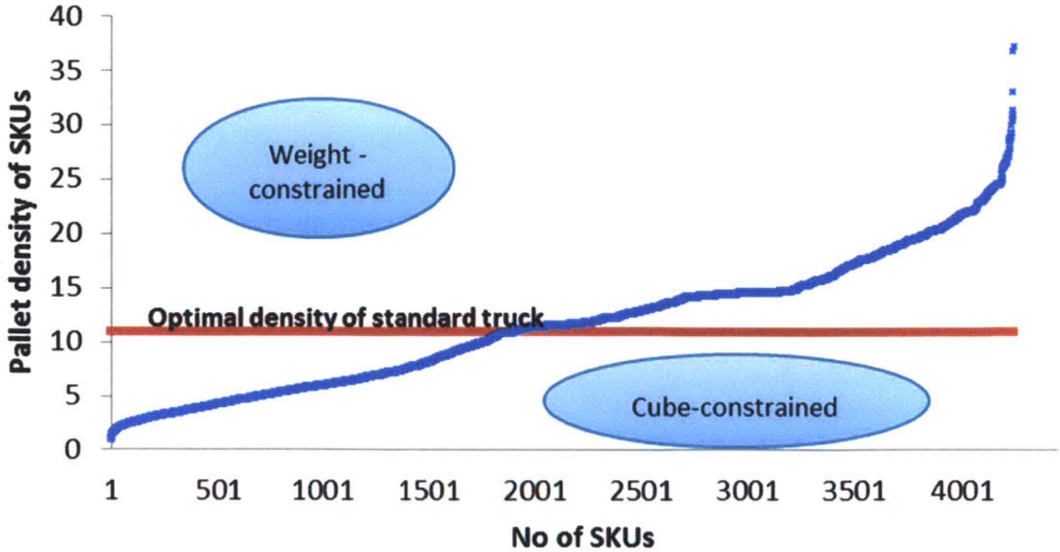

Figure 6: Pallet density spread around the optimal density

Table 6: Categorization of pallet density data

Category	Pallet density values (in lbs/Cube)	No of SKUs	% of SKUs
Cube-constrained	0.7-9.0	1649	39%
Neutral	9.1-12.5	826	19%
Weight-constrained	12.6-37.1	1770	42%
	Total	**4245**	**100%**

4.3.2 Utilization Calculation

Next, we calculate the number of pallets required for each SKU to meet the optimal cube and weight level for filling a truck. Naturally, some SKUs, when combined into homogeneous pallets, weigh out the truck before reaching 60 pallets, while others cube-out the truck. As such, we choose the minimum of the cube and weight quantities but we are always cognizant of the fact that the number of pallets is constrained by the 60 pallet maximum. We then aggregate the SKU statistics into each of our three categories, calculating the weight and cube utilization that each category would cause if it were loaded onto just one truck. These statistics are shown in Table 7.

Table 7: Utilization summary statistics for each density category

	Weight utilization			Cube utilization		
Category	Average	Maximum	Minimum	Average	Maximum	Minimum
Cube-constrained	46%	83%	6%	96%	100%	27%
Neutral	95%	100%	34%	93%	100%	40%
Weight-constrained	98%	100%	63%	65%	87%	29%

As one can see, the range for weight as well as cube utilization for each category varies greatly. To combat this variance, for the cube-constrained segment we remove SKUs that have cube utilization of less than 90%. For neutral and weight-constrained segments, we discard SKUs that have weight utilization of less than 90%. For neutral category, we chose weight utilization since most of the SKUs have higher weight rather than cube utilization. This leads to elimination of 235 outliers that skewed the distribution, reducing the sample size from 4245 to 4010 observations. The new category segmentation is shown below in Table 8.

Table 8: Revised frequency of pallet density

Category	No of SKUs	% of SKUs
Cube-constrained	1470	37%
Neutral	780	19%
Weight-constrained	1760	44%
Total	**4010**	**100%**

With this new truncated sample, we again compute the weight and cube utilization statistics as shown in Table 9. The variance in the data has been reduced significantly with the truncation and is now easier to work with. Removing the outliers also helps us to now see a pattern in the data, which is highlighted in Figure 7. This pattern shows that, for cube-constrained categories the SKUs meet the cube target. However, these same SKUs are unable to meet the weight limit by a high margin. The trend is opposite for the weight-constrained SKUs where weight is much better utilized than cube. For neutral categories, the SKUs fulfill both criteria simultaneously, making this category naturally optimal.

Table 9: Revised statistics for utilization of truckload category-wise

	Weight utilization			Cube utilization		
Category	Average	Maximum	Minimum	Average	Maximum	Minimum
Cube-constrained	48%	83%	6%	99%	100%	97%
Neutral	97%	100%	81%	94%	100%	80%
Weight-constrained	98%	100%	92%	65%	87%	29%

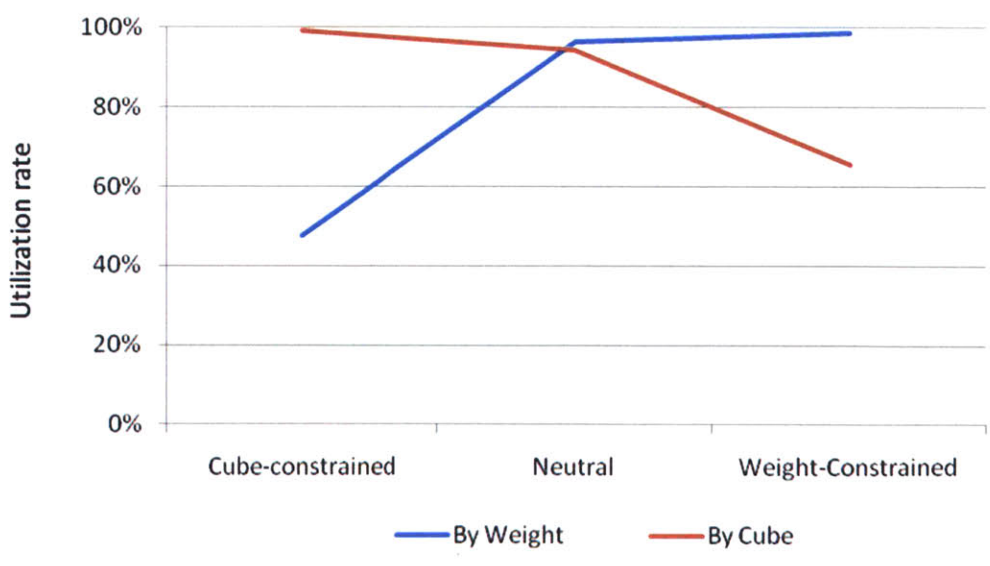

Figure 7: Average truck utilization for different types of SKUs

4.3.3 SKU Mixing Algorithm

Based on the findings discussed in section 4.3.2, we suggest a rudimentary mixing algorithm for Shipper. Obviously, the cube-constrained SKUs should be mixed with the weight-constrained SKUs in order to balance high weight utilization with low cube utilization. The neutral SKUs should be mixed with one another as they are already helping Shipper to achieve their target utilization. To illustrate the algorithm, some simple calculations follow.

First, assume that a truck is half-full (30 pallets) of SKUs of one particular category, which we shall refer to as, "primary category." The remaining half of the truck needs to be filled with a one of the other two categories, referred to as, "mixing category," which helps the primary category to achieve an optimal truckload density. As described earlier, the optimal density of a truck is 10.944. Now, the number of mixing category pallets can be less than 30 if the total combined density of both categories exceeds the optimal level. Our analysis spans three major product lines. Table 10 shows the best mix of SKUs for each line to achieve the optimal density for a truck. The best solutions have been highlighted in blue for each category combination. It is interesting to note that for weight-constrained category, the average density is so high that it forces a truck to contain less than 30 pallets. Therefore, we worked with just 20 pallets for that particular category. This category not only mixes well with cube-constrained category as expected but also with the neutral category. For the sake of simplicity, we work with discrete pallet numbers. For more encapsulating analysis, continuous numbers should be used in building pallets.

Table 10: Mixing table for different SKU categories

Primary category ordered			Mixing category			Density Utilization	Deviation from Optimal
Category	Average Density	No of Pallets	Category	Average Density	No of Pallets		
Cube-constrained	4.85	30	Weight-constrained	24.85	20	10.71	-0.24
Cube-constrained	4.85	30	Neutral	10.8	30	7.83	-3.12
Neutral	10.8	30	Neutral	10.8	30	10.80	-0.14
Neutral	10.8	30	Weight-constrained	24.85	13	10.78	-0.16
Weight-constrained	24.85	20	Cube-constrained	4.85	30	10.71	-0.24
Weight-constrained	24.85	20	Neutral	10.8	14	10.80	-0.14

Another approach to develop a SKU mixing algorithm is to look at a specific customer's ordering pattern and study their average SKU portfolio. To do this, we first arrange the demand data by the customer's locations and the density category statistics for those locations. One example is shown in Table 11 where we show the demand for two locations, Region 1 and Region 2. This data set is categorized by customer demand, in terms of number of pallets ordered for each category. For Region 1, we notice that most of the customers, A through M, order from just one category. This product is cube-constrained and will never achieve an optimal weight utilization level by itself. Thus, we recommend that its shipment be combined with a shipment for customer N that falls into weight-constrained density categories. For Region 2, the evident gains lie in mixing the weight intensive orders of customer W with the cube intensive orders of customer Z. It would thus be very beneficial to Shipper if the company treats customers W and Z as shipment partners. Meanwhile, customers X and Y, for example, have a wide portfolio of categories, which demonstrates that they could benefit greatly from simply mixing internally.

Table 11: Demand by density category (no of pallets)

Location: Region 1

Category / Customer Name	Cube-Constrained	Neutral	Weight-Constrained	Total
A	1.0			**1.0**
B	1.0			**1.0**
C	2.0			**2.0**
D	1.0			**1.0**
E	2.0			**2.0**
F	1.0			**1.0**
G	1.0			**1.0**
H	1.0			**1.0**
I	1.0			**1.0**
J	2.0			**2.0**
K	1.0			**1.0**
L	1.0			**1.0**
M	1.0			**1.0**
N	9.9	4.2	19.7	**33.7**
Total	**25.9**	**4.2**	**19.7**	**49.7**

Location: Region 2

Category / Customer Name	Cube-Constrained	Neutral	Weight-Constrained	Total
W	0.0	2.5	22.0	**24.5**
X	2.6	5.1	2.5	**10.2**
Y	73.8	65.7	32.7	**172.2**
Z	165.0	0.0	0.0	**165.0**
Total	**282.2**	**81.5**	**96.5**	**460.3**

Our proposed algorithm is based on the assumption that Shipper's data sample is representative of its annual demand. We do realize that the actual pattern may be far deviant from this assumption. However, Shipper can implement this same methodology by simply analyzing annual demand, as we have analyzed sample demand, in order to identify shipment partners based on customer location and ordering pattern. The important criteria would be to determine how many days of data would have to be accumulated before the opportunity for matching is explored by Shipper. For customers with wide ordering portfolio, Shipper can try to incentivize them into ordering the complimentary SKUs through ordering pattern-based discounts. This approach would help Shipper to obtain better truck utilization while simultaneously fulfilling the same demand with a given service level.

Of course, our algorithm is constrained due to the skewness of the demand data towards cube-constrained SKUs. The sample data reports that 68% of the orders fall in this category, as shown in Table 12. This highly skewed sample limits the implementation of the algorithm and restricts achievement of maximum efficiency. Again, the sample might not be a true representation of annual demand; and, analyzing the annual demand could give a different picture. We also recommend that Shipper further explore the density categorization based on the number of pallets ordered, rather than on the number of SKUs.

Table 12: Demand by pallet category-wise

Category	No of pallets	% of pallets
Cube-constrained	57045.0	68%
Neutral	8533.8	10%
Weight-Constrained	17845.2	21%
Total	**83424.0**	**100%**

5. CONCLUSION

Based on our analysis, we arrive at three recommendations. First, Shipper should re-examine whether to use floor position as a criterion for declaring when to ship a truck. As shown in our analysis, there is evidence that the calculation of cube fill already encompasses floor positions for the majority of Shipper's shipments. Eliminating the floor position metric as an operating heuristic makes the implementation of order guidelines easier.

Second, Shipper should implement linear programming to better mix their orders at a tactical level. As the number of decision variables becomes larger, it becomes impossible to use linear programming at a large strategic level. As such, we propose that Shipper use linear programming only for mid-size customers. In order to implement this, tactical personnel at the operations level would have to be trained in the usage of this tool.

Shipper can explore the opportunities presented by mixing SKUs that belong to different density categories. This would help Shipper to obtain much better average truckload utilization rates. There are a few SKUs that are outliers, and would have to be treated on a case-by-case-basis, but they are part of a minority. For a more robust order-mixing model, Shipper should explore the ordering pattern of their high volume customers. Shipper can then use those customers as a jumping off point for extending cross-mixing opportunities to other customers in the same geographic vicinity. Overall, these steps should help Shipper to achieve significantly better truckload utilization rates, while also reducing their carbon footprint.

Finally, the applicability of these methods extends beyond Shipper to other companies and other modes of freight transport. The usage of linear programming at an operations level can help achieve a greater mix of trucks; but, its applicability is restricted by the scale and scope of operations due to a ceiling on the number of decision variables. Strategic mixing, through a segmentation algorithm, should be further explored for a firm's most important customers, as it can help one to achieve much better utilization. It also provides greater value proposition to a company as it increases channel integration with its partners.

BIBLIOGRAPHY

Caputo, A. C., Fratocchi, L., & Pelagagge, P. M. (2006). A genetic approach for freight transportation planning. *Industrial Management & Data Systems, 106*(5), 719-738.

Christopher, M. (1992). *Logistics & Supply Chain Management*, Pitman, London.

Collins,Terry R. & Rossetti,Manuel D. (2004). Tracking logistics performance in the trucking industry: Institute of Industrial Engineers, Houston, TX, United states,1053-1054.

Don Taylor, G. & Meinert,Timothy S. (2000). Improving the quality of operations in truckload trucking: IIE Transactions (Institute of Industrial Engineers), 32, 6, 551-562, Kluwer Academic Publishers, Dordrecht, Netherlands.

Fawaz,Youssef M. & Garrison,William L. (1994). Truck and highway combinations for increasing trucking productivity in market niches: National Research Council, Washington, DC, United States, 1430, 10-18.

Glickman, T. S., & White, S. C. (2008). Optimal vendor selection in a multiproduct supply chain with truckload discounts. *Transportation Research: Part E, 44*(5), 684-695.

He,Daifeng; Nickerson,J.A. (2006). Why do firms make and buy? Efficiency, appropriability, and competition in the trucking industry: Strategic Organization, Sage Publications, UK, 4, 1, 43-69.

Ichoua,Soumia (2008). A collaborative-distributed framework for the long-haul truckload trucking industry: IEEE Technology Management Council; IEEE Portugal Section, Inst. of Elec. and Elec. Eng. Computer Society, Estori, Portugal.

MacCarthy, B.L., & Kritchanchai, D. (1999). Responsiveness of the order fulfillment process. *International Journal of Operations & Production Management, 19*(8), 812-833.

Olhager, Jan (2010). The role of the customer order decoupling point in production and supply chain management. *Computers in Industry, 61*, 863-868.

Preiss, K.I., Nagel, R.N., & Goldman, S.L. (1995). *Agile Competitors and Virtual Organizations: Strategies for Enriching the Customer*, Van Nostrad Reinhold, New York, NY.

Richardson,Barbara C. (2001). Freight trucking in a sustainable transportation system - Framework for analysis : National Research Council, 1763, 57-64.

Srethapakdi, S., & An, Y. (2006). Order promising/fulfillment and customer/channel collaboration in supply chain management. Massachusetts Institute of Technology.

Stank, T., Germain, R., & Davis-Sramek, B. (2010) The impact of order fulfillment service on retailer merchandising decisions in the consumer durables industry," *Journal of Business Logistics*

Tan, K.C., Liu, D.S., Huang, S.Y., Ho, W.K., & Goh, C.K. (2007). On solving multi-objective bin packing problems using evolutionary particle swarm optimization. *European Journal of Operational Research, 190*, 357-382.

Taylor, G. D., & Whicker, G. L. (2002). Optimization and heuristic methods supporting distributed manufacturing. *Production Planning & Control, 13*(6), 517-528.

Turner, T.J., Mendibil, K., Daisley, P., Breen, T.H.J., & Bititci, U.S. (2002) Improving the reliability of the customer order fulfillment process in a product identification company. *International Journal of Production Economics. 78*(1), 99-107.

Webster, S., & Weng, Kevin Z., (2007) Ordering and pricing policies in a manufacturing and distribution supply chain for fashion products. *International Journal of Production Economics. Volume 114, Issue 2,* 476-486.

Xie, M., Xiao, T., Shi, J., Jin, J., & Chen, G. (2009). Ordering, wholesale pricing and lead-time decision in a three-stage supply chain under demand uncertainty. *Computers & Industrial Engineering, 59*, 840-852.

Yan, R., & Wang, J. (2010). Service level, pricing strategy and firm performance in a manufacturer-giant retailer supply chain. *Journal of Product & Brand Management, 19*(1), 61-66.

Yildiz, H., Ravi, R., & Fairey, W. (2010). Integrated optimization of customer and supplier logistics at robert bosch LLC. *European Journal of Operational Research, 207*(1), 456-464.

Zenk, Edward (2007). Composites in the trucking industry: Society of Plastics Engineers, Cincinnati, OH, United States, 3, 1377-1382.

Zhang, L., Ma, Q., & Jiao, R. (2010). Accountability-cased order fulfillment process reengineering towards supply chain management. A case study at a semiconductor equipment manufacturer. *Journal of Manufacturing Technology Management, 21*(2), 288-305.