Real-Time Order Acceptance in Transportation Under Uncertainty

By

Hiral Nisar
B.S. Computer Engineering, University of Mumbai, 2005

And

Joshua Rosenzweig
B.S. Mechanical Engineering, University of Connecticut, 2011

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 2014

Signature of Author .... Signature redacted ....
Master of Engineering in Logistics Program, Engineering Systems Division
May 9, 2014

Signature of Author .... Signature redacted ....
Master of Engineering in Logistics Program, Engineering Systems Division
May 9, 2014

Certified by .... Signature redacted ....
Dr. Chris Caplice
Executive Director, Center for Transportation and Logistics
Thesis Supervisor

Accepted by .... Signature redacted ....
Prof. Yossi Sheffi
Director, MIT Center for Transportation and Logistic
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

Real-Time Order Acceptance in Transportation Under Uncertainty

By

Hiral Nisar and Joshua Rosenzweig

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

Without using any order acceptance criteria, retail companies distributing products with private transportation fleets are not able to maximize their profits because they are not adequately utilizing their capacity. The objective of this paper was to create and validate a model to determine if historical demand data can be used by retail firms operating private fleets to make effective real-time order acceptance/rejection decisions with the purpose of eliminating unprofitable orders in a short-haul transportation setting. A Java tool was generated to instantaneously decide whether or not to accept an order depending on the order location and time of receipt. The model was tested against optimal decisions using total demand knowledge and several alternative real-time decision-making strategies. The model was found to significantly outperform the alternative real-time decision making strategies and provided profits approximately eight percent lower than the optimal decisions. We conclude that using historical demand probabilities is useful in informing the decisions of retail firms seeking to utilize private fleets efficiently and increase profitability through cost reduction.

Thesis Supervisor: Christopher Caplice
Thesis Advisor: Francisco Jauffred

ACKNOWLEDGEMENTS

We would like to express our gratitude toward the people who gave us valuable guidance on our thesis. First, we want to thank Dr. Chris Caplice for his engagement and all of the long hours he spent sharing his expertise in the field of transportation. As we were not partnered with a thesis sponsor, Dr. Chris Caplice acted as a sponsor and chief project supervisor, providing valuable insights to thesis scope and analysis throughout the year. We would also like to give a special acknowledgement to Dr. Francisco Jauffred, a research affiliate at MIT. His vast knowledge and understanding of analytics was instrumental to our success.

Hiral Nisar and Joshua Rosenzweig

I am grateful to my husband, my parents and my brother for turning my dream into reality. This program would not have been possible without their unconditional love and support. Thank you for being there for me always. I am also thankful to Bruce, Jen, Allison, and all my friends without whom this past year wouldn't have been such a fun learning experience.

Hiral Nisar

I would like to take this opportunity to thank my mother, Laney, for her unconditional love and support throughout my developing career. I'd also like to give special thanks to my brother, Zach, for his encouragement. Finally, although my father was not able to see my admittance into MIT, I am confident that he would be proud of my accomplishments and I am grateful for all of his contributions toward my upbringing.

Joshua Rosenzweig

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Package delivery refers to the service of delivering small packages to end consumers. Delivery can be provided by postal services, parcel service companies, less-than-truckload third party transportation companies, and private company fleets. Each distribution service type is suitable for unique situations. Large parcel delivery companies, like UPS, can afford to accept all orders to achieve high customer service because they have resources and physical networks, which allow them to utilize economies of scale to decrease marginal cost of delivery. Alternatively, firms with private company fleets have the option to decide which orders to personally deliver. Although package delivery modes include ground, air, and water, for the purposes of this paper we focus on ground transportation.

1.1 Motivation for Research

The online retail industry has been rapidly growing over the past decade with a 16% increase in revenues from 2012 to 2013 (US Census Bureau, 2013). The choice to switch from third party transportation distribution to private company fleets is becoming increasingly important as increasing global competition continues to apply pressure on retail companies to provide faster delivery and higher service levels. In addition to lowering operating costs, retail companies can increase control over delivery by utilizing privately owned transportation vehicles. Increased transportation control gives these companies the power to maintain high service levels, which is critical in sustaining positive customer relationships. One retail company that is moving toward private fleet distribution is Amazon.com. In the Christmas holiday season of 2013, Amazon.com suffered widespread distribution failure when its key parcel delivery partner, UPS, failed to deliver large numbers of packages on time, which ultimately led to a large decline in stock prices. Although Amazon has stated that their distribution centers were prepared for the seasonal demand spikes, it is speculated that the cause of the failure was largely linked to Amazon's promise of two day delivery times to Amazon prime subscribers who pay additional fees for premium service. The agreement with these customers contributed to a large number of last minute purchases, which were not properly forecasted. UPS was therefore unable to meet this unexpected surge in demand due to lack of preparation and capacity constraints. Amazon's reputation was damaged greatly from the distribution failure, and executives sought to avoid any similar occurrences in the future. This contributed to Amazon's decision to implement a private fleet distribution to 30-40 US markets in 2014 (Banker, 2014).

In transportation, cost is minimized through shipment aggregation. Firms may have to reject certain incoming orders due to limited vehicle capacity. Accepting orders indiscriminately can put unwanted constraints on the ability to combine shipments. Despite this, minimal research has been done on order acceptance criteria involving probabilities of order deliveries in sub-regions based on historical demand. Because of this, many companies simply accept incoming loads to fill their capacities, foregoing many profitable future orders due to lack of available capacity. In this paper, we develop order acceptance criteria based on the probability distribution of orders across service regions, marginal cost of accepting the order, and truck capacity.

1.2 Research Question

Without using any order acceptance criteria, companies aren't able to maximize their profits because they are not adequately utilizing their capacity. Our research objective is to create and validate a model to determine if demand probabilities can be used by retail firms operating private fleets to inform order acceptance/rejection decisions with the purpose of eliminating unprofitable orders in a short-haul transportation setting.

1.3 Structure

This paper is organized as follows: Chapter 2 provides insights into decision making in other industries. Chapter 3 describes the methodology behind the model creation process. Chapter 4 outlines the testing and design phase by introducing various alternative strategies for comparison purposes. Chapter 5 shows the model output and provides detailed analyses of our model performance versus the performance of the alternative strategies. Chapter 6 summarizes the conclusions drawn from our model validation and suggests future research opportunities.

2. LITERATURE REVIEW

In many industries, a lack of standards in operational decision making processes is detrimental to the profitability of a firm. Sequential order acceptance in certain industries leads to orders being accepted that are not feasible due to capacity constraints. Moreover, an order may be feasible, yet unprofitable due to allocation of resources at the plant level. Although decision making processes have been evolving over the last several decades, opportunities for improvement are plentiful. Models for decision making typically involve quantitative methods such as linear programming, simulation, and regression analysis. Acceptance criteria are also frequently used by firms to prevent acceptance of orders that are infeasible or unprofitable. Several industries that benefit from making intelligent order acceptance decisions are the manufacturing industry and the trucking industry.

2.1 – Order Acceptance in Make-To-Order Manufacturing

Make-To-Order (MTO) manufacturing is an example of an industry where firms must make frequent decisions on how to handle incoming orders. Effective operational decision making is critical in overall firm success because production capacity provides a unique constraint, as firms must choose not only which orders to accept, but also order priority. Order acceptance decisions can incorporate a variety of factors including available capacity, technological ability, order revenue, and order due date.

Complementing the concept of lean manufacturing, the MTO industry has been using order acceptance criteria to make acceptance/rejection decisions to increase operational efficiency. Hemmati, Ebadian, and Nahvi (2012) discuss order acceptance criteria for accepting orders in an

MTO environment involving factors such as feasibility, production schedules, and delivery end dates. Accounting for these factors allows MTO companies to select the incoming orders that will maximize profit and facilitate a smooth production planning process. They propose a process for decision making that includes classification and prioritization of orders. By incorporating criteria into the decision to accept and assign priority to orders, the firm is able to optimize daily operations and avoid overcommitting capacity.

Mathematical simulation is also a valuable tool in determining order acceptance procedures in manufacturing environments. Nandi and Rogers (2004) created a simulation model that utilizes the "full amount of information on the shop floor status." The detailed information allows the model to predict current and future behavior by utilizing a systems approach. Nandi and Rogers, however, recognize that the model is limited by certain assumptions, including deterministic processing times and the inability to account for future orders during a simulation.

While many MTO companies use factors at the plant level to inform their decision making processes, other research suggests that taking customer value into consideration is the key in creating and maintaining maximum long-term profits through customer relationships. Hao, Yu, Wu, and Chen (2013) use linear programming optimization techniques to determine optimal order acceptance in an MTO environment. The research proposes order acceptance policies that focus not only on short-term gains from orders, but also on long term effects of relationship building tactics. These variables are represented by weights and customer-value based revenue parameters. Although Hao, Yu, Wu, and Chen found success incorporating the concept of relationships into their model, this is uncommon in mathematical models due to the qualitative nature of relationship building.

2.2 – Order Acceptance in the Trucking Industry

In the US trucking industry, third party logistics companies for less-than-truckload (LTL) shipments have difficulty with effective operational decision making due to high variability in pricing and high uncertainty in demand. In addition to high uncertainty, aggressive competition leads to low margins. This makes profit maximization crucial to the success of individual firms and leads to innovation, such as collaboration between carriers to reduce cost (Hernandez & Peeta, 2014).

Marginal cost and marginal revenue have been used as decision making tools to maximize profit through effective order acceptance. In a long distance trucking decision model, Powell and Sheffi (1988) effectively introduce marginal cost and marginal revenue of adding and removing truck capacity from locations upon receiving an order. In long haul trucking, profitability of an order is determined not only by the direct revenue and cost associated with the order, but also by the probability that the truck will be able to find another order upon delivery completion. For that reason, an order to a large, robust city is more attractive to a trucker because opportunities for outbound loads will be greater relative to areas with low population. To capture this idea quantitatively, Powell and Sheffi use a concept called Total System Contribution (TSC) to factor opportunity cost into each decision, where TSC is equal to order profit plus the marginal contribution of adding a truck to the destination minus the marginal contribution of subtracting a truck from the origin region. Marginal contribution is determined by assigning quantitative values, or end effects, to various cities. By including these indirect contributions into a decision, Powell and Sheffi are able to take profitability of future orders into account. This type of analysis, involving assigning values to possible outcomes, represents a trend of systems thinking that is evolving to take advantage of all factors that affect profitability of a firm.

2.3 – Complementary Research on Real-Time Trucking Delivery Acceptance

The problem investigated in this paper, deciding to accept or reject orders in package delivery, has previously been investigated in academia. Patrick Jailletz and Xin Lu (2012) provide an instantaneous decision making strategy for an optimal routing problem, known as the traveling salesmen problem, to minimize cost under uncertain conditions which are commonly used by taxi, trucking, and courier services. In the typical traveling salesman problem, orders are aggregated at the end of an acceptance period and routing decisions are made at one time. Jailletz and Lu consider a situation where orders are received online and an acceptance/rejection decision is made immediately upon order receipt. In Jailletz and Lu's model, the trucker has already determined a route and must make a decision to incorporate the new order into the current delivery route. This decision is made with an objective to minimize the total route distance and penalties for not accepting an order. Although Jailletz and Lu's research gives insight into the complexity of real-time decision making in a truck delivery setting, their model does not include historical customer demand as a consideration. Jailletz and Lu state that including a probabilistic model to forecast and optimize design routes is sometimes "not possible nor desirable," however, do not explain their rational further.

2.4 – Commonality of Research in Operational Decision-Making Processes

The research presented above shares one major commonality: it does not incorporate probabilistic demand when determining optimal order acceptance. In transportation, profitability of an order is determined by the manner in which orders can be aggregated in terms of capacity. The reason for this is that by taking advantage of order proximity, companies minimize cost of delivery given capacity constraints. Because of this, we do not seek to optimize route distance, but rather to incorporate demand probability into the decision making process of our model to maximize the effects of economies of scale on deliveries, minimizing marginal cost of each order.

3. METHODS

The purpose of this thesis is to provide an alternative to the current processes that govern the operational decision making of retail companies with their own private fleet. We develop a probabilistic model that uses order information to output a decision on order acceptance in real-time. The model utilizes historical demand data to determine the likelihood of profitability of delivering to specific locations. The goal of the model is to give a tool to package delivery companies to allow them to differentiate between profitable and unprofitable orders prior to order acceptance.

3.1 – Obtaining and Preparing Data

The data to create and validate the model was acquired from a small retail company in the Greater Boston area of Massachusetts. The company operated its own transportation fleet and delivered orders within a 20 mile radius of its distribution center. The data contained only information on customer location; however, exact addresses were unavailable for use. Hence, customers were segregated into 19 separately defined regions of varying customer distribution as seen in Figure 1 below.

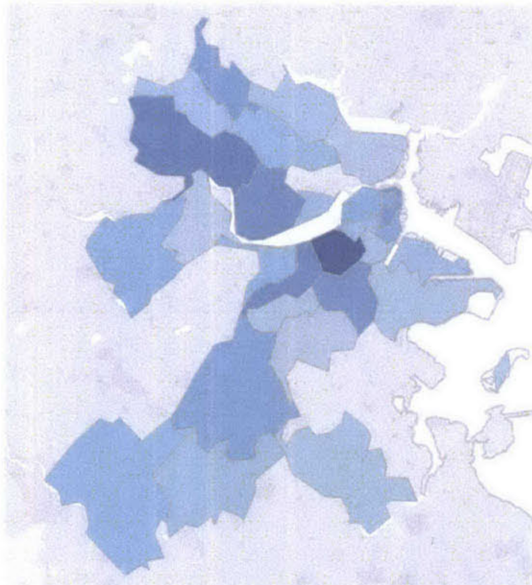

Figure 1: Data representation of the 19 regions of the Greater Boston area. Customer density is represented by strength of shading, with darker regions signaling higher customer density.

The company data acquired did not include samples of daily demand. For the purposes of model formulation and testing, representative demand samples were generated. To simulate daily demand samples that exceed our model capacity, an assumption is made that a unique five percent of customers submit an order every business day. This allowed for 100% of the customers to submit an order in a 20 day period. To determine the orders that occur on each random daily sample of demand, a random number between 1 and 20 was generated for each customer. The entire simulated data set therefore consisted of 20 simulated days of demand with orders assigned with the same number being received on the same day. Each order is also attributed with an order size, which is introduced as number of totes.

The model relies on the time an order is received; therefore time stamps for each order needed to be simulated. Equation 1, a Poisson process, was used to create stochastic time between orders (refer to Appendix A for a data sample).

$$\Delta T = -\frac{\ln(r)}{\lambda} \tag{1}$$

where,

ΔT: the time between orders

r: random number

λ: expected order frequency

Expected order frequency is found by dividing the number of orders in a day by the length of the order acceptance period. Using this method, a cumulative time stamp was created for each simulated demand day. For the purposes of this model, the order acceptance period is defined as the period of time that orders may be received, with an assumption that they are delivered immediately afterwards. The order acceptance period can be adjusted to account for specific business practices.

3.2 – Determining Probabilistic Model Criteria

To formulate the probabilistic model, we first determined the criteria that have large impacts on transportation cost. The criteria that appear to most directly affect profitability of an order are linked to that order's destination region. Profitability of an order requires that marginal revenue exceed marginal cost for a region; therefore a certain amount of order revenues must be received to offset the cost of delivering to a specific region. The breakeven number of orders for a region, where revenues equal costs, can be calculated using a transportation cost approximation formula.

The two major criteria for the model are the breakeven number of orders and the probability of receiving that many orders during the time left in the order acceptance period. Available truck capacity is also an important consideration when developing the model.

3.3 – Probabilistic Model Structure

After determining criteria, a decision-making tree (Figure 2) was developed to demonstrate the flow of each order through the model. The initial question determines whether or not orders have already been accepted into this region. In the event that orders are already accepted in the region, the fixed cost has already been incurred; therefore the order will always be accepted to decrease marginal cost of delivery to that region until truck capacity is reached.

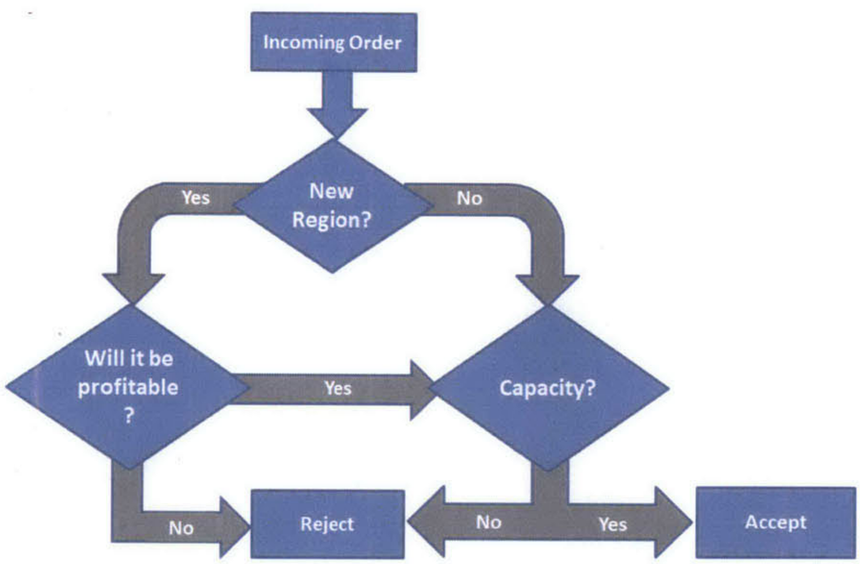

Figure 2: Decision making tree for the probabilistic model

The "will it be profitable?" node in the decision making tree above contains the most critical computation. At this node, the model will determine the profitability of adding capacity to a new region. In the event that a region has no prior orders, the model then determines the number of totes expected to be received in the remainder of the order acceptance period. If the model expects the breakeven point to be reached with the time left and capacity is available, it will accept the order and subsequent orders in the same region until capacity is fully utilized. For example, if an order is received for delivery to the Brookline region and there are no prior orders there, the model will determine how many orders must be accepted to reach a breakeven point. If the model finds that Brookline requires X orders to be profitable, and there are T hours left in the order period, it will determine how many orders are expected to be received in those T hours. If Brookline is expected to receive at least X orders in T hours, it will accept the order under the assumption that the breakeven number of orders will be met or exceeded. In the event that Brookline is expected to receive fewer than X orders in T hours, it will reject the order and all future Brookline orders for the rest of that order acceptance period.

3.4 – Primary Probabilistic Model Calculations

There are several key calculations at the core of the probabilistic model. First, the breakeven number of totes must be calculated to determine the number of orders required in a region to reach profitability. The breakeven number of totes requires a calculation of cost. Total cost comprises purchase cost, transportation cost, and a fixed operating cost. Purchase cost includes all costs incorporated with buying, handling, and storing inventory. In our case, we made an assumption that purchase cost was equal to 85% of tote revenue, which allows for 15% of tote revenue to be allocated toward transportation cost and profit margin. Fixed cost is the daily cost of operating the truck, including the cost of labor. Equation 2, below, describes the company's generalized cost structure:

$$Total\ Cost = 0.85 * Tote\ Revenue + Transportation\ Cost + Fixed\ Cost \tag{2}$$

The nature of the model requires that orders are constantly being added to regions, which will ultimately have an effect on the actual route driven at the end of the acceptance period. Because we could not include route optimization in our model due to data constraints, exact transportation cost was impossible to determine. For this reason, we modified an equation derived by Carlos Daganzo (2010) to allow for an approximation of transportation cost (TC) to be calculated upon addition of each order. Equation 3, below, uses number of customers and customer density of a region to approximate costs associated with load totes, travel to and from a region, and unload totes.

$$TC = C_s\left[E[n] + n_{tour} + \frac{1}{2}\right] + C_d\left(2\left[n_{tour} + \frac{1}{2}\right] d_{LineHaul} + \frac{E[n]k_{TSP}}{\sqrt{\delta}}\right) + c_{vs}E[D] \quad (3)$$

where,

$E[n]$: expected number of stops in a region

$E[D]$: expected demand in a region

$d_{linehaul}$: driving distance from the distribution center

n_{tour}: number of tours in the region

C_s: cost per stop

C_d: cost per distance

C_{vs}: cost per unit per stop

δ: number of stops per unit area

k_{TSP}: dimensionless network factor

Assumptions were made to determine values of constants used in the equation, which can be found in Appendix B. Transportation cost could now be updated upon receipt of each new order.

By incorporating total cost of a certain number of orders with total revenue, a calculation could be made to determine the number of orders that would create profitability in a region.

Figure 3, below, graphically demonstrates the breakeven point in Brookline, where the total number of orders required for Brookline to be profitable occurs at the intersection point between revenues and costs.

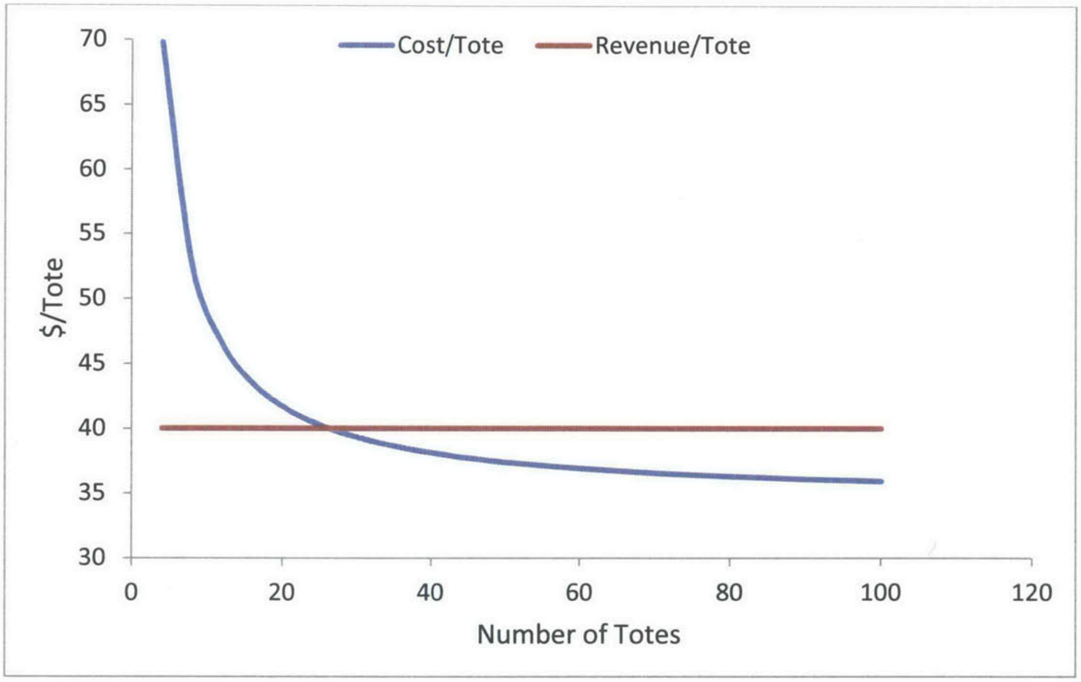

Figure 3: Illustration of Brookline breakeven graph. The point at which revenue and cost functions cross determines the breakeven number of totes. According to the graph, Brookline will require 28 totes in an order period to allow the region to become profitable.

After the breakeven number of orders is known for each region, the likelihood of receiving that number of orders must be determined. For the purposes of this model, an assumption was made on the frequency of deliveries per customer and kept constant. This assumption is crucial in allowing the model to determine the number of totes a region can expect to receive in a certain time period. By knowing the total number of customers in a region and assuming that each customer will order periodically, a Poisson distribution analysis will allow for the computation of the expected number of totes in a certain time interval. The Poisson distribution is calculated using Equation 4 as follows:

$$f(k;\lambda) = \frac{\lambda t^k e^{-\lambda t}}{k!} \tag{4}$$

where,

k: number of orders

λ: expected number of orders per unit time

t: time remaining in the order acceptance period

The Poisson output, f(k;λ), tells us the probability of receiving exactly k orders in a time period given an expected value, λ. In this situation, however, the model requires the probability of receiving at least k orders. The model does not only seek to meet the breakeven number of orders, but is also satisfied if the number of orders is exceeded.

The graph in Figure 4 depicts the likelihood of receiving totes within a time period in the Brookline region. Because probabilities are incorporated in calculations, the number of totes can never be found deterministically. To account for this, various lines are plotted to demonstrate different levels of confidence.

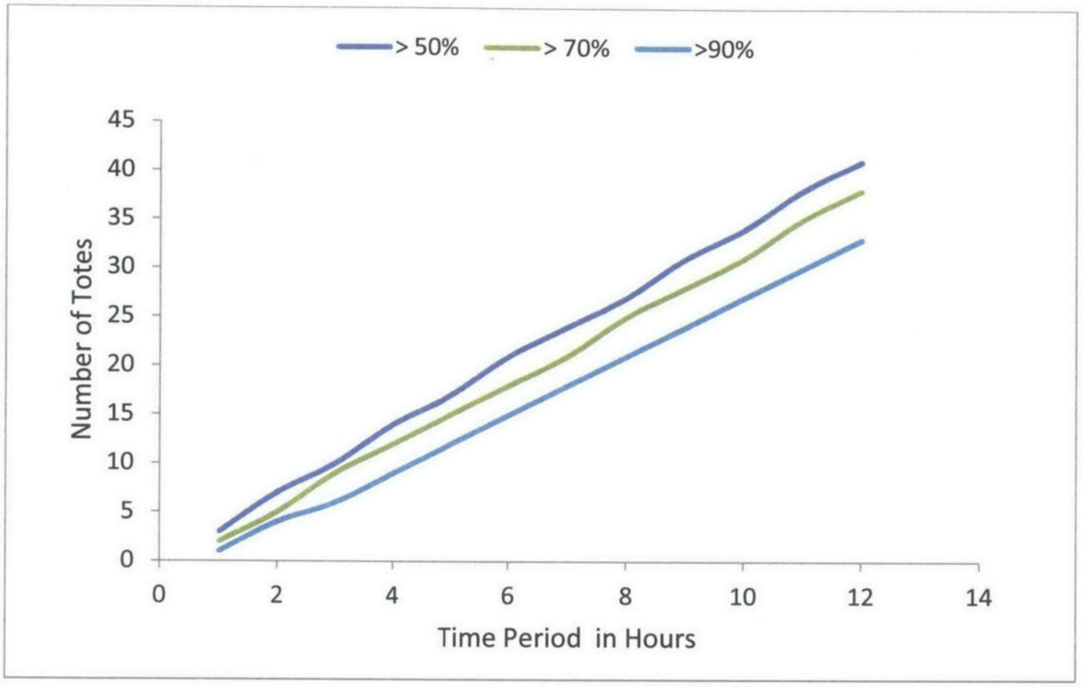

Figure 4: Illustration of expected orders over time in Brookline.

Combining the Poisson distribution analysis with the breakeven analysis, the model is able to determine the likelihood of receiving enough orders to reach the breakeven point for each region. Orders for regions that are not expected to reach the breakeven number of totes in the time left will be rejected.

3.5 –Assumptions, Justifications, and Limitations

The probabilistic model relies on assumptions that can be modified as needed. Several model iterations were tested to create the final model and their assumptions are outlined below.

3.5.1 –Fixed Order Size Probabilistic Model Assumptions

The assumptions and their justifications for the initial model are as follows:

- The order acceptance time period is 12 hours in length and is followed immediately by delivery of the orders accepted within that time frame. The model assumes next day delivery service. This follows the current trend in online delivery growth and decreasing lead times.
- The model accepts an order for a new region with 70% certainty that it will receive enough totes to reach a breakeven number of totes.
- Customers order once every 20 days, or about once a month. This assumption is based loosely on the idea that customers are ordering common household items that must be replenished frequently.
- Each order is four totes and each tote generates the same revenue of $40. This assumption greatly simplifies the model and can be modified by incorporating statistical analysis of order size and revenue.
- Each truck has a fixed capacity of 80 totes. For simplicity purposes, each truck serves regions individually to isolate costs associated with each region.
- The number of trucks that the carrier should have to maximize its profits is included as a parameter and determined through sensitivity analysis. Hence the model is tested for different fleet sizes and the profits are compared to find the ideal fleet size.
- Entire orders are shipped together, which implies that orders are equivalent to shipments.
- The profit calculation considers the operational cost of the trucks. This fixed cost is set at $100 per day per truck. The model also accounts for 60% of the fixed cost as a penalty for not utilizing a truck.
- The cost for rejecting an order is not considered in this model. The only loss from rejecting an order is the loss of revenue from that order.

3.5.2 –Variable Order Size Probabilistic Model Assumptions

The first iteration of the model gives some basic insights that are discussed further in Section 5.1. However, the model is greatly simplified due to the assumptions mentioned above. In the second iteration of the model, some of these basic assumptions were removed to match real-world scenarios. The following assumptions were relaxed for this enhanced model.

- Orders now vary by order size. Number of totes is applied in the simulated data as a uniform distribution between 1 and 7.
- Because orders now have a variable number of totes, the revenue from each order also varies as per the number of totes in the order.

3.6 – Model Validation

We used Java as a platform to create and validate the probabilistic model. The tool allowed for flexibility to easily alter parameters while maintaining underlying functionality. The tool allowed us to validate the effectiveness of our model through profit calculations by using 20 days of simulated data as inputs. The tool generated multiple echelons of outputs at the order, truck, and daily levels. The following pseudo algorithm describes the logic built in the aforementioned Java tool:

Notation:

currTruckList: List of trucks assigned to the different regions
truckCityList: List of trucks for a particular destination region
MaxTruckCapacity: Maximum truck capacity of each truck and is set to 80 totes
MaxNoOfTrucks: Maximum number of trucks
timeInAcceptancePeriod = time the order came in.

```
Algorithm:
For Every Order {
	destinationRegion = Order.getDestination();
	truckCityList = get a list of trucks going to the destinationRegion;
	For each Truck in truckCityList {
		truckCapacityAvailable = MaxTruckCapacity – Truck.getCapacityUsed();
		if (Order.getOrderSize() <= truckCapacityAvailable){
			Accept the Order and add the Order to the Truck;
			orderAccepted = true;
			Break out of this loop;
		}
	}
	If (orderAccepted is not true) {
		startNewTruck = true;
	}

	If (startNewTruck is true) {
		If (currTruckList.Size() < MaxNoOfTrucks){
			TimeInAcceptancePeriod = Order.getTime();
			feasibleToStartNewTruck =
			feasibleToStartTruck(TimeInAcceptancePeriod);
			if (feasibleToStartNewTruck is true){
				Start a new truck;
				Add the truck to currTruckList;
				Accept the order and add the order to the new truck;
			} else {
				Reject the order;
			}
		} else {
			Reject the order;
		}
	}
}


feasibleToStartTruck(TimeInAcceptancePeriod){
	RemainingTime = AcceptancePeriod – TimeInAcceptancePeriod;
	Probability = Calculate P[Receiving breakeven number of orders in RemainingTime]
	If (Probability >= 0.7){
		Return true;
	} else {
		Return false;
	}
}
```

4. ALTERNATIVE STRATEGIES

Determining the effectiveness of the probabilistic models required a comparison of results of various alternative strategies. The strategies chosen range from overly simple to complex. The Java tool described in Chapter 3 incorporates each strategy for comparison in the Results section. Strategy descriptions are described below.

4.1 – Optimal Model Overview

In our problem, decisions must be made instantaneously upon order arrival. In an optimal situation, we would have complete knowledge of all daily orders before making decisions. This model optimizes capacity utilization and maximizes profit. By comparing our model against results from the optimal situation, we can determine the effectiveness of the decision making process in our model. The optimal model, described below, is only optimal under the assumption of constant revenues per tote. With variable tote revenues, a more complex algorithm would be formulated. The pseudo algorithm for finding the optimal results is as follows:

Notation:
regionOrderMap: It contains (key,value) pairs, where key is the region and value is list of orders
orderList: List of orders for a particular region
cityTruckList: List of trucks for a particular region
allTruckList: List of all the trucks across all the regions
currTruckList: List of trucks assigned to the different regions
MaxTruckCapacity: Maximum truck capacity of each truck and is set to 80 totes
MaxNoOfTrucks: Maximum number of trucks

```
Algorithm:
regionOrderMap = Segregate the orders as per the destination regions;
For each City in regionOrderMap{
        orderList = get the list of orders for the City;
        Sort the orderList based on the revenue of the totes;

        For every Order in the sorted orderList{
                orderAddedToTruck = false;
                For every Truck in cityTruckList{
                        truckCapacityAvailable = MaxTruckCapacity – Truck.getCapacityUsed();
                        if (Order.getOrderSize() <= truckCapacityAvailable){
                                Accept the Order and add the Order to the Truck;
                                orderAddedToTruck = true;
                                Break out of this loop;
                        }
                }
                If (orderAddedToTruck is not true){
                        Start a new truck;
                        Accept the order and add the order to the new truck;
                        Add the truck to cityTruckList;
                }
        }
        Add all cityTruckList to allTruckList;
}

Sort allTruckList based on the profits for the truck;
Set currTruckList as the top (MaxNoOfTrucks) trucks from the allTruckList;
Accept all the orders in the trucks in currTruckList and reject the remaining;
```

4.2 – Myopic Manager Model Overview

In this model, we create a scenario where orders are accepted sequentially until capacity is filled. As in our probabilistic model, decisions are again made instantaneously upon receipt of orders. This model reflects the actions of a firm with no intelligent operational decision making process and simulates a worst-case scenario. The pseudo algorithm used to find the results for this strategy is as follows:

Notation:
currTruckList: List of trucks assigned to the different regions
truckCityList: List of trucks for a particular destination region
MaxTruckCapacity: Maximum truck capacity of each truck and is set to 80 totes
MaxNoOfTrucks: Maximum number of trucks

Algorithm:
```
For Every Order {
        destinationRegion = Order.getDestination();
        truckCityList = get a list of trucks going to the destinationRegion;
        For each Truck in truckCityList {
                truckCapacityAvailable = MaxTruckCapacity – Truck.getCapacityUsed();
                if (Order.getOrderSize() <= truckCapacityAvailable){
                        Accept the Order and add the Order to the Truck;
                        orderAccepted = true;
                        Break out of this loop;
                }
        }
        If (orderAccepted is not true) {
                startNewTruck = true;
        }

        If (startNewTruck is true) {
                If (currTruckList.Size() < MaxNoOfTrucks){
                        Start a new truck;
                        Accept the order and add the order to the new truck;
                        Add the truck to currTruckList;
                } else {
                        Reject the order;
                }
        }
}
```

4.3 – Binary Logistic Regression Model Overview

The binary logistic regression uses explanatory variables and previous decisions to predict the discrete outcomes of future decisions in order to make instantaneous order acceptance decisions. To determine the previous decisions to use for the regression formulation, we used binary outputs from the optimal solution model on a dedicated dataset containing 20 days of simulated demand. In our model, explanatory variables used to create the model are the incoming order's distance (miles) from the distribution center and order size (number of totes). Time that the order was received was not included in the regression analysis because the optimal model explained in Section 4.1 does not incorporate order time when making decisions; therefore any correlations would be purely coincidental. The parameters from the regression are used to create a function with outputs between 0 and 1. As orders are received, the function output is compared to an acceptance threshold, above which orders are accepted. The functions are determined by the following equation (Hosmer & Lemeshow, 2000):

$$E(Y/X_1, X_2) = \frac{1}{1 + e^{-B_0 - B_1 X_2 - B_2 X_2}} \tag{5}$$

where,

Y: optimal decision binary

X_1: distance of the order from the distribution center

X_2: the number of totes of the order

B_0: constant regression coefficient

B_1: distance regression coefficient

B_2: order size regression coefficient

The pseudo algorithm for the logistic regression strategy is as follows:

Notation:
currTruckList: List of trucks assigned to the different regions
truckCityList: List of trucks for a particular destination region
MaxTruckCapacity: Maximum truck capacity of each truck and is set to 80 totes
MaxNoOfTrucks: Maximum number of trucks
thresholdValue: The threshold value for the logistic regression equation and is set at 0.53

Algorithm:

```
For Every Order {
        destinationRegion = Order.getDestination();
        distance = get driving distance from the DC to destination region;
        orderSize = Order.getOrderSize();
        regressionValue = get regression value using the logistic regression equation for distance
        and order size.
        If (regressionValue > thresholdValue){
                orderToBeAccepted = true;
        }
        If (orderToBeAccepted is true){
                truckCityList = get a list of trucks going to the destinationRegion;
                For each Truck in truckCityList {
                        truckCapacityAvailable = MaxTruckCapacity – Truck.getCapacityUsed();
                        if (Order.getOrderSize() <= truckCapacityAvailable){
                                Accept the Order and add the Order to the Truck;
                                orderAccepted = true;
                                Break out of this loop;
                        }
                }
                If (orderAccepted is not true) {
                        startNewTruck = true;
                }

                If (startNewTruck is true) {
                        If (currTruckList.Size() < MaxNoOfTrucks){
                                Start a new truck;
                                Accept the order and add the order to the new truck;
                                Add the truck to currTruckList;
                        } else {
                                Reject the order;
                        }
                }
        }else {
                Reject the order;
        }
}
```

4.3.1 Binary Logistic Regression Model Derivation

Running the regression analysis, we obtain parameter values shown below in Table 1.

Table 1: Binary logistic regression parameters

Regression Coefficients	Explanatory Variable Units	Coefficient	Standard Error	p-Value
Constant (B_0)	-	-1.578	0.09895	< 0.0001
Line haul (B_1)	Miles	-0.08093	0.01143	< 0.0001
Order size (B_2)	Totes	0.5559	0.01860	< 0.0001

From Table 1, we see that p-Values for regression parameters are less than 0.0001. Using an alpha value of 0.05, we conclude that the regression parameters are statistically significant, which allows us to reject the null hypothesis that they do not explain the data sufficiently.

According to the regression function described in Section 4.3, positive coefficients dictate that as an input variable increases, the output increases. From our regression analysis, we find a line haul coefficient (B_1) of -0.08093 and an order size coefficient (B_2) of 0.5559. This tells us that large values for line haul will decrease the model output, increasing the likelihood of order rejection, while large values for order size will increase the model output, increasing the likelihood of order acceptance. Large sized orders that are in close proximity to the distribution center are therefore desirable. Table 2, below, describes the relative strength of the explanatory variables in computing logistic regression decisions.

Table 2: Heat map table of logistic regression output as a function of order size and line haul

Order size (totes)	Line haul (miles) 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
7	91%	90%	90%	89%	88%	87%	86%	85%	84%	83%	82%	81%	79%	78%	76%	75%
6	85%	84%	83%	82%	81%	79%	78%	77%	75%	74%	72%	70%	69%	67%	65%	63%
5	77%	75%	74%	72%	71%	69%	67%	65%	64%	62%	60%	58%	56%	54%	52%	50%
4	66%	64%	62%	60%	58%	56%	54%	52%	50%	48%	46%	44%	42%	40%	38%	36%
3	52%	50%	48%	46%	44%	42%	40%	38%	36%	35%	33%	31%	29%	28%	26%	25%
2	39%	37%	35%	33%	31%	30%	28%	26%	25%	23%	22%	20%	19%	18%	17%	16%
1	26%	25%	23%	22%	21%	19%	18%	17%	16%	15%	14%	13%	12%	11%	10%	10%

Table 2 demonstrates the comparative strength of the two explanatory variables used to create the model. The table shows that variability in order size has a much greater impact than line haul in the likeliness to accept an order. As line haul increases, marginal decreases in the likelihood of order acceptance are small relative. This tells us that the benefits from larger order sizes outweigh the disadvantages of larger line haul distances.

To determine an appropriate threshold for accepting and rejecting orders based on the function output, sensitivity analysis was done on the same dataset from which regression was created to minimize the number of incorrect values. For the entire dataset, the total number of incorrect values was calculated with Equation 6 as follows:

$$Min \sum abs(Y_{optimal} - Y_{regression}) \tag{6}$$

where,

$Y_{optimal}$: binary decision output from the optimal model

$Y_{regression}$: binary decision output from the regression model

By varying the threshold of acceptance, we find a threshold of 0.53 allows for the minimum number of incorrect answers. With the threshold set, we can now calculate a pseudo-R^2 value to determine goodness of fit of the regression curve. The pseudo-R^2 value is determined by Equation 7, shown below (Freese & Long, 2005).

$$Pseudo - R^2 = \frac{\#\ Correct}{Total\ \#\ of\ Decisions} \tag{7}$$

Given a threshold of 0.53, we find 3615 correct decisions out of a total of 4777 orders; therefore the regression model has a pseudo-R^2 value of 0.757. Additionally, the accuracy of the regression model as compared with the optimal model is shown in Figures 5 and 6 with goodness of fit curves. Explanatory variables were investigated separately to isolate the effectiveness of their contributions to the model.

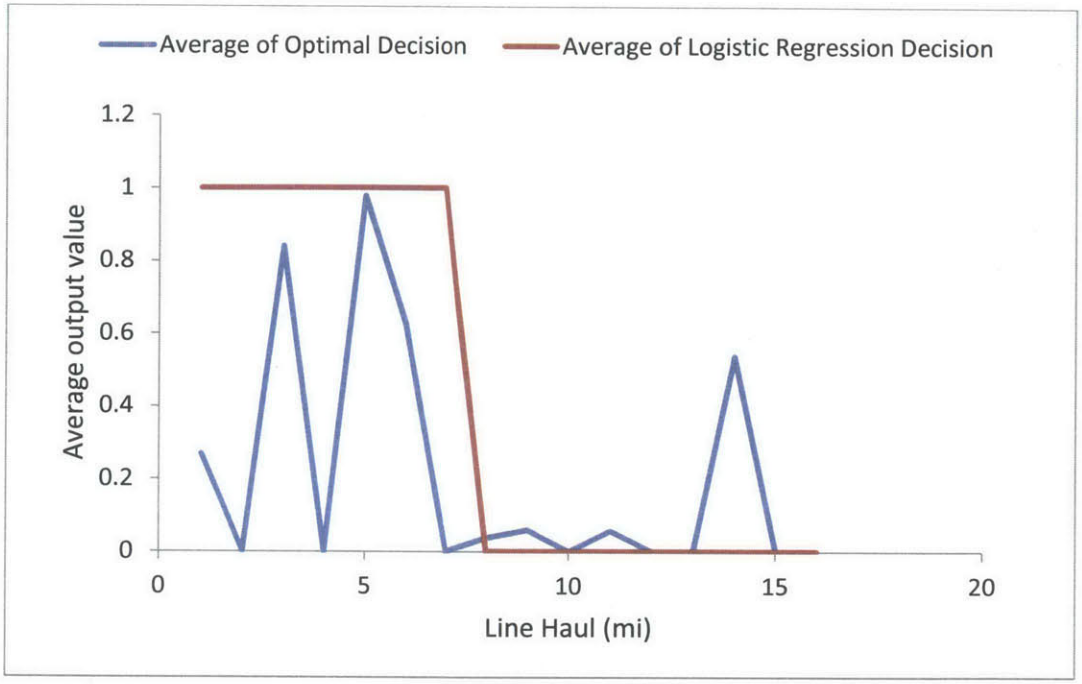

Figure 5: Binary logistic regression goodness of fit to the optimal model output. This graph compares the average regression and optimal outputs with a constant order size to isolate the line haul variable for evaluation purposes. The average order size of four totes was chosen to avoid bias from extreme values.

Figure 5 demonstrates the goodness of fit of the binary logistic regression as compared with the optimal model decision, specifically pertaining to the line haul explanatory variable. From the graph, we can see that while the optimal decision is highly variable across line haul distances, the logistics regression model strictly accepts or rejects orders based on a certain distance thresshold. The variations in the optimal model are due to information that is not captured by the explanatory variables used to create the logistic regrssion model. Overall, we see a basic trend for both models to accept orders in close proximity to the distribution center and reject orders with large line hauls.

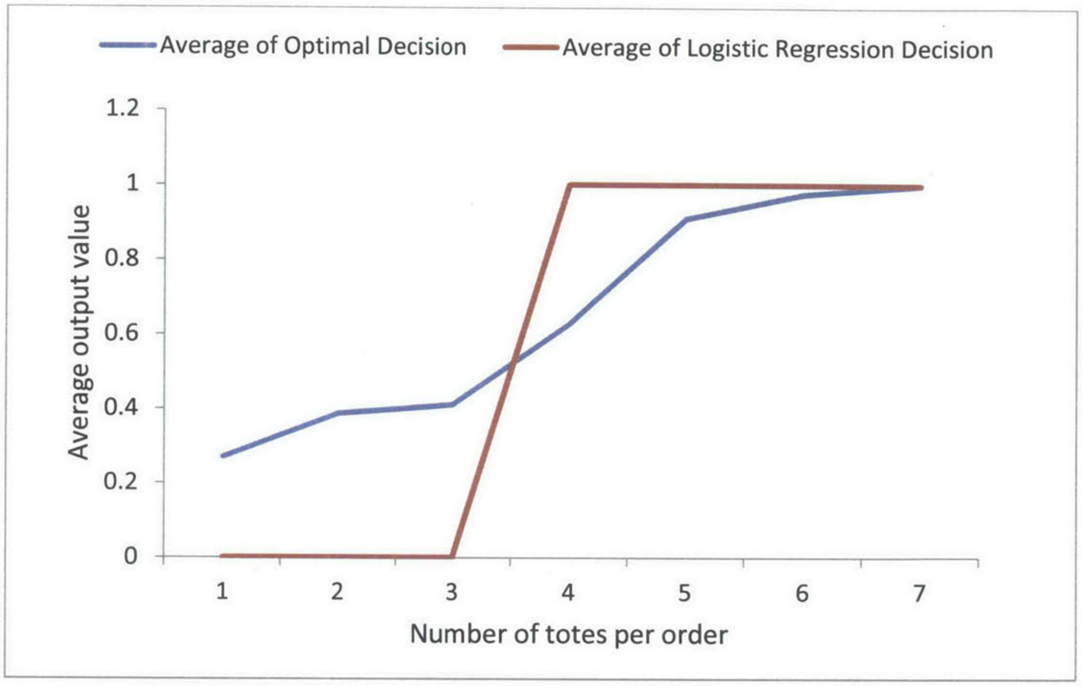

Figure 6: Binary logistic regression goodness of fit to the optimal model output. This graph compares the average regression and optimal outputs with a constant line haul from the distribution center to isolate the order size variable for evaluation purposes. A line haul approximately 6 miles was chosen to avoid bias from extreme values.

Figure 6 demonstrates the goodness of fit of the binary logistic regression as compared with the optimal model decision, specifically pertaining to the order size explanatory variable. Again we see that while the optimal decision is variable across order size, the logistics regression model strictly accepts or rejects orders based on a certain order size thresshold. For example, when a truck only has space for 2 totes, the optimal strategy will accept an order size of 2 totes. The binary logistics model, however, cannot take capacity into account, and therefore will never accept small orders Overall, we see a basic trend for both models to accept large orders and reject small orders.

5. RESULTS & ANALYSIS

The purpose of the probabilistic model is to allow companies with private fleets to allocate resources effectively and eliminate unprofitable orders. In order to determine the effectiveness of our proposed models, we tested them against the various possible strategies described in Chapter 4. We tested each strategy using the same randomly generated demand data and capacity constraints to determine the expected daily profits. The results were obtained by inputting 20 days of simulated demand into each model. Profits are reported as daily averages. Each model is also tested for different fleet sizes to provide insights on the ideal fleet size the company should have to generate the greatest profits. The results are dependent on values used for model parameters, which are listed in Appendix B.

This section displays the results obtained for the two model iterations mentioned in Section 3.5 and draws key insights from the results obtained from the various strategies. The first iteration, in which order size was kept constant, illustrated the key insight that our probabilistic model could help the firms to increase their profits by not accepting unprofitable orders. However, the model needed to be tested under real-world conditions to understand its potential. Therefore, the assumption of constant order size was removed. The variable order size probabilistic model confirmed our hypothesis that our order acceptance model based on the past history will benefit the company monetarily by accepting only profitable orders. The detailed insights for each of the models can be found in the below sub-sections.

5.1– Fixed Order Size Results and Analysis

Figure 7, below, compares the results for the first iteration of our model with the alternative strategies under the assumptions stated in Section 3.5.1. In this scenario, order size is kept constant. The graph depicts the average daily profit for each strategy across fleet sizes.

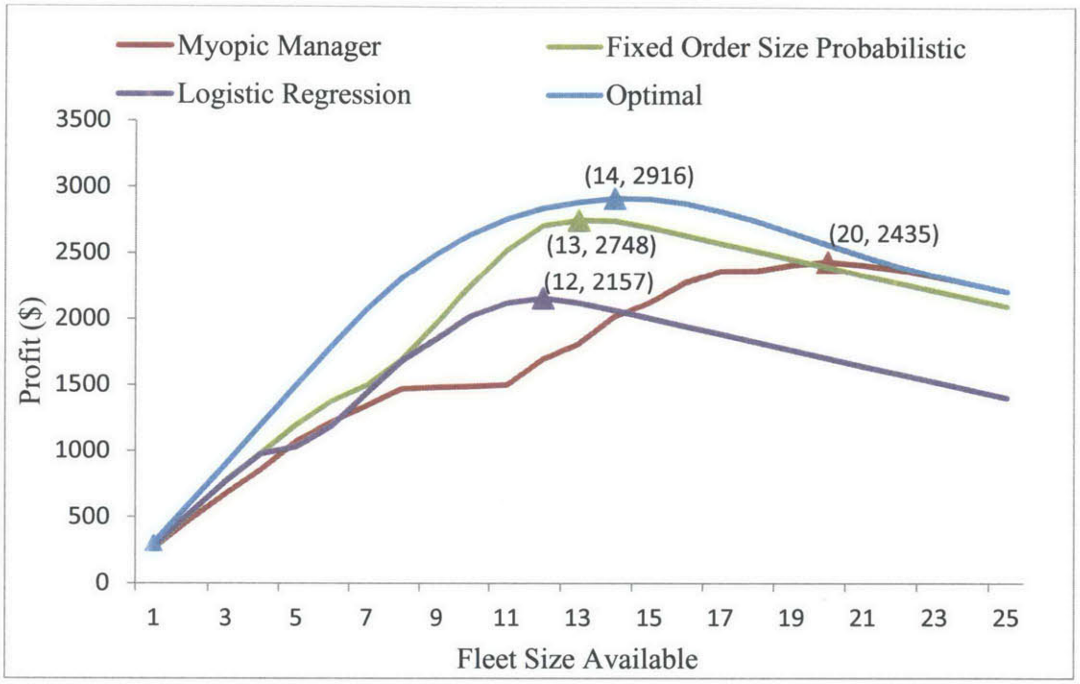

Figure 7: Average daily profit by strategy. The figure shows the profit earned by using different strategies with orders containing a constant number of totes. We can see from the graph that each strategy has a different optimal fleet size.

We can see from Figure 7 that the optimal model would allow the company to make a maximum profit of $2,916 when the fleet size is 14. The optimal model makes maximum profit as it makes a decision based on the complete demand data and only accepts order which would maximize its capacity. Among the three instantaneous decision models, the probabilistic model performs better than the myopic manager model and the logistic regression model. Detailed analyses of these results are described below. Figure 8, below, depicts the number of trucks the models actually use as compared to the fleet size and provides insight to the behavior of the models depicted in Figure 7.

Figure 8: Number of trucks used by fleet size for all strategies in a fixed order size scenario. Myopic manager and optimal models fully overlap. Each model linearly utilizes trucks until a certain fleet size, when additional trucks are no longer useful.

As can be seen from Figure 8, the number of trucks used by the myopic manager and optimal models increase linearly with the fleet size until they reach a capacity of approximately 22 trucks. Until 22 trucks, every truck added gives the model increased capacity to service new regions or deploy multiple trucks in densely populated regions. However, after 22 trucks, both the models do not fully utilize all the available trucks as they have enough capacity available on their currently deployed trucks to meet the daily demand. The probabilistic model also grows linearly with the increase in fleet size, but tapers when it reaches 13 trucks. This is because the model will not deploy new trucks after a certain point in the acceptance region as it will not be able to make any profit from the new trucks. The logistic regression curve also levels off after using approximately 13 trucks as it will accept only those regions that are closer to the distribution center and will eventually reach a point where it has enough capacity on its truck to service these closer regions.

The purpose of the first fixed order size probabilistic model was to verify whether our order acceptance model, based on likelihood of receiving breakeven number of orders in the acceptance period, allows the company with no decision making tool to increase its profitability. For the initial model comparison, order size was kept constant and the model was tested for different fleet sizes. The initial results are encouraging and indicate that using past data to determine the probability of receiving orders helps to improve the bottom line figures.

As can be seen from Table 3, below, the optimal strategy results are much higher than those of other strategies as it has complete knowledge of the daily demand and accepts only those orders that maximize truck utilization and daily profit. Because our probabilistic model is forced to make instantaneous informed decision under uncertainty, the profits earned are about 6% less than the profits earned using the optimal strategy. Our model, however, assists the firms in improving response time by making an instantaneous decision rather than having to wait for the entire order acceptance period before making the optimal decision.

Table 3: Comparison of maximum daily profit for all strategies with fixed order size

Model Type	**Ideal Fleet Size**	**Maximum Average Daily Profit ($)**	**Difference From Optimal ($)**	**Difference from Optimal (%)**
Optimal	14	2916	-	-
Myopic manager	20	2435	(481)	(16.5)
Fixed order size probabilistic	13	2748	(168)	(5.8)
Logistic regression	12	2157	(759)	(26.0)

In the fixed order size scenario, the probabilistic model allows the firm to make an additional $313 per day over the myopic manager model; an increase of approximately 12%. The difference in profits occurs because the probabilistic model rejects orders for certain regions that would cause the operating costs to be greater than the revenue, whereas the myopic manager's model would accept all orders, including those to regions which will not reach a breakeven number of totes.

The logistic regression model is the poorest performing real-time decision making model in the fixed order size scenario, earning approximately 26% less than the optimal profits. The results show that the explanatory variables used for the logistic regression model are not able to accurately capture the effects of demand behavior and capacity constraints.

The probabilistic model makes the greatest profits with fewer trucks than the optimal model, because of its inherent nature. The model accepts orders until it expects profits will not be reached. Due to the probabilistic nature, in some cases actual orders in a region will be lower than expected orders. When this occurs, profits are decreased. Alternatively, when demand is greater than a single truck capacity it will fill the truck and then begin a new truck, time permitting. When the last truck is deployed to a region, however, the model will not capture the profits from additional orders after the capacity is filled. This phenomenon decreases the daily profit when number of orders is less than expected, but does not increase the daily profit when number of orders is greater than expected, thus underestimating the ideal fleet size under optimal conditions.

The results obtained from the fixed order size iteration of our model provide important insights of how a company can use past history to predict the probability of receiving enough orders in a region. However, these results should be substantiated further to assess their real-life capability. The next iteration seeks to account for this by incorporating variable order sizes into the demand dataset. The results obtained from the next iteration of the model are discussed in the following section.

5.2 – Variable Order Size Results and Analysis

The second iteration of the model, which includes real-world conditions of variable order size, provides deeper insights into the strength of our model. In this iteration, because the orders vary in size, firms can increase their potential revenues by accepting larger orders as transportation cost is decreased when multiple totes are delivered to the same location. Each model was tested with the same dataset, which included a uniform distribution of order size ranging from 1 tote to 7 totes. We also evaluated the optimal fleet size that a company should have to maximize its profit through sensitivity analysis. Detailed results of all strategies are reported below in Figures 9 and 10.

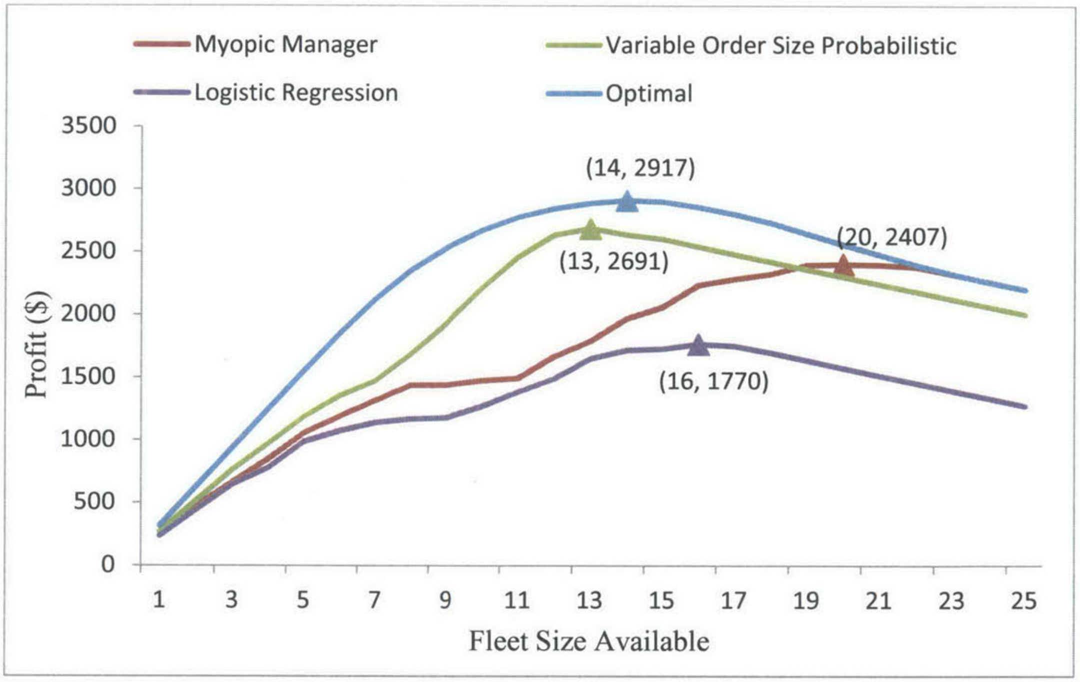

Figure 9: Average daily profit by strategy. The figure shows the profit earned by using different strategies with orders containing a variable number of totes. We can see from the graph that each strategy has a different optimal fleet size.

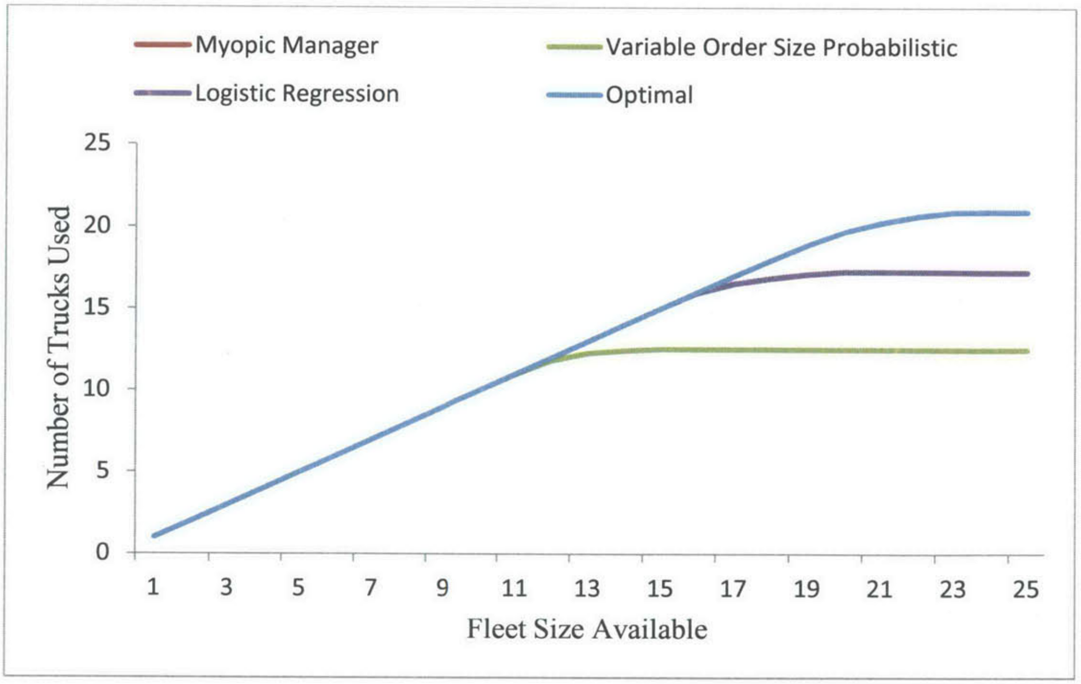

Figure 10: Number of trucks used by fleet size for all strategies with variable order size. As with the fixed scenario, myopic manager and optimal models fully overlap. Again, each model linearly utilizes trucks until a certain fleet size, when additional trucks are no longer useful.

The following sections provide comparison analyses for each strategy in the variable order size iteration.

5.2.1 – Optimal Model Analysis

The combined results in Figure 9 show that the maximum profit earned for the optimal strategy is much higher than that of any other strategy as it has complete information of the demand and only chooses those orders that earn higher revenue and allow for utilizing maximum truck capacity. The model maximizes profit when the fleet size is 14 trucks. When the fleet size is greater than 14 trucks, the profits start decreasing as the cost of operating the new trucks is higher than the revenues earned from accepting the orders. The profits earned decrease as the trucks are underutilized and not able to reach the breakeven number of orders required to meet the operational cost of running the truck. Figure 10 shows that the optimal model curve grows linearly with the fleet size until a fleet size of 21. The explanation for the linear growth is that with more trucks available, the model has increased ability to service new regions or service the same region multiple times. When the fleet size becomes greater than 21 trucks, the existing trucks have sufficient capacity to service the regions and fulfill the demand.

5.2.2 – Myopic Manager Model Analysis

The results in Figure 9 indicate that for the myopic manager model the profit increases as the fleet size increases. The reason for this behavior is that more trucks enable the firm to service more regions and increase its potential to accept more incoming orders. However, after 21 trucks there is capacity available in each of the trucks to meet the daily demand for a majority of regions, hence new trucks will not be utilized and will incur a penalty. This accounts for the behavior of the myopic manager curve in Figure 10, explaining why the curve levels off at 21 trucks used. The myopic manager model behaves like an optimal model at the point where the profits converge because all the daily demand can be accepted as there is effectively no capacity constraint.

The profits for this strategy are lower compared to the optimal model and the probabilistic model when the fleet size is less than 20. This is because the model deploys trucks in the regions for earlier incoming orders. As a result of this, it can accept an order for a region which does not have a high customer density and will cause most of the truck's capacity to be underutilized. As a result, the company will lose money in these unprofitable regions and report lower profits. As there are up to 19 regions to service, the model will make maximum profit if it can meet the demand sufficiently in each region. Some regions require more than one truck and some regions may not have enough demand to justify a single truck. Through sensitivity analysis, we found the ideal fleet size for this model should be 20 trucks to allow the model to secure enough capacity in high volume regions. Having a fleet size of more than 20 trucks results in some trucks not being deployed and incurring a penalty.

5.2.3 – Variable Order Size Probabilistic Model Analysis

Figure 9 provides useful insights about the validity of the probabilistic model. Average daily profit is maximized when the fleet size is 13 trucks. This indicates that the company's ideal fleet size should be 13 trucks. Having more than 13 trucks does not help to increase profits or the probability of accepting new orders as the model is highly influenced by the order acceptance period. The model will not deploy a new truck after a certain point in the acceptance period as the likelihood of meeting the breakeven number of orders falls below 70%. Hence, the profits start decreasing after they reach their peak as the model incurs the penalty for not utilizing the available trucks. This behavior of not utilizing all the available trucks when the fleet size is greater than 13 is also reflected in Figure 10 above.

5.2.4 – Binary Logistic Regression Model Analysis

Compared with all other strategies, the logistic regression analysis offers the poorest performance, as shown in Figure 9. For capacities greater than 16 trucks, profits decrease due to the penalty for not utilizing trucks. Figure 10 shows that the number of trucks used grows linearly with fleet size until the available trucks reach a count of 17. For 18 or more trucks, the curve flattens as the existing trucks can hold all of the orders that the binary logistic regression model will accept. One factor causing low profits is that the model did not adequately describe the situation with the available explanatory variables. This contributed to the fairly low pseudo-

R^2 value of 75.7% correct decisions described in Section 4.3.1. The regression model actually performs worse than the myopic manager model because the nature of the regression parameters allows for some regions to be ruled out entirely due to high line haul distances that may otherwise be profitable due to high demand. Similarly, orders with only one tote will most likely not be accepted even if there is already available capacity to the same region.

Another driver for this phenomena is that the regression model does not change acceptance threshold based on capacity; therefore it cannot adjust its decision making process throughout the day. In our probabilistic model, the decision to accept a new region is highly dependent on the time an order is received; however in the logistic regression model, the time an order received is irrelevant, which creates a poor decision making process.

5.2.5 – Combined Analysis

Figure 9 combines the individual model results to demonstrate visually how the strategies compare with each other. From the graph we can see that our probabilistic model outperforms all other instantaneous decision making models.

Table 4, below, shows that the maximum profit earned by applying the myopic manager strategy is 17.5% less than the optimal results and the probabilistic model performs better than the myopic manager model. The profit earned by applying our strategy is within 8% of the optimal profits and also requires one less truck.

Table 4: Comparison of maximum daily profit for all strategies with variable order size

Model Type	Ideal Fleet Size	Maximum Average Daily Profit ($)	Difference from Optimal ($)	Difference from Optimal (%)
Optimal	14	2917	-	-
Myopic Manager	20	2407	(510)	(17.5)
Variable Order Size Probabilistic	13	2691	(226)	(7.5)
Logistic Regression	16	1770	(1147)	(39.3)

The results indicate that the logistic regression model is not an ideal choice for an order acceptance model as it offers a significantly lower profit than all other strategies. Comparing the instantaneous decision making models, the myopic manager model provides 10.5% less profit than our probabilistic model and the logistic regression model provides 34.2% less profit. We can conclude from Table 4 that incorporating demand history, as in the probabilistic model, is useful for increasing profitability through informed decision making processes.

Table 5, below, compares the profits gained by each model when the fleet size is fixed at 13 trucks, which is the ideal fleet size of the probabilistic model. This allows us to compare each strategy against the best case scenario for the probabilistic model developed in this thesis. This comparison is useful from a realistic viewpoint, as companies will not have perfect flexible control over fleet size, and must make decisions that will allow for the best utilization of current available resources.

Table 5: Comparison of average daily profit earned with fleet size of 13

Model Type	Average Daily Profit	Difference from Optimal	Percent Difference from Optimal
Optimal	2894	-	-
Myopic Manager	1792	(1102)	(38.1)
Variable Order Size Probabilistic	2691	(203)	(7.0)
Logistic Regression	1652	(1242)	(42.9)

From Table 5 we can see that in a scenario with fixed capacity of 13 trucks, the probabilistic model provides only 7% less profit than the optimal model, while the myopic manager and logistic regression models provide 38.1% and 42.9% less, respectively. These results further prove that a real-time order acceptance decision model based on likelihood of receiving orders in a region performs nearly as well as the optimal model.

6. CONCLUSION

To determine the effect of historical demand on the effectiveness of real-time operational decision making, a probabilistic model was created to maximize profit through rejection of unprofitable orders for retail firms with its own operating fleet. As orders are received for service regions with no available capacity, the model determines if expected daily revenues justify the cost of deploying an additional truck to those regions. The expected revenues are determined by probabilistic analysis of customer data. When trucks are deployed to a region, all orders from that region are accepted until truck capacity is full.

To provide a baseline for profit analysis, an optimal model was created that was able to use all demand data simultaneously. To further test our probability model, two alternate real-time order acceptance strategies were run in parallel: the myopic manager model and a binary logistic regression model. Java tools were created for all four strategies to facilitate model validation. By inputting the same simulated demand data sets into each Java tool, we were able to determine that our probabilistic model provided approximately 8% less profit than the optimal solution; however, it outperformed all other instantaneous decision making models. In the variable order size scenario, average daily profits for the myopic manager model and the logistic regression model were about 11% and 34% lower than the probabilistic model, respectively. We determine that the myopic manager and logistic regression models simply do not offer the insights required to match supply with demand in an effective way.

We are able to conclude that demand probabilities are an important variable for use in real-time operational decision making processes. This research shows that demand probabilities through historical demand patterns should be considered by companies with private fleets who have to make decisions to accept or reject orders under capacity constraints.

6.1 – Probabilistic Model Limitations

For practicality, several assumptions were made to create the probabilistic model to account for the lack of availability of detailed demand data for a retail company operating its own private transportation fleet. The underlying assumptions that remove strength from the model are as follows:

- Each customer orders once every 20 days. This assumption was crucial in allowing straightforward statistical analysis; however, it does not accurately depict the distribution of demand that a real firm will face under uncertainty.
- Each order provides constant revenue per tote. By assuming constant revenue pet tote for the probabilistic model, we were able to simplify the profit equations and focus the effects of the model on cost minimization.
- Each truck can only deliver to one region. This assumption was used to allow for simplicity in cost calculations in order to separate breakeven costs by region. In reality, regions that are close together provide further opportunity for shipment aggregation.
- There is no cost for rejecting an order. In reality, a retail company will not reject the customer's order, but instead opt to hire a third party logistics transportation company to make the delivery. To model this accurately, these outsourcing costs must be modeled and then incorporated into the total cost equation described in Chapter 3. Additionally, by passing delivery to third party providers, the company relinquishes control and service levels may be affected.

- Transportation cost parameters were estimated based on experience and common sense, rather than determined through data analysis. Therefore, actual operational cost may not be accurately illustrated by the transportation cost estimation equation described in Chapter 3, which may affect calculations for breakeven number of totes.

6.2 – Suggestions for Future Research

Using probability to create a decision making process for order acceptance has been shown to increase profitability under the assumptions dictated in this paper. To understand how the model would function in a more realistic setting, however, more research must be done to amend the model to remove naïve assumptions.

First, acquiring real package delivery demand data would allow for the model to be tested more accurately. Assumptions surrounding customer order frequency, order size, and order revenue would be refined to include realistic probability distributions to be included in the creation of demand samples.

Second, the model should be adjusted to allow trucks to deliver to more than one region. This would allow for the possibility of additional aggregation of orders, which would cause the model to be more favorable to orders from low customer density regions if they are in close proximity to high density regions.

Third, the cost for order rejection must be included in the model. Using parcel delivery cost data from third party transportation companies; we can determine accurate cost of outsourcing tote delivery and incorporate this cost into the regional breakeven analysis.

Last, parameters for transportation cost must be more accurately determined. A thorough analysis of transportation data will provide more insights into the true values and allow for more accurate transportation cost estimations.

In addition to relaxing assumptions, future research should be done to create accurate models of current practices of retail companies operating private fleets. In our thesis, we simplify the actions of these companies with the myopic manager model, a worst-case scenario in which companies will simply accept every order until capacity is unavailable. This is most likely not true for many companies, as they may adapt strategies and develop decision making heuristics over time. By investigating the real processes used by companies, we can more accurately assess the benefits of the model. Finally, the model should be tested in different locations, other than the Greater Boston area, to conclusively determine its effectiveness across various geographic areas.

REFERENCES

Banker, S. (2014, February 24). Amazon Takes One Step Forward, Two Steps Back. . Retrieved May 7, 2014, from http://logisticsviewpoints.com/2014/02/24/amazon-takes-one-step-forward-two-steps-back/

Daganzo, C. (2010). One-to-Many Distribution. *Logistics systems analysis* (4th ed.,). Berlin: Springer.

Freese, J., & Long, S. (2005). Estimation, Testing, Fit, and Interpretation. *Regression Models for Categorical Dependent Variables Using Stata* (2nd ed.,). : Taylor & Francis.

Hao, J., Yu, J., Wu, M., & Chen, X. Customer-value-based order acceptance policy in make-to-order manufacturing. *International Conference on Computational and Information Sciences*, 426-429. Retrieved April 3, 2014, from the Inspec database.

Hemmati, S., Ebadian, M., & Nahvi, A. A new decision making structure for managing arrival orders in MTO environments. *Expert Systems with Applications*, *39*, 2669-2676. Retrieved November 10, 2013, from the Inspec database.

Hernandez, S., & Peeta, S. A carrier collaboration problem for less-than-truckload carriers: characteristics and carrier collaboration model. *Transportmetrica A: Transport Science*, *10*, 327-349. Retrieved March 19, 2014, from the Compendex database.

Hosmer, D. W., & Lemeshow, S. (2000). Logistic Regression. *Applied logistic regression*(2nd ed.,). New York: Wiley.

Jailletz, P., & Lu, X. Online Traveling Salesman Problems with Rejection Options. *Networks*, *58*, 137-146. Retrieved November 12, 2013, from the Inspec database.

Nandi, A., & Rogers, P. Using Simulation to Make Order Acceptance/Rejection Decisions. *Simulation*, *80*, 131-142. Retrieved April 3, 2014, from the Inspec database.

Powell, W., Sheffi, Y., Nickerson, K., & Butterbaugh, K. Maximizing profits for North American Van Lines' truckload division: a new framework for pricing and operations. *Interfaces*, *18*, 21-41. Retrieved October 20, 2013, from the Inspec database.

Quarterly Retail E-Commerce Sales. (2014, February 18). . Retrieved November 14, 2014, from https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf

APPENDIX A – Data Sample

Order Number	Customer ID	City	State	Day	Time Between Orders	Cumulative Time (Hours)	Number of Totes
1	31781	Boston	MA	1	0.020	0.02	7
2	9234	Boston	MA	1	0.037	0.06	6
3	8002	Boston	MA	1	0.011	0.07	4
4	44511	Newton	MA	1	0.050	0.12	7
5	43819	Dorchester	MA	1	0.006	0.12	4
6	47458	Jamaica Plain	MA	1	0.006	0.13	1
7	34804	Newton	MA	1	0.018	0.15	3
8	38792	Boston	MA	1	0.085	0.23	7
9	36072	Roxbury	MA	1	0.014	0.25	7
10	45196	Boston	MA	1	0.029	0.28	1
11	39512	Boston	MA	1	0.312	0.59	2
12	14374	Cambridge	MA	1	0.021	0.61	3
13	9904	Boston	MA	1	0.030	0.64	7
14	14173	Allston-Brighton	MA	1	0.016	0.65	7
15	35431	Boston	MA	1	0.071	0.72	1
16	44415	Boston	MA	1	0.036	0.76	4
17	20154	Boston	MA	1	0.104	0.86	6
18	465	Somerville	MA	1	0.021	0.89	7
19	33562	Boston	MA	1	0.025	0.91	6
20	41815	Cambridge	MA	1	0.113	1.02	1
21	45751	Brookline	MA	1	0.090	1.11	2
22	18686	Dorchester	MA	1	0.039	1.15	3
23	33828	Boston	MA	1	0.104	1.26	6
24	263	Brookline	MA	1	0.063	1.32	2
25	47268	Boston	MA	1	0.052	1.37	2
26	12838	Cambridge	MA	1	0.086	1.46	2
27	45822	Charlestown	MA	1	0.012	1.47	4

APPENDIX B – Model Parameters

Constant	Value
Fleet capacity	20 trucks
Revenue per tote	\$40
Cost per tote	85% of revenue/tote
Fixed cost per truck	\$100
Penalty for not using the truck	\$60
Order acceptance period	12 hours
Average number of totes per order	4 totes
Average revenue per order	4 totes * \$40 = \$160
Number of local tours per region	1
Network constant, k	0.765
Cost per distance	0.75 / mile
Cost per stop	\$2
Cost per unit per stop	\$0.25
Constant regression coefficient	-1.578
Line haul regression coefficient	-0.08093
Number of totes regression coefficient	0.5559
Threshold value for logistic regression	0.53
Threshold probability for meeting breakeven number of orders	0.7

APPENDIX C – Model Output Summary

Type	Day	Fleet Size	Trucks Used	Orders Accepted	No of Totes	Revenue	Cost	Profit
Myopic Manager	8	1	1	14	60	2400	2091	209
Our Model	8	1	1	14	60	2400	2091	209
Optimal	8	1	1	13	80	3200	2783	317
Logistic	8	1	1	7	44	1760	1528	132
Myopic Manager	8	2	2	37	140	5600	4899	501
Our Model	8	2	2	37	140	5600	4899	501
Optimal	8	2	2	27	160	6400	5571	629
Logistic	8	2	2	21	121	4840	4213	427
Myopic Manager	8	3	3	43	160	6400	5619	481
Our Model	8	3	3	47	183	7320	6414	606
Optimal	8	3	3	43	240	9600	8360	940
Logistic	8	3	3	35	201	8040	6998	742
Myopic Manager	8	4	4	53	203	8120	7134	586
Our Model	8	4	4	65	263	10520	9208	912
Optimal	8	4	4	59	320	12800	11153	1247
Logistic	8	4	4	49	281	11240	9788	1052
Myopic Manager	8	5	5	71	283	11320	9927	893
Our Model	8	5	5	68	272	10880	9543	837
Optimal	8	5	5	74	400	16000	13945	1555
Logistic	8	5	5	54	313	12520	10915	1105
Myopic Manager	8	6	6	74	292	11680	10262	818
Our Model	8	6	6	87	352	14080	12344	1136
Optimal	8	6	6	88	480	19200	16751	1849
Logistic	8	6	6	55	318	12720	11115	1005
Myopic Manager	8	7	7	93	372	14880	13063	1117
Our Model	8	7	7	106	432	17280	15161	1419
Optimal	8	7	7	119	560	22400	19576	2124
Logistic	8	7	7	67	396	15840	13849	1291

APPENDIX D – Model Output at Order Level

Order No.	City	Day	Cumulative Time	No. of Totes	Myopic Manager Decision	Optimal Decision	Our Model Decision	Logistic Regression Decision
1	Boston	1	0.02	7	1	1	1	1
2	Boston	1	0.06	6	1	1	1	1
3	Boston	1	0.07	4	1	1	1	1
4	Newton	1	0.12	7	1	1	1	1
5	Dorchester	1	0.12	4	1	1	0	0
6	Jamaica Plain	1	0.13	1	1	1	1	0
7	Newton	1	0.15	3	1	1	1	0
8	Boston	1	0.23	7	1	1	1	1
9	Roxbury	1	0.25	7	1	1	1	1
10	Boston	1	0.28	1	1	1	1	0
11	Boston	1	0.59	2	1	1	1	0
12	Cambridge	1	0.61	3	1	1	1	0
13	Boston	1	0.64	7	1	1	1	1
14	Allston-Brighton	1	0.65	7	1	1	1	1
15	Boston	1	0.72	1	1	1	1	0
16	Boston	1	0.76	4	1	1	1	1
17	Boston	1	0.86	6	1	1	1	1
18	Somerville	1	0.89	7	1	1	1	1
19	Boston	1	0.91	6	1	1	1	1
20	Cambridge	1	1.02	1	1	1	1	0
21	Brookline	1	1.11	2	1	1	1	0
22	Dorchester	1	1.15	3	1	1	0	0
23	Boston	1	1.26	6	1	1	1	1
24	Brookline	1	1.32	2	1	1	1	0
25	Boston	1	1.37	2	1	1	1	0
26	Cambridge	1	1.46	2	1	1	1	0
27	Charlestown	1	1.47	4	1	1	0	1
28	Cambridge	1	1.5	6	1	1	1	1
29	Jamaica Plain	1	1.52	4	1	1	1	0
30	Newton	1	1.57	3	1	1	1	0

APPENDIX E – Model Output at Daily Level

Our Model Output

****************** Day 1 ************************

Truck 1: Boston
No of Orders: 19
Capacity Used: 80
Revenue: 3200.0
Cost: 2799.11
Profit: 300.88

Truck 2: Newton
No of Orders: 21
Capacity Used: 73
Revenue: 2920.0
Cost: 2582.47
Profit: 237.52

Truck 3: Jamaica Plain
No of Orders: 9
Capacity Used: 36
Revenue: 1440.0
Cost: 1276.64
Profit: 63.35

Truck 4: Roxbury
No of Orders: 13
Capacity Used: 59
Revenue: 2360.0
Cost: 2069.45
Profit: 190.54

Truck 5: Cambridge
No of Orders: 24
Capacity Used: 80
Revenue: 3200.0
Cost: 2806.55
Profit: 293.44

Summary
Total No of Orders Accepted: 86
Total No of Totes Accepted: 328
Total Revenue: 13120.0
Total Cost: 11534.24
Total Profit: 1085.75