Carrier Strategies in the Spot Trucking Market

by

Paul Jeffrey R. Leopando
A.B. Anthropology
Harvard University, 2009

and

Kyle A.C. Rocca
B.A. Psychology
B.S. Biological Sciences
University of Notre Dame, 2009

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUL 17 2014
LIBRARIES

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

© Paul Jeffrey Leopando and Kyle Rocca. All rights reserved.

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

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

Signature redacted

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

Signature redacted

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

Signature redacted

Accepted by .....
Dr. Yossi Sheffi
Elisha Gray II Professor, Engineering Systems
Professor, Civil and Environmental Engineering
Director, MIT Center for Transportation and Logistics

Carrier Strategies in the Spot Trucking Market

by

Paul Jeffrey R. Leopando and Kyle A.C. Rocca

Submitted to the Engineering Systems Division on May 9, 2014
in Partial Fulfillment of the Requirements for the Degree of
Masters of Engineering in Logistics

Abstract

How an owner-operator chooses a specific load is a relatively unstudied field in transportation literature. Stakeholders in the decisions, such as freight brokers, stand to benefit from a better understanding of the selection process. Using load board data from a freight brokerage, we identified four parameters available to a carrier when a load is presented: length of haul, revenue per mile (RPM), the probability of finding an onward load from the destination, and the required mileage to reposition to the shipment origin. We also identified preferences of the owner-operators based on experience, literature, and the data, such as owner-operators' preference for long haul routes. We tested selection strategies that disintegrated the four load parameters and incorporated owner-operator preferences in a computerized simulation. We found that strategies combining two or more of the identified parameters provide better results in terms of revenue and utilization (% loaded) maximization. Furthermore, we found that including consideration of the empty repositioning distance was critical to success. Our simulated carriers outperformed peers in the dataset by up to 16%. Carriers can apply these insights to improve their operating strategies. Freight brokerages can apply the quantitative approach to advise their carrier clients and optimize the matching of freight with available carrier capacity.

Thesis Supervisor: Dr. Chris Caplice
Title: Executive Director, Center for Transportation and Logistics

Acknowledgements

My heartfelt thanks goes to my partner, Kyle, for her acumen, energy and dedication throughout this project. Kyle, trying to keep up with you has been a great learning experience for me.

I also wish to pay tribute to my *Nanay*, who passed away at 100 earlier this school year. *Nanay* took care of me, fed me, scolded me when I needed it, and made me feel loved, always. In fact, she could not talk about me—even to strangers—without tears coming to her eyes. Because I loved her too, and knew how deeply special I was to her, I felt I could conquer the world.

-Jeff

I would like to thank my Grannabelle and dedicate this paper to her. She is the model of a humble and giving woman. Throughout her life, she has demonstrated strength, resilience, and the power of hard work. She has inspired my passion for adventure with her own travels and supported my insatiable thirst for knowledge. Grannabelle, thank you for paving the road for me and making this all possible.

And of course, I am sincerely indebted to my partner, Jeff. Jeff, thank you for sharing you patience, calm nature, and genius with me throughout the year. Thank you for your willingness to power through 10 hours of thesis work at a time, and for committing seemingly endless Saturdays to simulations. I could not have done this without your partnership, and I know amazing things lie ahead for you.

-Kyle

We would also like to thank our corporate sponsors, specifically Bill, David, Amit, and Ella, for their invaluable support. We appreciate the comments of our coach and editor, Thea Singer, and the instruction and guidance provided by Lenore Myka. We would also like to thank our classmates, firstly for their peer reviews, but mostly for their company and friendship along the way.

Finally, we would like to thank Chris Caplice, our advisor, for sharing his expertise limitedly and allowing us to muddle through on our own, pushing us to write an impactful paper that he could have written in his sleep.

-From both

Contents

Acknowledgements 3
Contents 4
List of Figures 6
List of Tables 7
1 Introduction 8
1.1 Trucking Industry 9
1.2 Motivation 11
1.3 Problem Statement 12
1.4 Partner Company 13
1.5 Chapter Summary 14
1.6 Looking Ahead 14
2 Literature Review 15
2.1 Freight Transportation Marketplace 15
2.2 Shipper Perspective 17
2.3 Strategies of Carriers 18
2.4 Owner-Operator Characteristics and Strategies 21
2.5 Chapter Summary 23
3 Research Approach and Data Preparation 24
4 Data Profiling and Mining 25
4.1 Methodology of Data Profiling and Mining 25
4.2 Results of Data Profiling and Mining 26
4.3 Chapter Summary 33
5 Simulation 34
5.1 Constructing the Load Board 35
5.2 Examining and Choosing Loads 36
5.3 Initial Testing Scenarios 46
5.4 Strategy Design and Testing 54
5.5 Chapter Summary 62
6 Discussion 63
6.1 Discussion of Results and Features of Data Profiling and Mining 63

6.1.1 Data consistencies and inconsistencies with spot trucking market 62
6.1.2 Limitations of data mining analysis 64
6.2 Discussion of Features and Results of the Simulation 65
6.2.1 Discussion of exploratory simulation runs 65
6.2.2 Discussion of comparison of tested carrier strategies 65
6.2.3 Limitations of simulation model 67
6.3 Chapter Summary 70
7 Conclusion 72
7.1 Summary of Main Findings 72
7.2 Contributions and Areas for Future Research 74
References 76

List of Figures

Figure 1 Load decisions available to a carrier at a given time and location 14
Figure 2 Distribution of length of haul (miles), which follows a lognormal distribution..... 27
Figure 3 Average percentage of deadhead miles for five major U.S. carriers..... 28
Figure 4 Average revenue per mile (RPM) earned compared to length of haul for loads in the dataset. 29
Figure 5 Distribution of carriers by fleet size from the dataset provided by the third-party logistics (3PL) provider..... 30
Figure 6 Median revenue per mile (RPM) for carriers, segmented by fleet size, over different lengths of loads hauled 31
Figure 7 Average percentage of deadhead miles for five major United States carriers..... 37
Figure 8 Flow diagram showing simulation progression..... 38
Figure 9 Average length of haul per *L* value 41
Figure 10 Poisson distribution for the probability of loads departing from Massachusetts each day based on the dataset..... 43
Figure 11 Probabilities of at least one load, PX > 0, departing from each state 44
Figure 12 Comparison of revenue ($) and utilization (% loaded) for the five home cities across complete months..... 49
Figure 13 Comparison of revenue ($) and utilization (% loaded) for long-haul preference and short-haul preference for the August simulation run..... 51
Figure 14 Comparison of revenue ($) and utilization (% loaded) across different simulation wait times (hrs) for the August simulation run..... 52
Figure 15 Comparison of revenue ($) and utilization (% loaded) across difference simulated truck speed (miles per hour, mph) for the August simulation run, with linear trend lines. 53
Figure 16 Total revenue ($) and utilization (% loaded) for each home city resulting from the simulated strategies over five months..... 56
Figure 17 Average length of haul (mi/load) and average revenue per mile (RPM) ($/mile) for each home city resulting from the simulated strategies over five months. 59
Figure 18 Average length of haul (mi/load) and average revenue per mile (RPM) ($/mile) for each home city resulting from the simulated strategies over five months. 67

List of Tables

Table 1 Daily Average (Avg) Number of Departing Loads from a Given State 33
Table 2 Summary of Initial Case and Alternative Cases Tested During Exploratory Simulation Runs 48
Table 3 Summary of Strategies Tested in Final Simulation Runs 55
Table 4 Heat Map Showing Total Revenue ($) and Average Utilization (% Loaded) for Each Home City Resulting from the Simulated Strategies Over Five Months 57
Table 5 Heat Map Showing Length of Haul and Revenue Per Mile (RPM) for Each Home City Resulting from the Simulated Strategies Over Five Months 60

1 Introduction

The United States freight market is valued at nearly $650 billion, with trucking accounting for nearly 80% of the commercial market revenues and 70% of freight tonnage (Costello, 2013; Kirkeby, 2014, pp. 11-12). Industry carriers are segmented into two groups, private carriers and for-hire contract carriers, with roughly equivalent revenues earned and freight tonnage carried by each group. Among the for-hire contract carriers, it is estimated that there are nearly 50,000 unique carriers, the majority of whom operate five or fewer motor units (Kirkeby, 2014, p. 12, Mergent Intellect, 2014). Understanding the dynamics of how these carriers select freight would allow freight logistics brokers not only to match shipments and carrier capacity for the majority of their business more efficiently, but also to provide insights into optimal operating models to their carrier customers.

Our study examines the strategic freight-selection decisions by owner-operators in the spot trucking market to draw conclusions about effective choice making. The spot market is defined by its transactional nature, matching a single load to a single carrier. This is in contrast to the contract market, where a shipper projects a certain volume between a particular destination and origin and chooses a carrier to service that volume for an extended period of time. Using fleet size as a proxy, we defined owner-operators as those carriers owning five or fewer vehicles. Owner-operators are the major carrier in the spot market by number of businesses, and theoretically could apply nearly 50,000 unique strategies.

Using masked freight shipment data, we characterized the important features of a shipment. We tested different strategies for prioritizing these features, and compared the results to determine the impact of such strategies on the operating success of an individual owner-operator. Although we focus on this subset of the market, our methodology could be widened to include other carrier and market types.

1.1 Trucking Industry

The freight market involves two crucial parties, the shipper and the carrier. Often, these groups will use a third-party logistics (3PL) provider to manage their transportation and other logistics services. These 3PLs provide freight brokerage as a service, matching carrier capacity and shipment availability. In addition to the more than $600 billion United States freight transportation market, brokerage and 3PL services generate more than $140 billion in revenues (Armstrong & Associates Inc., 2014; Goddard, 2014). Although the 3PL is not an essential party like the shipper and carrier, the services they provide have been a growing, and increasingly key, element of the market.

Shippers use a variety of contractual arrangements with carriers to move freight, and the nature of these arrangements influences the involvement of the 3PL. Shippers either contract with a carrier on a long-term basis, or make one-time transactions with them. In a long-term contract, a shipper usually pays a carrier an agreed rate per load to cover a lane over a time period of one or two years. The role of 3PL providers ranges from complete management of a contractual arrangement, including Request-for-Proposals (RFPs), bidding by carriers, assignment of lanes, and tracking performance indicators, to no involvement whatsoever.

Alternatively, when freight movement is not covered by an existing contract, shippers have the option of entering into one-off transactions with carriers on the spot market. Each of these transactions is negotiated separately, resulting in rate volatility. In the spot market, carriers may rely on a broker's assistance to find loads or bypass them entirely by accessing electronic load boards directly.

When the shipper has relatively consistent and predictable volume and frequency of freight, the contract market is used. Here, the shipper will assign lanes, each consisting of an origin and destination pair (for example, from Houston to Chicago), to a particular carrier, who will receive an agreed upon rate for shipments on this lane over a certain time horizon. This is the more common arrangement in the for-hire marketplace, though the spot market gained significant share between 2010 and 2011 with growth around 30% (TransCore, 2011a; 2011b). A nuance of the freight transportation market is that carriers have the option to decline individual shipments on their contracted lane, though this is outside of the scope of this thesis, and is covered by an existing body of research (for example, see Kim, 2013, *Analysis of truckload price and rejection rates* (Master's Thesis) for more information on the line of research). As a result, a shipper will have multiple carriers to cover a specific lane at differing rates, and if all these carriers reject the load, the freight may end up in the spot market.

Additionally, shippers depend on a variety of carrier types to move freight. A private fleet is a trucking operation owned by a company that needs to ship material goods but whose primary business is not trucking; retailers, for example, often operate a private fleet to efficiently deliver their products to consumers. Private fleets represent nearly 37% of the commercial transportation revenues and 34% of freight tonnage; therefore private fleets comprise nearly half the trucking market in both tonnage and revenues (Kirkeby, 2014, p. 11). Additionally, shippers

can outsource the transportation using a dedicated fleet. The carrier owns the vehicles in a dedicated fleet, but the fleet is licensed for exclusive use by one or more shippers. Dedicated fleets ensure that carrier capacity is always available, and are used exclusively under contractual arrangements.

Alternatively, shippers can procure capacity by using independent carriers-for-hire. These carriers operate in both the contract and spot markets, collectively moving more than 4.5 billion tons of freight each year, accounting for 33% of the commercial freight market (Kirkeby, 2014, p. 11; TransCore, 2011b, p. 3). Shipments can either be truckload (TL), with loads greater than 10,000 pounds, or less-than-truckload (LTL), with shipments weighing 10,000 pounds or less. Typically, several LTL shipments can be combined, whereas there will only be one TL shipment per motor unit per journey. For-hire market revenues are roughly divided into 85% earned by TL carriers and 15% by LTL carriers (Kirkeby, 2014, pp. 11-12). Further, the majority of TL shipments are tendered to national carriers, with the largest earning more than $5 billion in total revenues (p. 12). The remainder of the market is highly fragmented, with nearly 40,000 carriers earning less than $1 million in revenues, and more than 70% having five or fewer vehicles (Mergent Intellect, 2014; Kirkeby, 2014, p. 12).

1.2 Motivation

This thesis focuses on loads brokered in the spot market to one of these independent carriers-for-hire operating less than six vehicles. When a broker is used to match carrier capacity to available freight, the shipper pays the broker, and the broker will subsequently pay the carrier. The broker creates revenue by securing a carrier at a price lower than that paid by the shipper, keeping the margin as a fee. Brokerage fees drive the more than $50 billion in revenues

generated by 3PL services, and therefore brokers are motivated to find new and better ways to optimize their matching procedures to serve both their shipper and carrier customers (Goddard, 2014; Mergent Intellect, 2014). When a carrier operates efficiently, he is able to offer a lower cost to the broker, and the broker can maintain his margin while passing along savings to the shipper. Effective matching enables the carrier to achieve greater utilization and revenues, and distinguishes the broker as an effective business partner, helping him gain more profit through increased business.

An understanding of how carriers operate in the spot market will provide insights that can improve the assignment of freight in the spot market. This intelligence can be used to counsel carriers on the financial benefits of adjusting their operating procedure. Furthermore, brokers can use their knowledge of a carrier's individual strategy to find freight that is beneficial to both the carrier and the shipper.

1.3 Problem Statement

Given a set of available loads, an individual carrier must decide which, if any, to take. The decision depends on characteristics of the load, such as revenue, length of haul, and destination, as well as characteristics of the carrier, such as current location. Depending on the choice, new loads will be available at the next location in time. Figure 1 shows a carrier positioned at Location A at Time 1, faced with the decision to stay at Location A, to move empty to Location B, to take a load to Location C for Time 2 or to take a load to Location D for Time 3. This choice then results in a new decision point. Moreover, additional loads may become available that are not currently known. Therefore, a carrier seeks to choose the best load at any given time that will also lead to known or likely optimal loads at the next decision point.

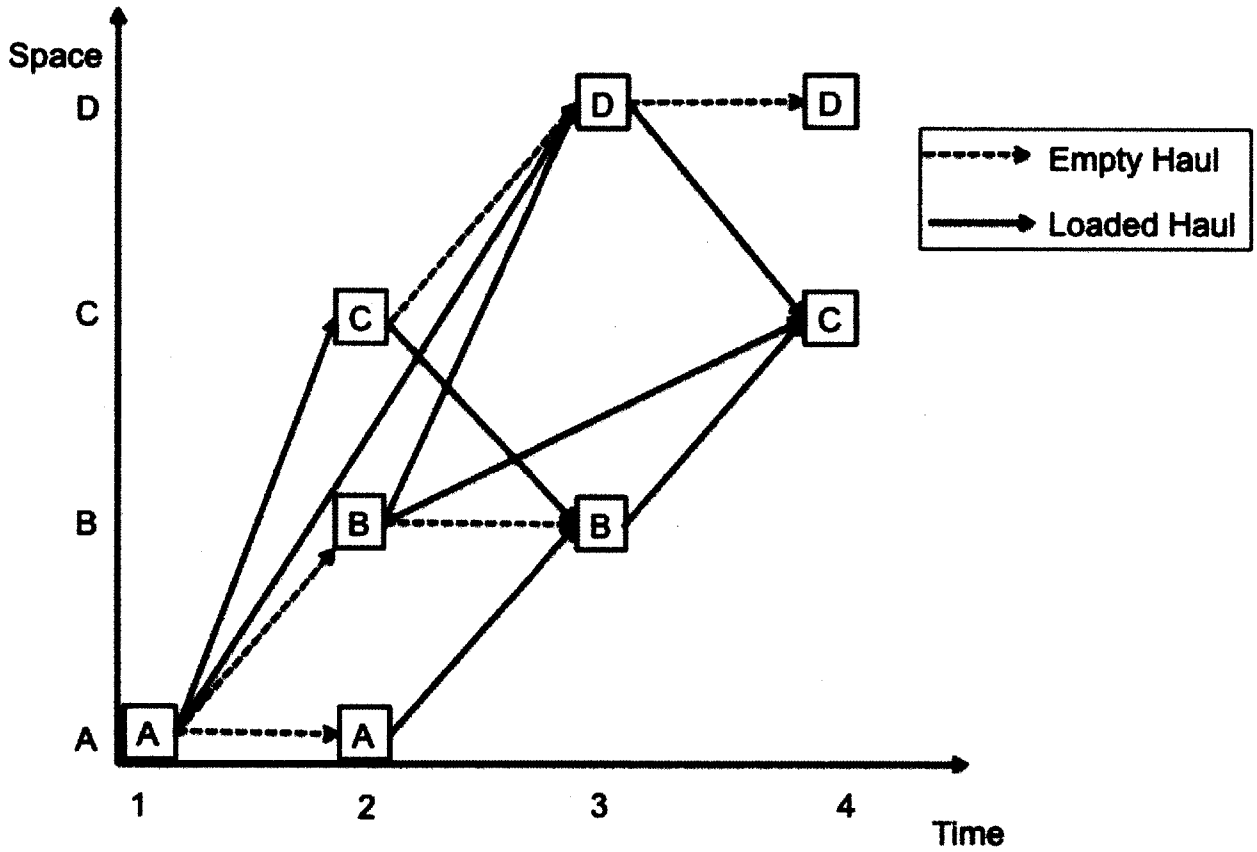


Figure 1 Load decisions available to a carrier at a given time and location. A carrier located at Point A at Time 1 faces four options: stay in place until Time 2, move without freight to Point B, carry a load to Point C that needs to be dropped off by Time 2, or carry a load to Point D by Time 3. Upon making his decision, the carrier is faced with a new set of options. A carrier aims to find the profit-maximizing combination of movements.

1.4 Partner Company

To conduct our research, we partnered with a non-asset owning freight brokerage firm and 3PL provider. The 3PL delivers consultative and operational services to its shipper and carrier clients, including Fortune 500 companies and numerous carriers. Specifically, one such service that the 3PL offers is matching available freight in the marketplace with available capacity from their 40,000 carrier clients using cold-call techniques.

1.5 Chapter Summary

There are an estimated 30,000 owner-operators with five or fewer vehicles servicing the for-hire TL market. Freight brokerages depend on these operators to provide capacity for shipments in the volatile spot market. Brokers, keen to reduce the time required to match shipments with carriers, seek to understand the strategies employed by carriers as well as to provide insights to help their carrier clients operate better. Therefore, constructing and testing possible strategies used by the owner-operators, the majority class by number, and comparing these strategies to one another allows us to analytically evaluate these strategies and derive valuable insights.

1.6 Looking Ahead

The materials in this thesis are organized into six chapters. Chapter 2 covers literature relevant to our research question, and highlights the gap that our research intends to fill. Chapter 3 covers our research approach and data preparation. Chapter 4 discusses the methodology and results of data mining analysis, and Chapter 5 reviews the simulation design, its initial application, and the results of testing eight distinct carrier strategies. Chapter 6 discusses the results presented in Chapters 4 and 5, including possible explanations for observations. Chapter 7 closes the thesis, highlighting key results and providing vision for future applications and additional research.

2 Literature Review

The literature reviewed is grouped into four thematic sections: an overview of the marketplace for freight over road transport, the shipper perspective, carrier load selection strategies, and information specific to the carrier type of interest to this thesis, the owner-operator.

2.1 Freight Transportation Marketplace

There are many different arrangements matching shippers with carriers, using different carrier types and contractual agreements. Furthermore, there are also many different marketplaces used. In the years following the deregulation of the United States transportation industry in the 1980s, barriers to entry were removed and carriers flooded the market with excess capacity. Internet-based electronic exchanges abounded to enable shippers to manage the volume of carriers, often matching shipments with the lowest bidding carrier. Optimization-based auctions followed in the late 1980s, first applied by Reynolds Metal Company. Reynolds offered their projected freight volume over origin-destination lanes for bidding and made a centralized decision in awarding these projected shipments to carriers. To find the optimal carriers, Reynolds employed an optimization-based auction using mixed integer programming, also known as the *winner determination problem* (WDP) (Moore Jr., Warmke, & Gorban, 1991). Complex combinatorial auctions followed, allowing carriers to bid on packages of lanes rather than on each origin-destination pair individually. Today, variations of combinatorial auctions are still applied and are enabled by the use of specialized software (Ledyard, Olson, Porter, Swanson, & Torma, 2000; Ledyard, Olson, Porter, Swanson, & Torma, 2002; Caplice, 2007, p. 426).

According to Caplice (1996), a carrier's bid is dependent on hedging against several contributors of uncertainty (p. 15). When bidding for a particular lane, the carrier must estimate the likelihood of finding a load from the destination point of the lane being reviewed to a new destination point. Loads over the initial lane are known as headhauls, and subsequent loads are known as follow-on loads. The carrier does so to cover the cost of both servicing the lane as well as any waiting time or miles driven empty without a revenue-earning load (deadhead distance). If the shipper can increase the likelihood of a follow-on load for the carrier, this uncertainty can be managed. While early optimizations such as the WDP did not incorporate these compexities, combinatorial auctions allow for all-or-nothing bids for return routes on lanes or other tours, packages of lanes that fit the carrier's operating network (p. 16). A shipper's ability to match complementary lanes for the carrier during the bid process is critical to the carrier (Caplice & Sheffi, 2003, p. 115). Moreover, the shipper must actually tender loads over the contract period, as payment to the carrier is made per shipment and not for the projected volume (p. 115).

Relatedly, the carrier's bid will depend on the accuracy of the shipper's forecast; carriers will buffer their costs against this uncertainty, again adding additional margin and driving up shipping cost (Caplice, 1996, p. 16; Caplice & Sheffi, 2003, p. 115). While this contributing factor to uncertainty can be managed by investment in better forecasting, the spot market is designed to address demand that is unexpected. As a result, auctions are typically applied to the contract shipping market.

In the spot market, open electronic markets such as TransCore's Load Board, and private electronic markets such as those operated by brokerages, are the primary means of matching loads and carrier capacity (anonymous owner-operators and 3PL employees, personal

communication, October 4, 2013). These are simple tools where shippers and brokers can post shipment origins and destinations and even a suggested price, and carriers and brokers can post the location or projected location of a truck at a specific time. All parties can then engage in matching freight with capacity.

2.2 Shipper Perspective

The excess carrier capacity relative to shipments, allows the shipper to filter and ultimately choose carriers that not only suit its geographic distribution network, but also align with company cost restrictions and reliability standards (Powell, Sheffi, Nickerson, Butterbaugh, & Atherton, 1988, p. 24). Shippers primarily use the contract market; it is estimated that 85% of freight available to for-hire carriers is transported by contracted carriers (by tonnage) (TransCore, 2011b, p. 3). From the shipper perspective, then, the focus of the trucking operation is on optimal carrier selection in the contract market, and thus most of the published literature covers optimal carrier selection by shippers using combinatorial auctions for longer-term contracts.

Nonetheless, we are able to make some generalizations about shippers' values. Shippers are motivated to minimize total inventory and transportation costs (Garrido, 2007, p. 1069). While agreements in the contract market will include stipulations about reliability and performance, cost seems to be the only driver in the spot market (Marcus, 1987, p. 6; Stephenson & Stank, 1994, p. 6).

Minimizing cost in the spot market can be difficult for shippers because of the market dynamics. Spot market loads are characterized by short lead times, volatile market prices, and their transactional, rather than contractual, nature (anonymous owner-operators and 3PL

employees, personal communication, October 4, 2013). Service availability and reliability are not guaranteed in such urgent and unexpected markets (Tsai, Saphores, & Regan, 2011, p. 921). A lane that places a driver in a destination where he is not likely to get a follow on load is likely to be rejected by many drivers unless the potential revenue is sufficient to cover the round trip, and therefore the rates can be extreme (anonymous owner-operators and 3PL employees, personal communication, October 4, 2013). Also, unbound by a contract for future business, the driver may be less diligent at making his delivery at the specified time (Tsai, Saphores, & Regan, 2011, p. 921). Both early and late deliveries are worrisome for a shipper, as unexpected and avoidable costs will be incurred. Shippers and carriers often rely on brokers to qualify operators and match capacity with freight under short time constraints (anonymous owner-operators and 3PL employees, personal communication, October 4, 2013). However, broker services, too, carry a cost. Where shipping demand is uncertain, shippers increasingly depend on the spot market, driving up overall costs (Tsai, Saphores, & Regan, 2011, p. 921).

2.3 Strategies of Carriers

Following deregulation, excess capacity in the market forced carriers to offer substantial discounts; competition was rife and shippers enjoyed low prices and better service (Stephenson & Stank, 1994, p. 6). In the spot market, where carriers largely, if not solely, compete on price, these dynamics continue to prevail. Therefore, carriers focus on finding loads and contracts that maximize revenue and squeezing out operating costs, and only limited attention is paid to relationship management and service standards, relative to the contract market carriers.

A primary cost driver is a deadhead, a movement made without revenue-paying freight, to reposition equipment to more favorable areas. Alternatively, a carrier can offer a rate below

the market price to reposition, preferring some revenue to none (Garrido, 2007). Pairing these movements results in a lower total cost to the shipper and the carrier. This cost advantage is an example of economies of scope, where the cost to serve a set of lanes using one carrier is lower than it would be if there were multiple carriers (Caplice & Sheffi, 2003). The carrier, then, ought to view his movements as comprising a tour rather than single, disjointed actions. He must consider the probability of finding a profitable follow-on load at the destination in addition to the revenue-earning potential of the load at hand (Powell et al., 1988, p. 23).

Carriers can mitigate the uncertainty of securing onward movements by balancing spot market loads and the more predictable contract loads. Research indicates that carriers earn the most revenue when they balance their portfolio of contract and spot market hauls (TransCore, 2011a, p. 2). There is however debate over how extensively carriers use the spot market and load boards, and what the ideal balance is. An industry report stated that spot market loads represent less than 10% of a carrier's business (Kirkeby, 2014, p. 5). However, TransCore (2011a) found that a for-hire carrier found 42% of his loads on electronic load boards, typically via a broker, with the remainder tendered via contract or direct contact with the shipper (pp. 2, 5). TransCore operates a load board and the statistics it cites are from its own marketing material; however the material was based on responses primarily from small sized carriers, more closely aligned with the carriers of interest in this research. TransCore (2011a) also found that carriers who use load boards moderately, procuring between 31% and 60% of their loads from the boards, earn the highest monthly revenue per truck, outperforming the average by nearly 8% ($1,378) (p. 4). These carriers also have better utilization – more freight-carrying miles driven and longer hauls – than their peers. Carriers who use load boards infrequently drive fewer miles overall and per haul, and more miles empty, resulting in the lowest utilization; this results in higher than average

revenue per mile (RPM) but lower overall revenue. Carriers who use load boards frequently have much fewer empty miles, but the longest average haul, and the lowest RPM and overall revenue.

The length of the haul influences carrier behavior in choosing subsequent loads. In their research on this phenomenon, Kafarski and Caruso (2012) observed that carriers are willing to front haul 100 to 140 miles without a corresponding backhaul due to sufficient equipment utilization (p. 42). This carrier would be able to return to the demand-rich area he started out from and then pick up another load on the same day. He could again backhaul empty, achieving profits for the day despite two empty backhauls. Beyond the 100- to 140-mile range, carriers tend to look for a backhaul load at the same time as securing their front haul to cover costs. Above 400 miles, carriers are more than likely committed to a multi-day journey, and can use their time on the road to secure their next load.

Other characteristics of a load beyond location and length of haul factor into a carrier's willingness to accept the job. Kafarski and Caruso (2012) focused on low- and medium-volume contracted lanes and rejection rates. They found that shorter lead time, long dwell time, adverse weather, and low suggested prices were key factors in driving rejection rates (p. 48). Lead-time is the time from when a load is offered to the pick-up time, whereas dwell time is the time when the truck remains in one position waiting for loading or unloading. These findings are supported by the research of Caldwell and Fisher (2008) in the contract market, who found that, up to a point, a longer lead time drove down real-time rejection rates and thus lowered cost to the shipper, and allowed the carrier to more effectively plan prior and subsequent movements.

Carrier behavior may also be simulated in other types of auction markets. Because the truckload market is highly competitive and typically has excess capacity, shippers see it is a commodity market (Garrido, 2007; Marcus, 1987, p. 6; Powell et al., 1988, p. 24). These

features allow comparison of this market to that for substitutable retail goods. For example, Zeithammer (2003) found that bidders bid lower on an item when the good is expected to be available again in the next few auctions (p. 84). This is a parallel situation to a shipper offering two loads on the same route, and could contribute to the rate volatility seen in the spot market. Goes, Karuga, and Tripathi (2012) found a distinction between the strategies of buyers with multi-unit demand and those of bargain-hunters with single unit demand. In this case, buyers with multi-unit demand employed late bidding strategies, while those who sought just one unit bid early. This may extend to carriers looking for a specific backhaul route as opposed to any haul with a sufficiently high rate.

2.4 Owner-Operator Characteristics and Strategies

Characteristics of owner-operators relate to their defining feature – their independence. Surveys of owner-operators reveal a preference for being one's own boss rather than constrained by a dispatcher (Wyckoff & Maister, 1985, p 64). However, research shows that owner-operators are often excluded from markets where coordination of multiple movements is required (Nickerson & Silverman, 2003, p. 116). This may explain the volume of owner-operators in the truckload spot market, as well as their comparatively limited involvement in the LTL and contract markets. Cantor, Celebi, Corsi, and Grimm (2013), however, suggest that shippers partner with owner-operators on a longer-term basis, as owner-operators have lower incidences of crashes. The authors suggest this may be due in part to the owner-operators' dependence on having a functional truck and safe reputation to continue in the business. Additionally, they found that this incidence wanes the more parties the carrier is involved with, again supporting a role for owner-operators in the contract market.

Briefly, from a shipper perspective, using owner-operators can be both more convenient and less costly (Peoples & Peteraf, 1995). Cantor, Celebi, Corsi, and Grimm (2013) found that shippers appreciated the flexibility afforded by using owner-operators, who are less constrained than company drivers. Additionally, owner-operators are less costly to shippers than company drivers for-hire (Belzer, 1994; Nickerson & Silverman, 2003, p. 94). A new owner-operator could make more money driving for a company, but operating independently allows him to build equity despite limited capital (Wyckoff & Maister, 1985).

This need for independence is influential in driving strategy and, ultimately, profitability. An owner-operator only receives payment for work that he does, and therefore utilization is one of the key components to his success. He is driven to find his next load in order to keep moving. Further, he is highly responsive to the local market, and will relocate his truck to where he is more likely to find a backhaul. The business model of the majority of owner-operators is to seek long hauls and to keep moving, often relying on load boards and brokers and predominantly focusing on spot market loads (anonymous owner-operators, personal communication, October 4, 2013).

The evidence in the literature reviewed supports this emphasis on utilization and willingness to take long hauls; however, the definition of long hauls is not agreed upon. The industry norm for haul lengths is different from that proposed by Kafarski and Caruso (2012); short is less than 500 miles, medium less than 1,000 miles, and long hauls are those 1,000 or more miles in length (Costello, 2013, p. 198). Most trucks are used in the short- and medium-haul category, but owner-operators are more likely to assume long hauls, and represent over 90% of the companies engaged in long distance trucking (Cherry & Adelakun, 2012, p. 5; Costello, 2013; Rivera, 2014). There is a further push towards utilization specific to the owner-operator;

these individuals typically have loans on their power units, and as a result may be desperate to take a load at any price to gain a short-term cash flow. This further complicates the market, causing large swings in accepted prices for the same service week-over-week (Wyckoff & Maister, 1985, p. 27).

The owner-operator's independent nature seems to have a direct effect on his net income; on the one hand, this translates to a willingness to accept a net income that is less than what he could earn in other roles, such as working as a company driver or in an alternate career. On the other hand, a successful owner-operator is successful because of personal drive and incentive, energy, and an entrepreneurial spirit (Wyckoff & Maister, 1985, p. 145). Ultimately, it is possible to be both an owner-operator and a profitable carrier, but the strategies and guidelines for achieving this have not yet been defined.

2.5 Chapter Summary

The freight market is well studied, but the body of research on trucking is focused mainly on reducing shipper costs in the contract market. Although the spot market makes up a smaller portion of the for-hire trucking industry, it is essential for covering unexpected traffic patterns and demand. Due to excess capacity overall, there is strong competition in this market, leading shippers to view freight movement as a commodity service. Carriers must distinguish themselves with their ability to operate efficiently at a lower cost in order to succeed. Owner-operators, because of their flexibility in routing, lower overhead, and independence, are uniquely positioned in this market. However, there is a vast gap in the research relating strategy, behavior, and revenue maximization. This research aims to close this gap, and provide insights relating operating success and load choice to brokers and carriers.

3 Research Approach and Data Preparation

To determine successful load selection strategies, we explored a set of masked shipment data from the sponsoring third-party logistics (3PL) provider. Using data exploration, we identified defining features of loads and then formulated a utility function that incorporated these features. We simulated the load selection process, applying variable weights on the terms of the utility function to test distinct carrier strategies. Ultimately, we scrutinized how strategic decisions impact carrier success in terms of metrics such as revenue earned and utilization achieved.

Initially, the dataset provided consisted of 258,596 loads offered for tender through the freight brokerage company over a several month period in 2013. The data for each load included fields for load identification (ID) number, equipment required, location coordinates, revenues earned by the carrier, and times when each load was entered into the system, accepted by a carrier, scheduled for pick up, and picked up by the carrier.

The raw data needed to be cleaned for the purposes of our research. We excluded anomalous records, such as those in which a pick-up occurred prior to entering the load into the system. Additionally, we focused on loads carried by the more prevalent dry-van equipment; alternative equipment types, such as refrigerated units, have unique cost structures and were not suitable for comparison. The cleaned and focused dataset, henceforth known simply as the dataset, included 207,592 loads and 13,117 unique carriers. Subsequently, this data was analyzed using data mining techniques (Chapter 4) and further utilized in a simulation (Chapter 5).

4 Data Profiling and Mining

This section of the thesis presents our methodology for, and results from, analyzing the dataset as previously described in Chapter 3. The methodology for this line of inquiry is presented first, followed by presentation of results and initial commentary. The analysis covered here enabled us to derive insights about the freight market and segment it by fleet size to draw conclusions about different carrier types. Further analysis is described in Chapter 6.

4.1 Methodology of Data Profiling and Mining

We conducted summary statistical analyses in order to validate the data and detect important features of loads and carriers. For example, we tested how representative the dataset was of the industry by checking both the distribution of length of haul and plotting revenues against length of haul. To understand the owner-operator segment of the carriers, we segmented the data by vehicle fleet size and isolated carriers having fewer than six vehicles. We determined the average length of haul for all carriers as well as for owner-operators, and we calculated the proportion of loads carried by the latter segment. These analyses were designed to test the assertion that owner-operators carry the majority of loads available in the spot market. We also examined carrier traffic patterns to understand the spatial dispersion preferences of an individual carrier, such as willingness to leave his home state or home region. In addition, we compared average revenue per mile earned by operators exclusively covering short, moderate, and long haul routes similar to those defined by both Kafarski and Caruso (2012), and industry norms (Costello, 2013). Finally, we calculated the average number of loads leaving each state daily to investigate consistency with industry trends.

This data analysis was supplemented by our conversations with employees of the 3PL provider and a few carriers that they worked with. Although no survey was conducted to ensure that opinions were representative, these interviews provided a source for several generally accepted insights regarding the spot market and carrier behavior.

4.2 Results of Data Profiling and Mining

In examining the dataset, we were interested in drawing conclusions about data representativeness, carrier behavior, and load characteristics.

Our initial objective was to determine whether the data was consistent with industry trends. Figure 2 is a distribution of the lengths of hauls. The distribution follows a lognormal distribution, which is aligned with industry trends (anonymous owner-operators and 3PL employees, personal communication, October 4, 2013).

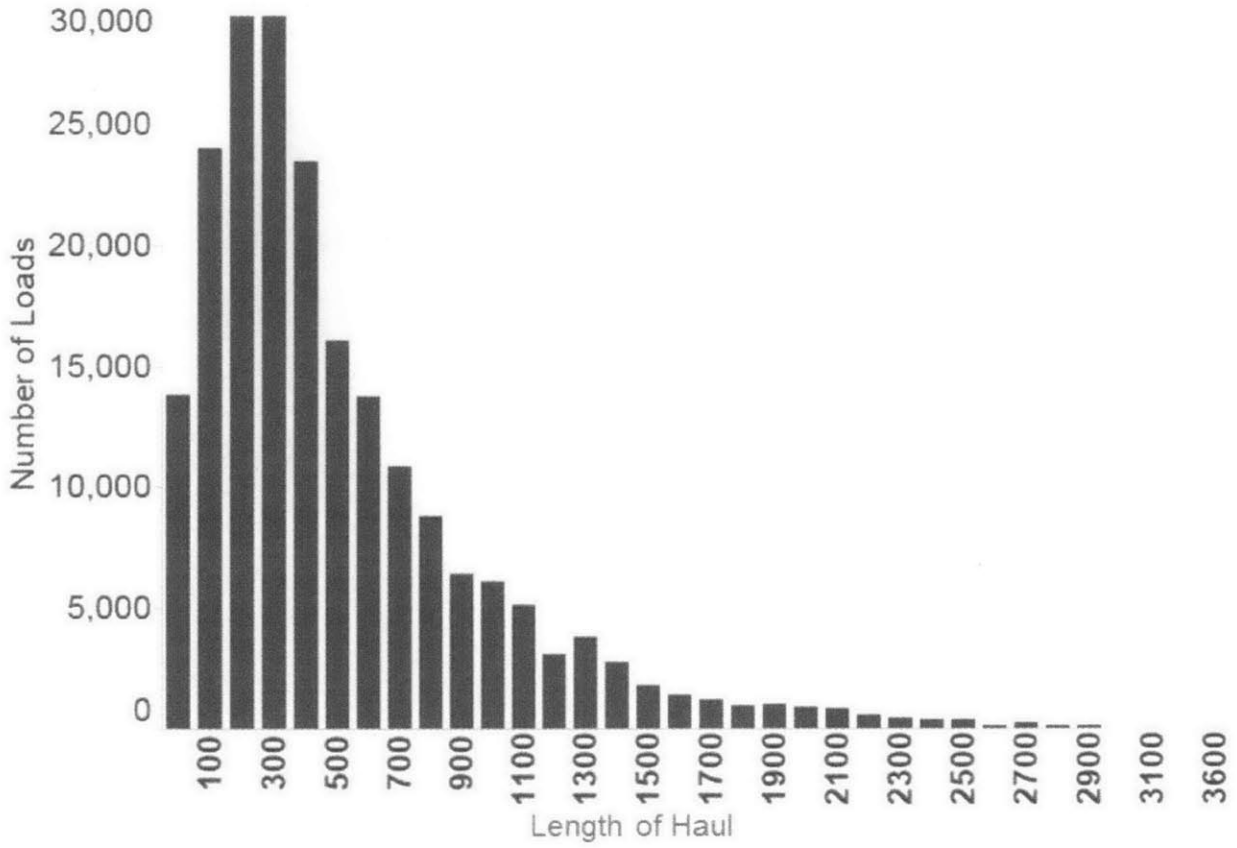


Figure 2 Distribution of length of haul (miles), which follows a lognormal distribution.

Figure 3 compares the average revenue earned to the length of haul, while Figure 4 compares the average revenue per mile (RPM) to the length of haul. These results are also aligned with industry trends, supporting the assertion that the dataset was a representative sample of the industry as described by Caplice (2013a).

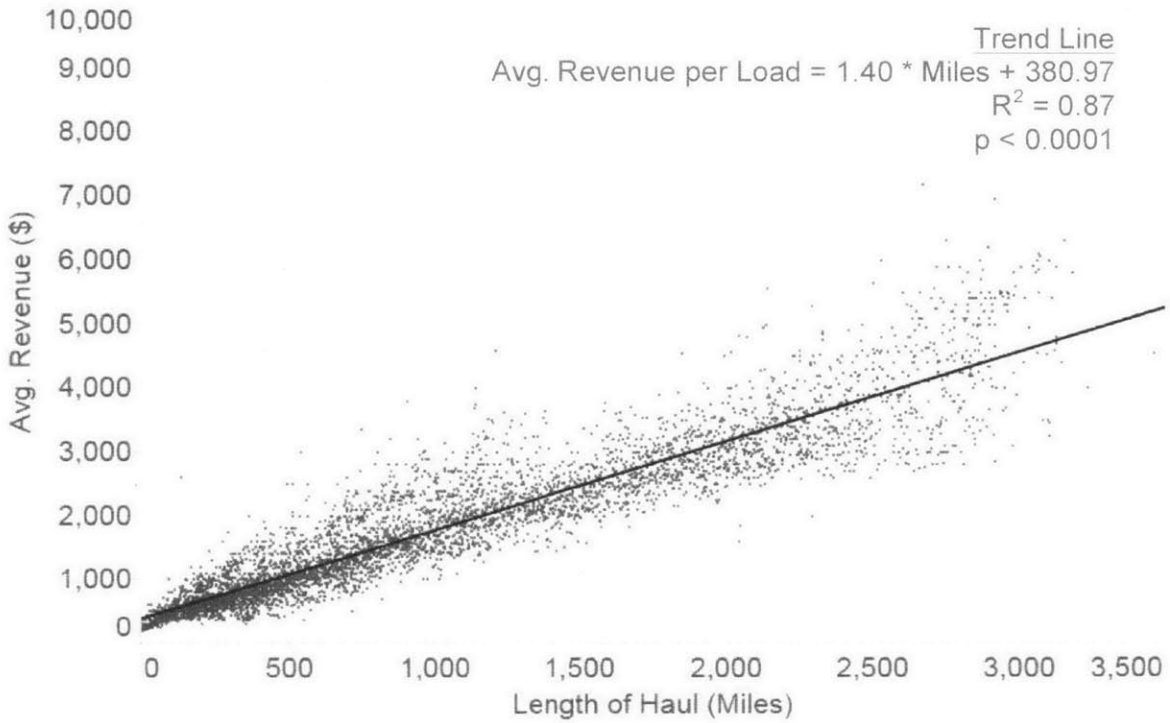


Figure 3 Revenue paid to carrier versus length of haul for all loads in the dataset. Figure 3 also shows a linear trend line with corresponding equation and significance of $p < 0.0001$.

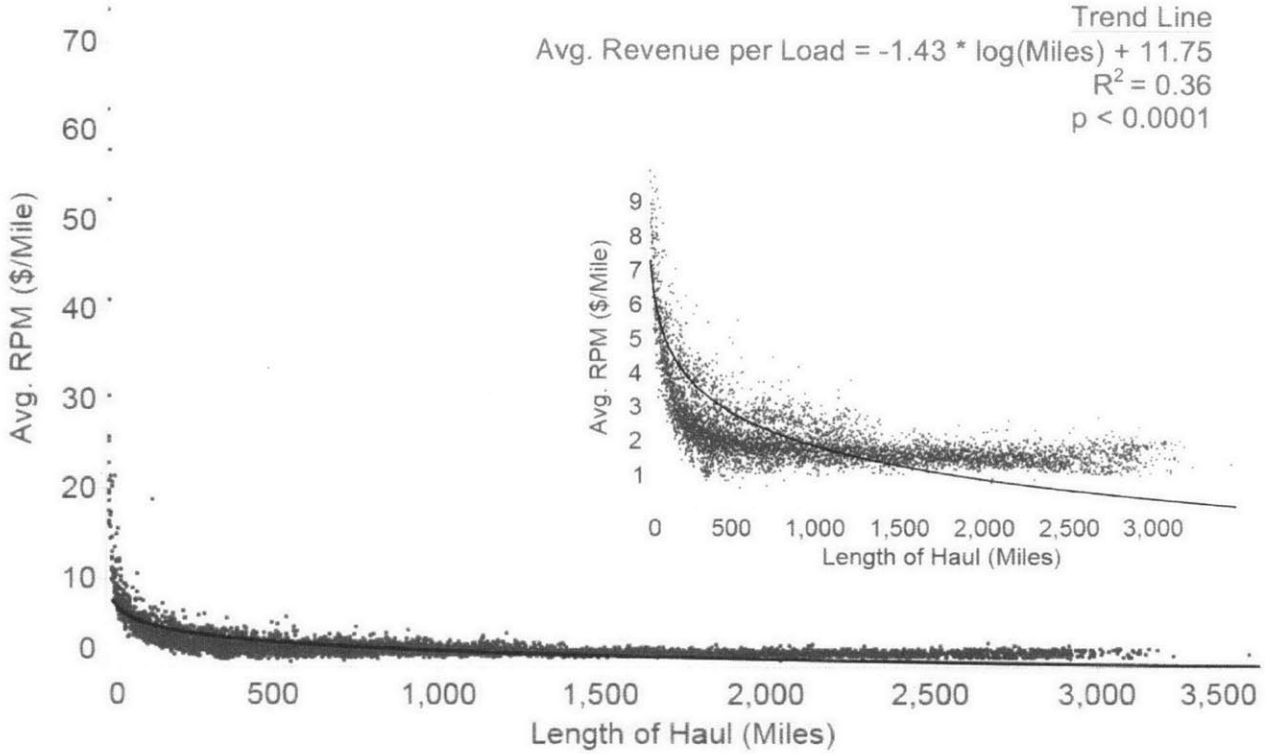


Figure 4 Average revenue per mile (RPM) earned compared to length of haul for loads in the dataset. Inset shows a more focused view for hauls less than $10 per mile, eliminating loads where revenue is dispersed over a short mileage, resulting in a high RPM. Both images show a logarithmic trend line; the trend line equation is also displayed with significance $p < 0.0001$.

We examined the dataset by carrier, segmenting and comparing relevant metrics. We found that, on average, carriers in our dataset hauled 15 loads for the period (four and a half months). When we segmented the data by vehicle fleet size, we found that operators with five or fewer motor units carried only nine loads for the period, and that these loads accounted for 35% of all loads. These owner-operators numbered 7,590, and therefore comprised the majority provider class (58%) (see Figure 5).

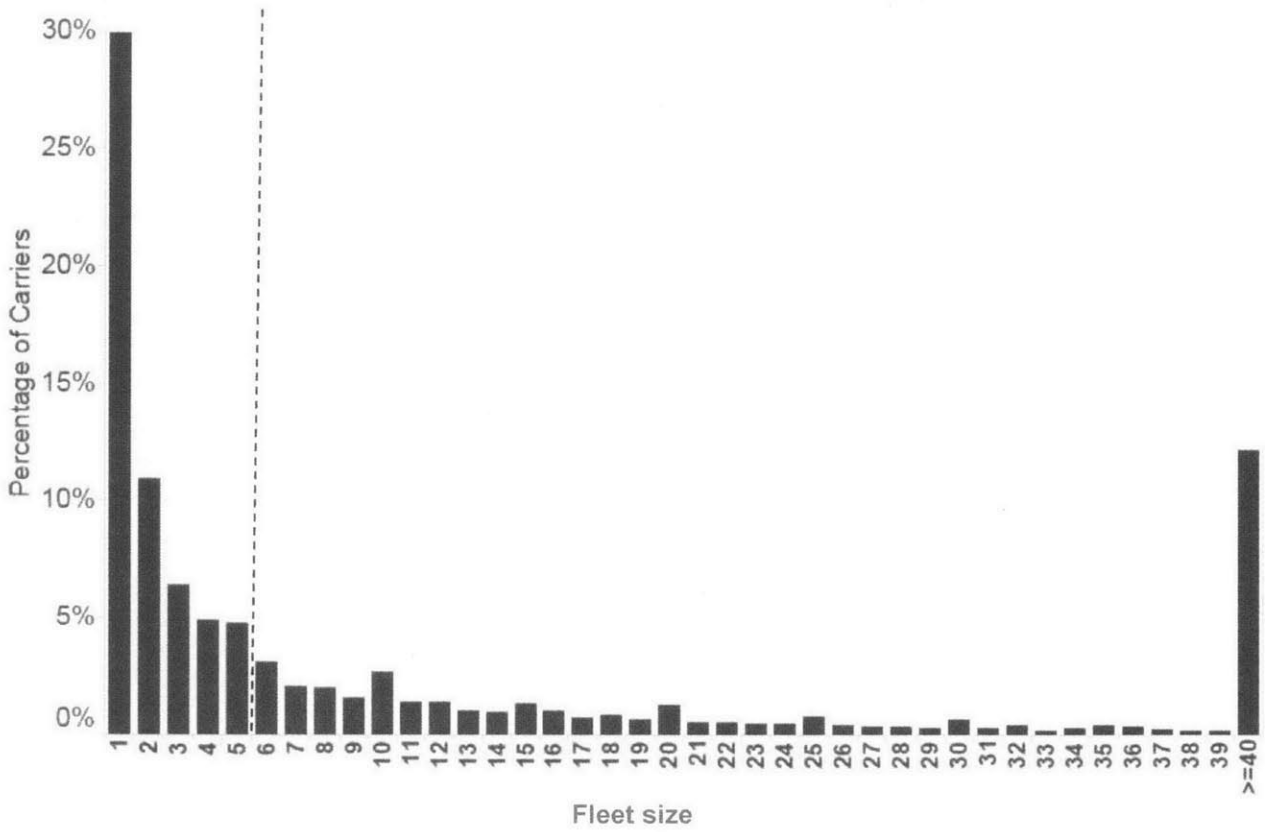


Figure 5 Distribution of carriers by fleet size from the dataset provided by the third-party logistics (3PL) provider. Of the 13,117 carriers, 58% were categorized as owner-operators. The bars left of the dotted line represent the owner-operator segment.

We found that owner-operators had an average length of haul of 593 miles, while the average over the whole dataset was 619 miles. Carriers hauling loads less than 150 miles exclusively were making on average $6.86 per mile, those hauling between 150 and 400 exclusively, $2.46 per mile, between 400 and 1,000 miles exclusively, $1.91 per mile, and above 1,000 miles exclusively, $1.60 per mile.

When compared to carriers with six to ten motor units and carriers with more than ten units, owner-operators enjoyed a higher median RPM for loads over 400 miles (see Figure 6).

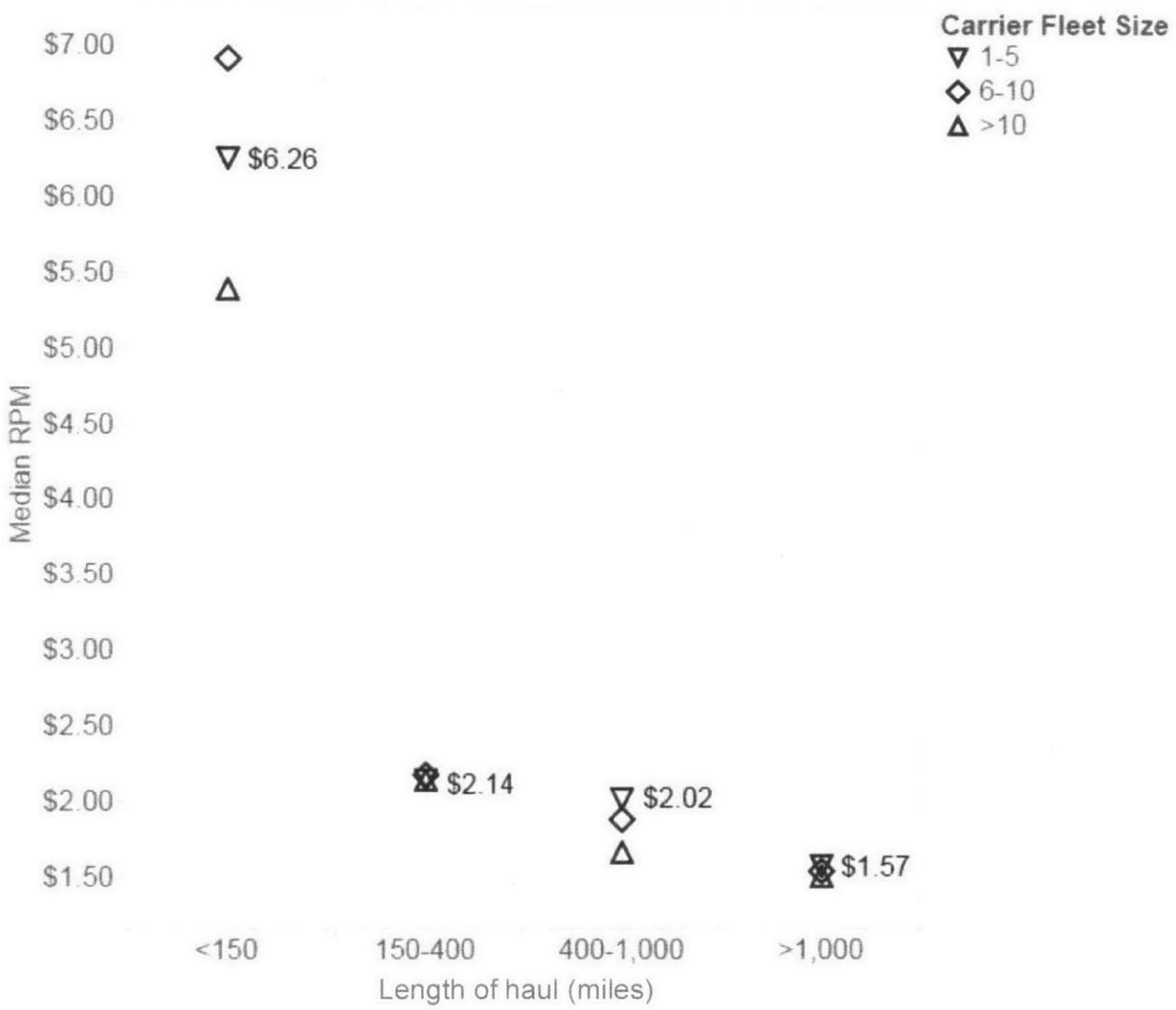


Figure 6 Median revenue per mile (RPM) for carriers, segmented by fleet size, over different lengths of loads hauled. Mark labels indicate the median RPM in each length of haul band for carriers operating five or fewer vehicles, or owner-operators.

Regarding carrier behavior, we found that 9% of carriers stayed within one state, and 1,958 (16%) carried the bulk (72%) of interstate loads. Owner-operators were even more likely to stay within one state; nearly 12% of owner-operators exclusively carried intrastate loads. This was a surprising result; industry knowledge would have anticipated owner-operators traveling long distances.

We checked the assumption that loads with short build-to-book times were more attractive. We found that the average RPM of loads booked within a day of becoming available was 23% higher than the overall average RPM. This conformed to our finding from interviews that many owner-operators use RPM as a main metric for evaluating loads (anonymous owner-operators and 3PL employees, personal communication, October 4, 2013).

Last, we looked at the average number of loads leaving a given state to understand the sources and sinks of loads relative to the geography. We calculated the average number of loads per day that left each of the states in the contiguous United States (see Table 1). Although several loads in the dataset had Canadian or Mexican destinations, the dataset did not include any loads originating from these regions, and are therefore excluded from this particular area of analysis.

Table 1

Daily Average (Avg) Number of Departing Loads from a Given State

State	Avg	State	Avg	State	Avg	State	Avg	State	Avg
AL	20.42	IA	9.94	MI	30.34	NM	6.65	TN	28.84
AR	14.07	ID	2.49	MN	11.90	NV	4.15	TX	225.28
AZ	15.36	IL	88.44	MO	25.02	NY	36.12	UT	2.03
CA	129.61	IN	50.98	MS	8.77	OH	58.50	VA	48.73
CO	15.96	KS	12.09	MT	0.36	OK	5.83	VT	1.10
CT	3.62	KY	26.98	NC	49.24	OR	10.36	WA	12.05
DC	0.03	LA	17.36	ND	1.94	PA	73.30	WI	28.79
DE	1.52	MA	7.30	NE	2.44	RI	0.70	WV	1.91
FL	73.75	MD	15.74	NH	0.64	SC	23.48	WY	1.30
GA	108.18	ME	4.40	NJ	49.99	SD	1.19		

4.3 Chapter Summary

We determined through data analysis that our dataset was consistent with trends in the trucking industry, and that owner-operators consistently earned above average RPMs across all haul distances. We also looked at the average number of loads originating in a given state to visualize the geographic distribution of loads. In subsequent investigation, we used these insights to develop and test strategies through simulation.

5 Simulation

Based on the results of our data analysis, we identified characteristics of a load that a carrier uses to make his decision: revenue, length of haul, and destination. Additionally, after our conversations with carriers and industry experts, we found that a feature of the carrier, his distance from a load's origin, was also a critical input. We built a simulation model in Microsoft Excel and Visual Basic for Applications that imitated a carrier's decision-making process, incorporating these features of the load and carrier. Adjusting the model parameters enabled us to test particular behaviors and priorities for the individual carrier. We analyzed the resulting choices of loads and determined how different strategies compared to each other in terms of realized revenue and truck utilization. Truck utilization was defined as a percentage according to the following equation:

$$\frac{Miles\ loaded}{Total\ miles\ traveled} * 100\% \qquad (1)$$

and is henceforth referred to as *truck utilization (% loaded).*

Methodologies and results discussed in this section are organized thematically. Sections 5.1 and 5.2 cover the building of the simulation model. Once constructed, we employed the simulation of carrier decisions in the spot market in two stages. In the first stage, we conducted a large number of trial simulation runs varying the model parameters individually to test their respective effects, if any, on the output metrics (see Section 5.3). Guided by the insights from these exploratory runs, we focused on testing and comparing eight load selection strategies in the second stage (5.4). A chapter summary is provided in Section 5.5.

5.1 Constructing the Load Board

To simulate the real-world process of load selection, our model mimicked a load board showing available loads hourly to a virtual carrier. The load board was built using the data provided to us by the sponsoring 3PL. Loads in this data originally included the following time stamps:

- Build: the time and date when the load was made available, and on which the 3PL was looking to match the freight with carrier capacity, also known as Entry date
- Pick-up Appointment: the time and date when the shipper desired the freight to be picked up by the carrier
- Booked: the time and date when a carrier accepted the freight

We manipulated these time stamps to create a virtual load board for our simulation. A load was made available on the build date indicated in the original dataset, with some minor changes. To ensure consistency with the majority of the dataset, loads entered into the live system more than one day prior to their pick-up appointment were re-timed so that the build date was the day prior to the pick-up appointment, enabling a sequential decision process. We also determined how long, in hours, a load was available before it was booked. We conjectured that this duration reflected the attractiveness of the load, and used it to determine how many times a load would be offered for booking. Specifically, the duration was rounded to the nearest hour, n, and the load was presented $n + 1$ times in the dataset. The most attractive loads disappeared after just one appearance, while others recurred hourly for up to one day. This manipulation aimed to replicate reality, where attractive loads would be accepted by the first available driver and therefore not be offered to any other driver.

The virtual load board only presented loads from 8 a.m. to 5 p.m., standard business hours when a broker-administered board would be maintained. The loads were batched by month for processing convenience, resulting in datasets for June, July, August, September, and half of October.

5.2 Examining and Choosing Loads

At the beginning of each simulation run, a home location, time window for reviewing loads, truck speed, and deadhead ceiling were set. The home location was also the origin point of the simulated driver. During the time window, the simulated driver was presented with all loads on the virtual load board from the time when the truck was available until the maximum wait time was reached; this criterion is also called *wait time* in this paper. Because truck speed affected the trip durations and therefore the times when loads were reviewed, we tested various truck speeds within legal limitations. The deadhead ceiling, c, was the maximum proportion of the length of the haul under consideration that the driver would be willing to drive empty to reach the pickup location. For example, if the load being considered had 450 miles between origin and destination, and $c = 20\%$, the driver could drive up to 90 miles to reach the load origin, else the load was not considered further. Figure 7 shows deadhead mileage percentage for five leading companies. Market data for total empty hauling mileage by carriers is generally cited at around 10%; we tested this value as well as less restrictive bounds, accounting for the possibility that owner-operators do not manage deadhead miles as well as the leading companies do (Kochar, Monahan, Peters, & Ward, 2012; TransCore, 2011a, p. 3).

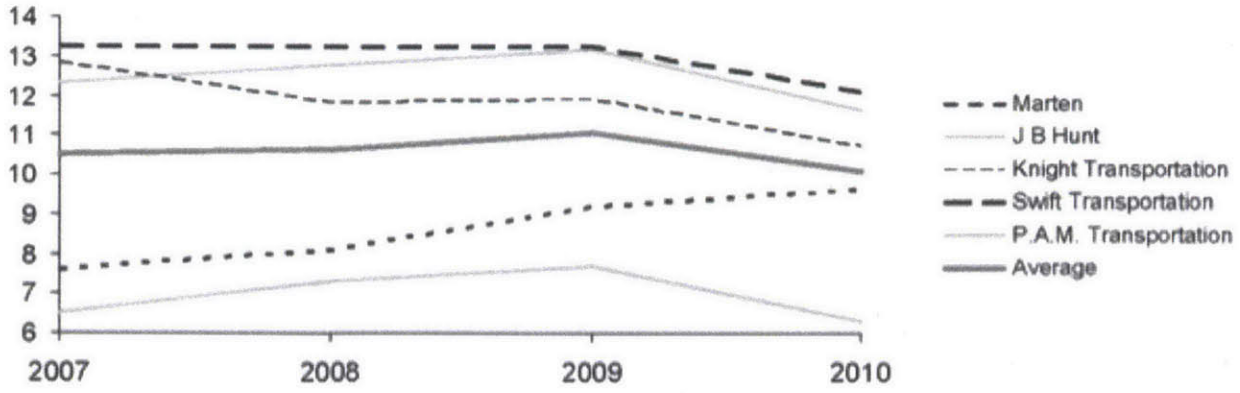


Figure 7 Average percentage of deadhead miles for five major United States carriers.

Source: Kochar, Monahan, Peters, & Ward (2012)

We also programmed an option to exclude loads that originated or terminated in specified states. After interviewing owner-operators and reviewing carrier profiles, we noted that some drivers were unwilling to go to certain locations for business or personal reasons (anonymous owner-operators, personal communication, October 4, 2013). In order to analyze the strategies under less restrictive conditions, we did not use this feature; however, we were aware that this was a critical factor in reality.

Following the entry of the home location and wait time, we initiated the simulation. The carrier examined each load sequentially in a two-step process, checking feasibility and calculating a score (see Figure 8).

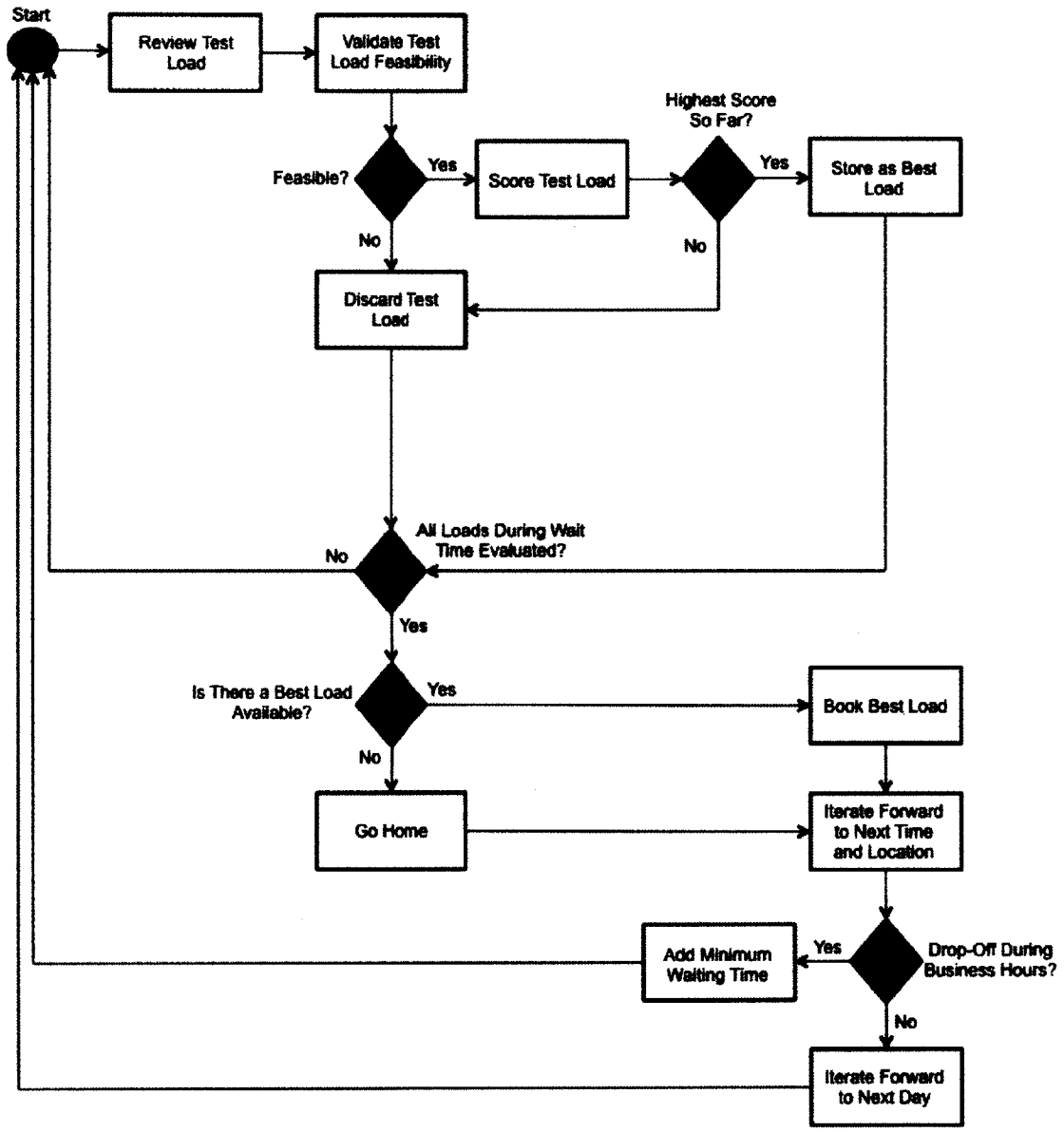


Figure 8 Flow diagram showing simulation progression. The first load available during the review period was considered. The simulation then checked the load against hard criterion. If eligible, the load was then scored according to the strategy in place. The simulation reviewed all loads in the waiting period, and booked the highest scoring load. The simulation then advanced forward to the new location and time as specified by the booked load, and the process began again. In the absence of a qualifying load, the truck would drive home empty, and would resume reviewing loads from the home location.

A load was considered for scoring only if certain hard conditions were met. The equipment type for the freight had to match the carrier's capabilities. Although our dataset included only dry-van loads, the model included further functionality to allow for a broader dataset including loads requiring refrigerated units, flat beds, etc. Additionally, the simulated carrier had to be capable of covering the distance between the current location and the pick-up location by the scheduled appointment time while driving at the set speed, or else the load was discarded. The speed limits selected were arbitrary, though no greater than 55 mph, the lowest maximum allowable speed on highways in the contiguous United States. Research shows that although owner-operators tend to drive more slowly than their company-driver counterparts in order to maximize fuel efficiency, the driving speed is greater than the speed limit on most highways (Kvidera, 2009, p. 29; Cantor et al., 2013, p. 43). In order to promote safe driving, we chose to implement speeds at 55 mph or less. Distance calculations were calculated using the Great Circle Distance Equation:

$$Distance = 1.2 * 3959\left(arccos\left[sin(Lat_O)\, sin(Lat_D) + cos(Lat_O)\, cos(Lat_D)\, cos\left(Long_O - Long_D\right)\right]\right) \quad (2)$$

where *Lat* and *Long* are latitude and longitude, and *O* and *D* are origin and destination, respectively. The radius of the earth is 3,959 miles, and in the United States, road travel is 20% longer, on average, than the straight-line distance between two points (Caplice, 2013b). Finally, the distance from the truck's current location to the origin point of the load being considered had to fall within the deadhead threshold, *c*. Each of these hard criterions had to be satisfied for a load to be considered further.

Loads that met the previous criteria were subsequently scored using soft constraints as follows:

$$Score(L, R, P, D) = \alpha L + \beta R + \gamma P + \delta D \tag{3}$$

where:

$L \in [0,1]$, an index value for the length of the haul

$R \in [0,1]$, an index value for the revenue per mile (RPM)

$P \in [0,1]$, an index value for the daily probability of onward load from the destination state

$D \in [-\infty, 1]$, an index value for deadhead miles to the pickup location

$\alpha, \beta, \gamma, \delta \in [0,1]$, relative weights of the index values

The score is a function of four variables, *L, R, P,* and *D,* which are index values for four key properties of each load: the length of haul, the revenue per mile (RPM), the probability of finding an onward load from the destination state, and the deadhead distance to the pickup location, respectively. Based on our data analysis, existing research, industry knowledge provided by 3PL employees, and carrier interviews, we deemed these four variables to be relevant and available to a driver when deciding whether to book a particular load (anonymous owner-operators and 3PL employees, personal communication, October 4, 2013). The utility function employed is linear by design, allowing for an intuitive weighting process in the style of additive von Neumann-Morgenstern utility (1953).

To determine the L value for each load, we assigned every load its percentile rank and rounded this figure to two decimal places. We chose a ranking approach to lessen the impact of very high outliers. Expressly, a load with a length of haul roughly greater than 1% of all load lengths in the dataset was assigned $L = 0.01$; a load with length of haul roughly greater than that of 2% of all loads was assigned $L = 0.02$; and so on. The index variable L thus took on values in increments of 0.01 from 0 to 1, with the resulting distribution shown in Figure 9. Long hauls were assigned higher L values and thus favored under this design, consistent with general market intelligence on the operating procedures of owner-operators, though we later tested the impact of reversing this preference.

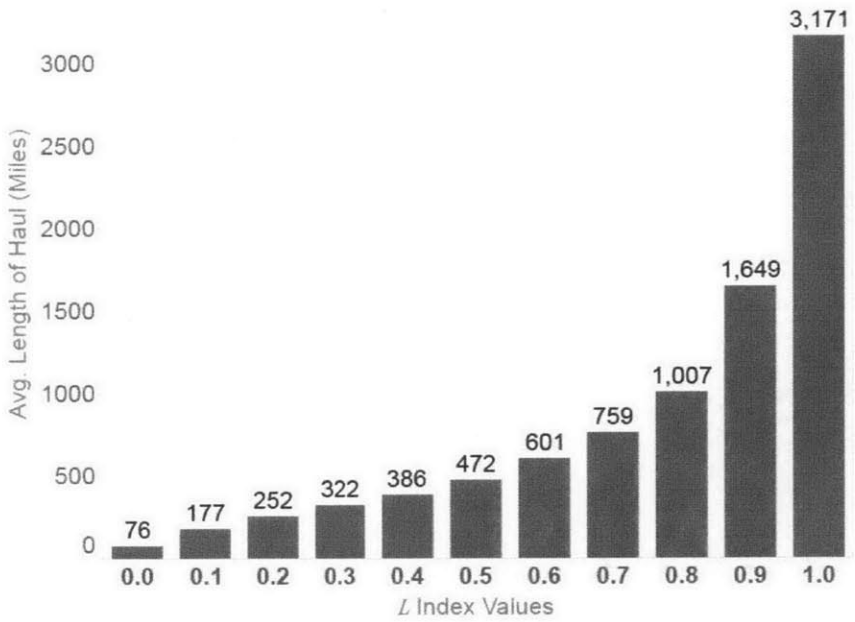


Figure 9 Average length of haul per L value. L values are defined as the percentile rank of the length of haul of each load. In this figure, L values are binned in intervals of 0.1, and the horizontal axis indicates the lower bound of each bin. The average length of haul for these bins is represented on the vertical axis, as well as labeled above each bar.

The value of R for each load was determined through a similar process, except that we accounted for the variability of RPM by region and length of haul. Our data analysis found that total revenue was directly correlated to haul length and that there were regional differences in expected load revenue and RPM. These findings are consistent with industry knowledge and research (Caplice, 1996, pp. 155-187). Thus, each load was ranked not against the full dataset but within a subset of data that included only other loads that terminated in the same region of the United States and whose length of haul fell within a certain range. We grouped all the loads by destination state into five regions, Northeast, South Atlantic, Midwest, South Gulf and West, based on the U.S. Federal Highway Administration's classification (2014), and further subdivided each set into three subgroups by length of haul (<150 miles, 150-400 miles, and >400 miles), roughly in accordance with mileage bands established by Kafarski and Caruso, which distinguish loads requiring more than one days travel from others (2012). Within each of the resulting fifteen subgroups, we calculated every load's RPM percentile rank and rounded this figure to two decimal places as before. The index variable R thus takes on values in increments of 0.01 from 0 to 1, and considers regional and haul length nuances.

The value of P for each load was determined by calculating the probability that at least one load departs from the destination state per day. We assumed that load departure was a discrete random variable following the Poisson distribution, with the probability mass function:

$$Probability[k\ loads\ leaving\ in\ time\ t] = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \in [0,1] \qquad (4)$$

where:

λ: Average number of loads departing state X per time t,

t: Time period, and

k: Number of departures of interest

Values for t and k in (4) were one day and one departure, respectively. Assuming that loads leaving a given state are a Poisson distributed event, these average departures discussed in Table 3 are equivalent to λ in the Poisson distribution equation.

With these values of λ, we derived Poisson probability distributions for departures from each state. To illustrate, Figure 10 shows a partial Poisson distribution for departures from Massachusetts, which has $\lambda = 7.30$. The probability of exactly zero loads ($k = 0$) leaving Massachusetts in a day is $P(X = 0) \approx 0.0003$, and thus, the probability of at least one load leaving the state in a day is $P(X > 0) = 1 - P(X = 0) \approx 0.9993$. Figure 11 captures the probability of at least one load leaving each and every state.

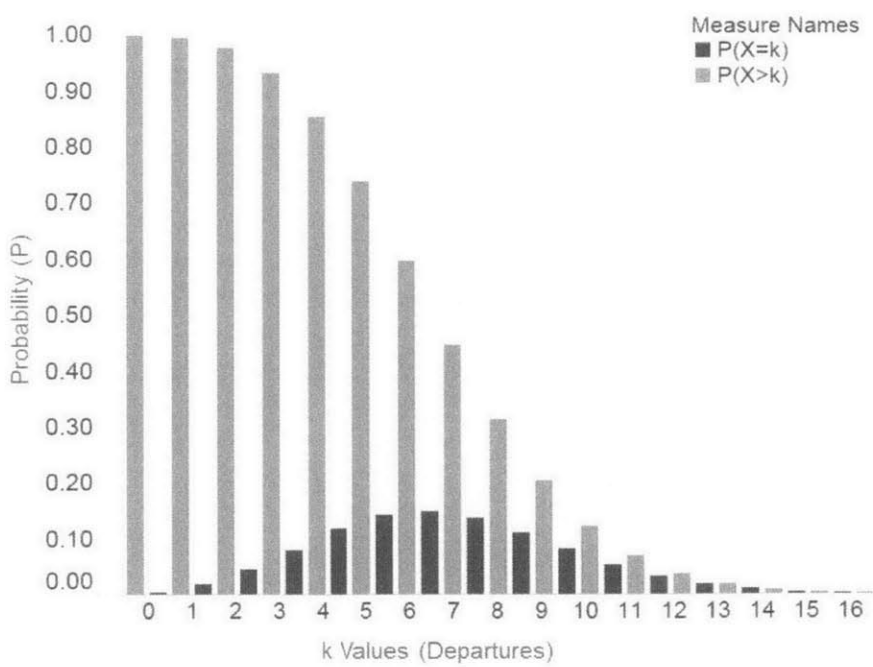


Figure 10 Poisson distribution for the probability of loads departing from Massachusetts each day based on the dataset. The probability of one or more loads leaving Massachusetts is $\mathrm{P}(\mathrm{X} > 0) = 1 - \mathrm{P}(\mathrm{X} = 0) \approx 0.9993$.

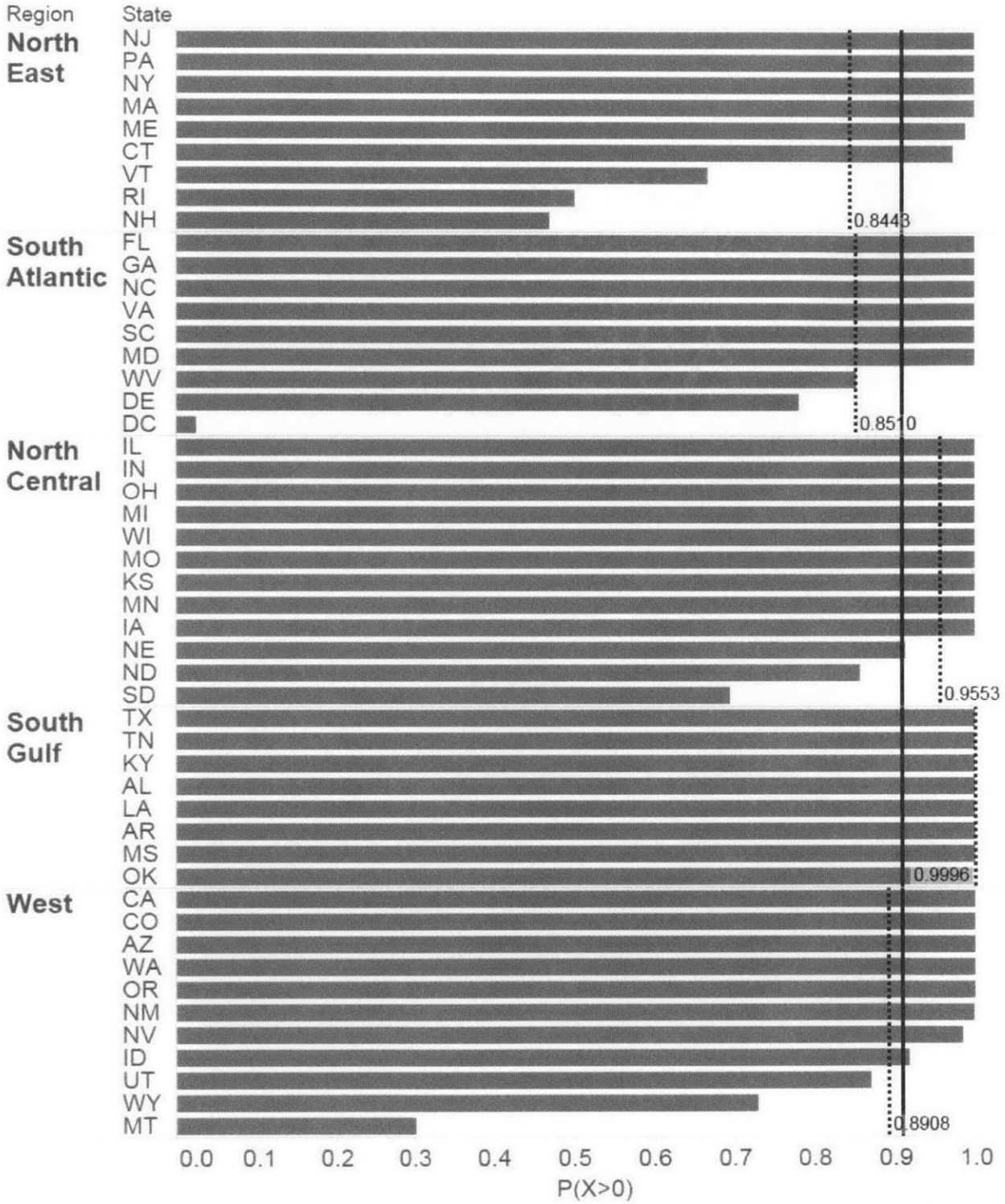


Figure 11 Probabilities of at least one load, $P(X > 0)$, departing from each state. States are grouped by regions as defined by the U.S. Federal Highway Administration (2014) and arranged in order of descending probability in each group. Regional averages are shown in dotted lines, and the national average of 0.9085 is shown in a solid line.

The value of D for each load was defined by the following formula:

$$D = 1 - \frac{Deadhead\ distance}{c * Haul\ length} \in [-\infty, 1] \tag{5}$$

where c was the deadhead ceiling. A load that required no deadhead miles to reach its point of origin achieved a perfect $D = 1$. As deadhead distance increased to the maximum proportion of haul length set by c, the value of D fell to 0. Beyond this threshold, D assumed a negative value, and continued to $-\infty$. The maximum deadhead distance, however, is bounded around 4,500 miles as there are fewer than 4,500 miles spanning the commercial regions of Mexico through the United States and Canada, corresponding to the geographic distribution of loads in the dataset.

In summary, each load was scored as a function of four variables L, R, P, and D, which represented characteristics of the load.

The $\alpha, \beta, \gamma,$ and δ parameters in (3) are the weights assigned to these four variables. To make the resulting score intuitive with a maximum possible value of 1, we bounded the values of these parameters as follows:

$$\alpha + \beta + \gamma + \delta = 1 \tag{6}$$

Manipulating the values of these parameters enabled us to examine the impact of different strategies on a carrier's success. These values can be thought of as how heavily each individual load characteristic (L, R, P, and D) weighs on the attractiveness of a load. For example, a carrier focused solely on taking on the longest haul would put 100% importance on α and 0% on the remaining three factors. Adjusting the weights of $\alpha, \beta, \gamma,$ and δ reflects variations in a carrier's strategy, prioritizing the four characteristics of each load differently.

At the end of the carrier's waiting time, the load with the highest score within that time window was booked to the simulated carrier, and the simulation model advanced in time to the delivery time of the booked load. This time advance was calculated as the quotient of total miles traveled (the sum of the deadhead mileage to the pickup location and the haul mileage) and the 55 mph driving speed. An additional hour of waiting time was added before the simulated carrier could begin looking for the next load; this covered the administrative and personal time a driver would take between hauls.

Following drop-off and the mandated waiting time, new prospective loads were reviewed and scored as above, with the highest scoring load booked. In the event that all available loads in the time window exceeded the deadhead threshold, the simulated truck would drive home and start anew from there. This process continued until the virtual load board was exhausted.

The simulation output showed the loads as they were reviewed, and at the completion of a run, displayed all loads selected. For each scenario, we collected the summary data of the selected loads, including the breakdown of total loaded and deadhead miles, revenue, and maximum, minimum, and average score. By comparing these measures across a variety of scenarios, we were able to clarify the influence of certain parameters on benchmarks of successful operation.

5.3 Initial Testing Scenarios

We aimed to use the simulation to determine how month of data, home location, length of haul preferred, wait time, truck speed, deadhead ceiling, and α, β, γ and δ weights impacted carrier revenues and utilization.

We ran the simulation for each month of loads separately to compare results for each month as well as to sum or average output metrics across all the months. Simulations were run separately on at least three complete months of data for initial findings and for all months for final analysis to detect temporal patterns. For home location, we selected a metropolis from each of the five regions identified by the U.S. Federal Highway Administration – Boston (Northeast), Atlanta (South Atlantic), Chicago (Midwest), Dallas (South Gulf), and Los Angeles (West) (2014). In addition to being geographically dispersed, these metropolises each have a unique operating economy with different costs of materials, capital expenditures, and other expenses (C. Barnes & Co., 2013).

To isolate the effects of an individual variable's impact in each run, we systematically varied one parameter, while holding other parameter values constant. For each variable, we tested an initial case as well as one or more alternatives. For example, we tested the impact of a short- versus long-haul strategy by reversing the percentile rankings for length of haul, L, such that short haul lengths were favored. We tested waiting times between one and nine hours to determine if there was a trade-off between revenues and utilization. We assumed that longer review periods would result in higher scoring loads and greater revenues earned, but utilization would be negatively impacted due to extended time sitting empty. As previously discussed, we also tested various driving speeds. We tested various combinations of α, β, γ and δ weights. To simulate a strategy that considered all load features equally, we set $\alpha = \beta = \gamma = \delta = 0.25$ as the initial case, whereas to examine the effect of focusing on each feature individually, we assigned values of 1 to each weight in series, and held the others at 0. Because of the varying deadhead ceilings seen in operation, we tested varying c values, from the industry average 10% up to 25% in 5% increments. The variable values tested are summarized in Table 2.

Table 2

Summary of Initial Case and Alternative Cases Tested During Exploratory Simulation Runs

Model parameter	**Initial case**	**Alternative cases**
Haul length preference	Long	Short
Waiting time (hours)	1	2, 3, 5, 7, 9
Speed (mph)	55	40, 45, 50
Deadhead ceiling, c (%)	10	15, 17, 20, 25
α	0.25	0, $0.\overline{3}$, 0.5, 1
β	0.25	0, $0.\overline{3}$, 1
γ	0.25	0, $0.\overline{3}$, 1
δ	0.25	0, $0.\overline{3}$, 0.5, 1

We ran the initial case values as enumerated in Table 2 for all five months of load boards and for all the five home cities. The resulting revenues and utilization (% loaded) achieved for October were much lower than the other four months since less than half as many loads were available in this month than in each of the other four. For a meaningful comparison, only the results of the four complete months from June to September are illustrated in Figure 12.

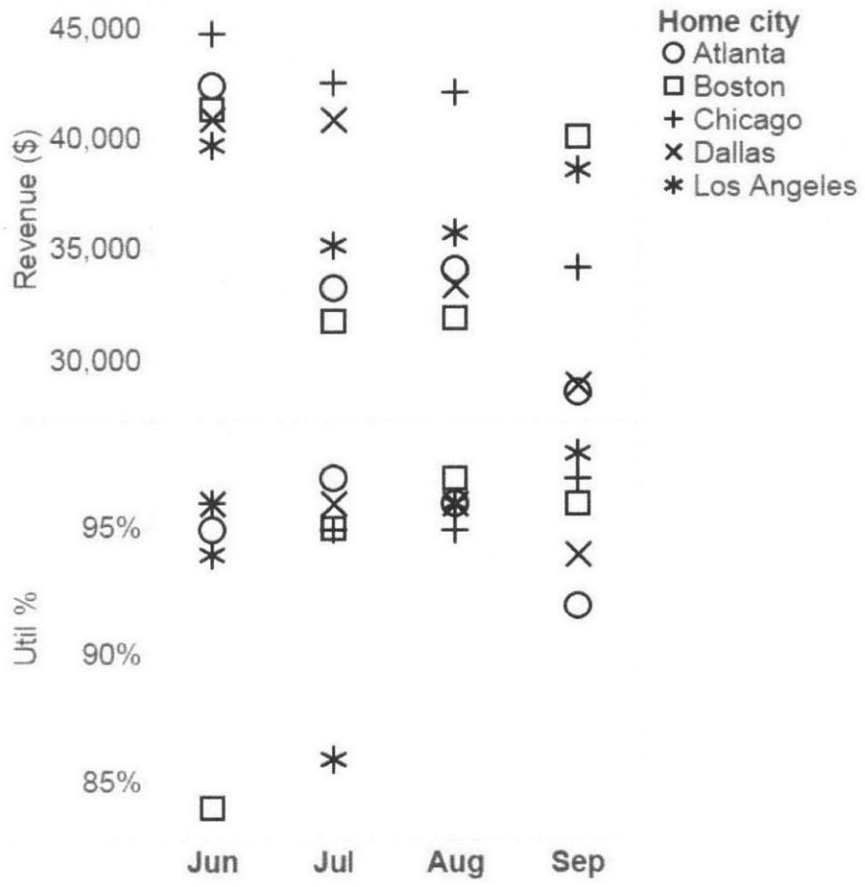


Figure 12 Comparison of revenue ($) and utilization (% loaded) for the five home cities across complete months.

The simulation runs originating from Chicago achieved the highest revenue in three of the four months shown. Aside from this, there were no consistent trends for the other cities. The variance in performance across cities supported further investigation of each geographic location.

For visual simplicity moving forward, we present the results of the exploratory simulation runs for the month of August alone, which, as seen in Figure 10, produced utilization with the narrowest variance and revenues with the second narrowest variance among the four complete months. However, any trends discussed reflect results across the complete time horizon.

The results of a short-haul- and a long haul-preferred strategy are shown in Figure 13. Recall that the utility function is initially engineered to score long-distance loads more highly; the alternative case, known as short-haul preference, favors short hauls. Pitting these strategies against each other illuminates whether short, local routing is inferior to a preference for longer hauls. The alternative case yielded markedly lower revenues and utilization than the initial case for the length of haul preference.

Figure 13 Comparison of revenue ($) and utilization (% loaded) for long-haul preference and short-haul preference for the August simulation run. In the long-haul condition, the longest haul in the dataset is given the highest ranking; the opposite is true for the short-haul strategy.

Extending the wait time, the time window when drivers could review available loads, from the initial case of one hour up to nine hours yielded the results shown in Figure 14. Contrary to our hypothesis that longer review periods would result in higher revenues earned and lower utilization, there was no clear trend in these two metrics with respect to time.

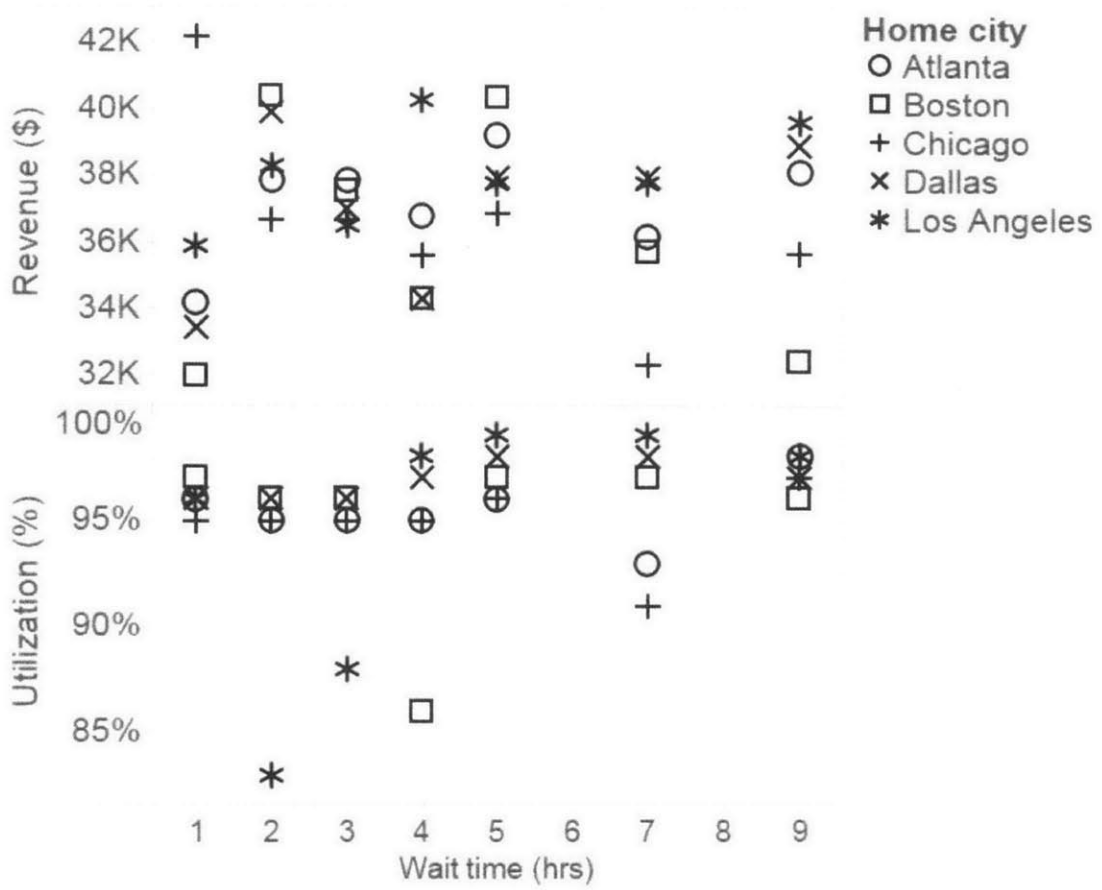
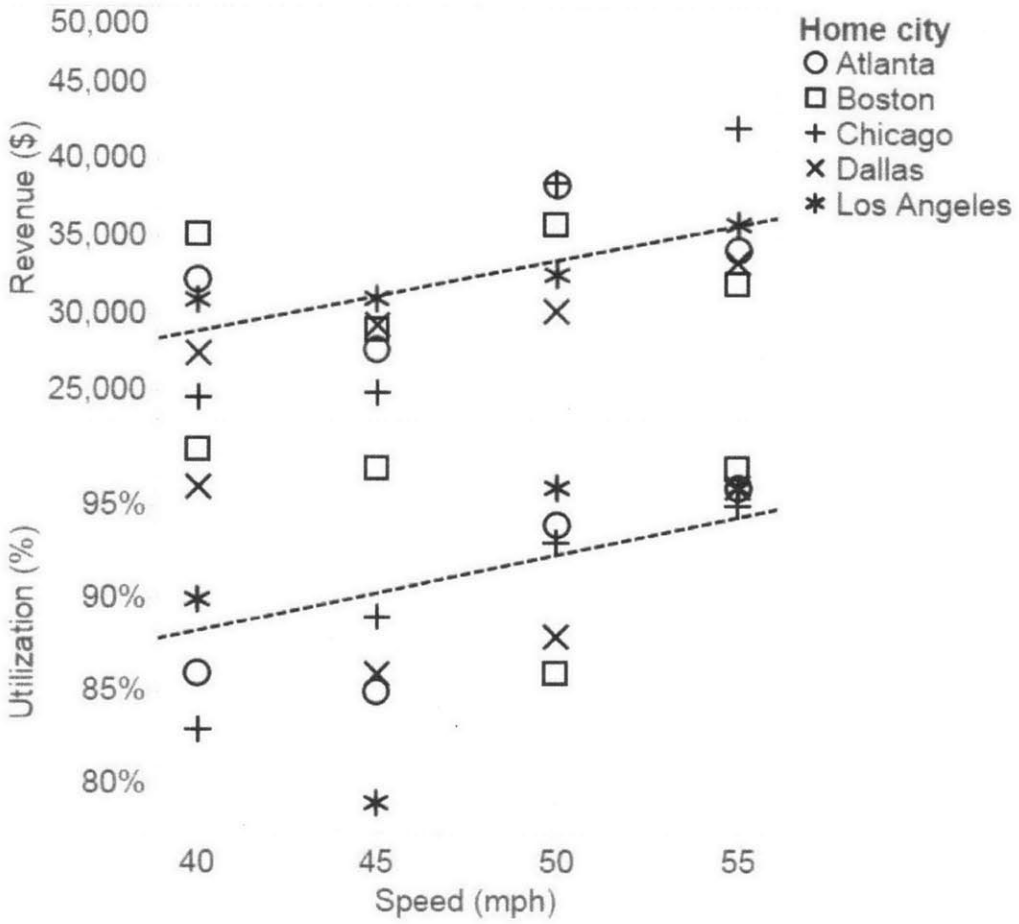


Figure 14 Comparison of revenue ($) and utilization (% loaded) across different simulation wait times (hrs) for the August simulation run.

Testing truck speeds from 40 mph to the initial case of 55 mph in intervals of 5 mph yielded the results shown in Figure 15. Both revenue and utilization trended upwards as truck speed increased.



Revenue Trend Line
Revenue = 459.53*(speed) + 10,475 p = 0.01

Utilization Trend Line
Utilization = 0.004*(speed) + 0.71 p = 0.07

Figure 15 Comparison of revenue ($) and utilization (% loaded) across difference simulated truck speed (miles per hour, mph) for the August simulation run, with linear trend lines.

Testing the deadhead ceiling values in the initial and alternative cases yielded results for August that exhibited an upward trend in revenues and a downward trend in utilization as the deadhead ceiling rose. However, these trends did not hold in other months, so we deemed the

August results to be non-indicative of a consistent relationship between deadhead ceiling and the chose not to show the August results here.

Finally, we looked at the impact of varying the weights of the load parameters in the utility function, α, β, γ and δ, from considering all variables equally to considering only one variable at a time. We found, in general, that strategies combining the four utility function variables, *L, R, P,* and *D,* yielded better results than strategies focusing exclusively on one of the four.

5.4 Strategy Design and Testing

Based on the insights from our data mining and initial simulation runs, we tested eight strategies in the final simulation iterations.

As month of data and home city did not generate results fitting a clear pattern, we tested all other variables using each city and month. Having determined that a long haul preference and truck speed of 55 mph yielded higher revenues and utilization above 90%, we fixed these across all strategies. Moreover, we fixed wait time at two hours since this was the shortest wait time at which most cities perform above average on revenue and utilization. We chose to set our deadhead ceiling, c, at a higher level than the industry average of 10% to account for the possibility that owner-operators do not manage deadhead miles as effectively as the leading companies (Kochar, Monahan, Peters, & Ward, 2012; TransCore, 2011a, p. 3). We decided to set c = 20% since increasing the threshold to 25% did not significantly improve the output metrics. Table 4 outlines parameters defining the strategies tested in the final simulation runs.

Table 3

Summary of Strategies Tested in Final Simulation Runs

Parameters \ Strategy	L	R	P	D	LD	LRD	LPD	LRPD
α	1	0	0	0	0.5	$0.\overline{3}$	$0.\overline{3}$	0.25
β	0	1	0	0	0	$0.\overline{3}$	0	0.25
γ	0	0	1	0	0	0	$0.\overline{3}$	0.25
δ	0	0	0	1	0.5	$0.\overline{3}$	$0.\overline{3}$	0.25
Haul Length Preference	Long							
Wait Time (hrs)	2							
Speed (mph)	55							
Deadhead Ceiling, c (%)	20							

Henceforth, strategies L, R, P, and D are referred to as *simple strategies* and LD, LRD, LPD, and LRD as *complex strategies*. The results of simulating these eight strategies across the five cities and five months are depicted in Figure 16 and summarized in Table 4. Total revenue reflects the sum of revenues over the four-and-a-half-month period, and utilization is the percent quotient of total loaded miles and total miles traveled.

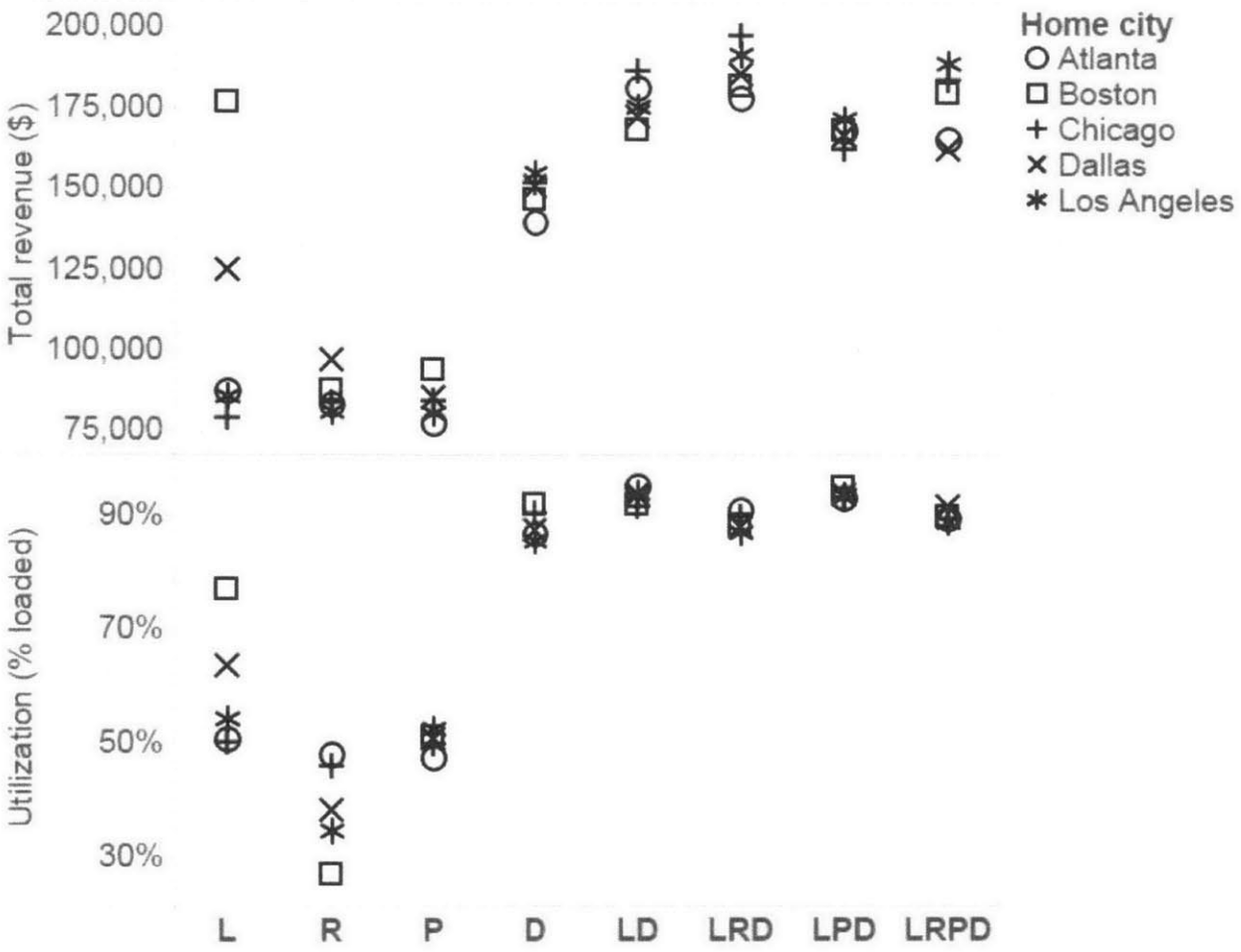


Figure 16 Total revenue ($) and utilization (% loaded) for each home city resulting from the simulated strategies over five months. Utilization is the sum of all loaded miles divided by the sum total of all miles driven.

Table 4

Heat Map Showing Total Revenue ($) and Utilization (% Loaded) for Each Home City Resulting from the Simulated Strategies Over Five Months.

		L	R	P	D	LD	LRD	LPD	LRPD
Total Revenue ($)	Atlanta	87,692	83,402	77,615	139,709	181,450	178,043	168,220	165,460
	Boston	176,819	88,132	94,132	146,726	168,525	182,363	168,255	179,913
	Chicago	79,329	84,034	84,944	151,960	186,729	198,256	162,535	184,545
	Dallas	125,100	97,611	85,907	151,002	172,672	186,108	167,083	162,693
	Los Angeles	86,148	81,264	80,887	155,093	175,649	191,923	172,018	189,170
	Mean	**111,018**	**86,889**	**84,697**	**148,898**	**177,005**	**187,339**	**167,622**	**176,356**
	StDev	**40,911**	**6,490**	**6,229**	**5,944**	**7,191**	**7,951**	**3,400**	**11,719**
	CoV	**0.37**	**0.07**	**0.07**	**0.04**	**0.04**	**0.04**	**0.02**	**0.07**
Utilization (% loaded)	Atlanta	51%	48%	47%	87%	95%	91%	93%	90%
	Boston	77%	27%	51%	92%	92%	89%	95%	90%
	Chicago	50%	46%	50%	91%	92%	90%	93%	90%
	Dallas	64%	38%	51%	88%	94%	88%	94%	92%
	Los Angeles	54%	35%	52%	86%	94%	87%	94%	89%
	Mean	**59%**	**39%**	**50%**	**89%**	**94%**	**89%**	**94%**	**90%**
	StDev	**11%**	**9%**	**2%**	**3%**	**1%**	**1%**	**1%**	**1%**
	CoV	**0.19**	**0.22**	**0.04**	**0.03**	**0.02**	**0.02**	**0.01**	**0.01**

Note:
The mean, standard deviation and coefficient of variation were calculated across the five cities for each strategy and metric.

The results show the impact of considering the four variables, *L, R, P,* and *D*, separately and together, on revenue earned and utilization achieved. The low coefficients of variation (<0.20) for 14 of the 16 groups of results suggested relatively robust outcomes across cities. In general, simple strategies yielded poor results, with the exception of an outlier total revenue value for Boston. Strategy L did not always maximize revenue, even though length of haul and revenue earned are correlated to each other. Strategy D yielded better results than the other three simple strategies. While complex strategies yielded almost universally better results, there was not a single best strategy for both output metrics. Applying strategy LRD earned the highest total revenue for four of the five home cities, whereas strategies LD and LPD achieved the highest utilization.

Carriers often structure their decisions with a length of haul and load RPM threshold (anonymous owner-operators and 3PL employees, personal communication, October 4, 2013). As a result, RPM results are presented as total revenue per total loaded miles, rather than total revenue per total miles. We are thus able to assess whether a strategy that focuses on load RPM is successful in terms of total utilization and revenues. We checked the average length of haul and loaded RPM that resulted from our strategies and summarize the results in Figure 17 and Table 5.

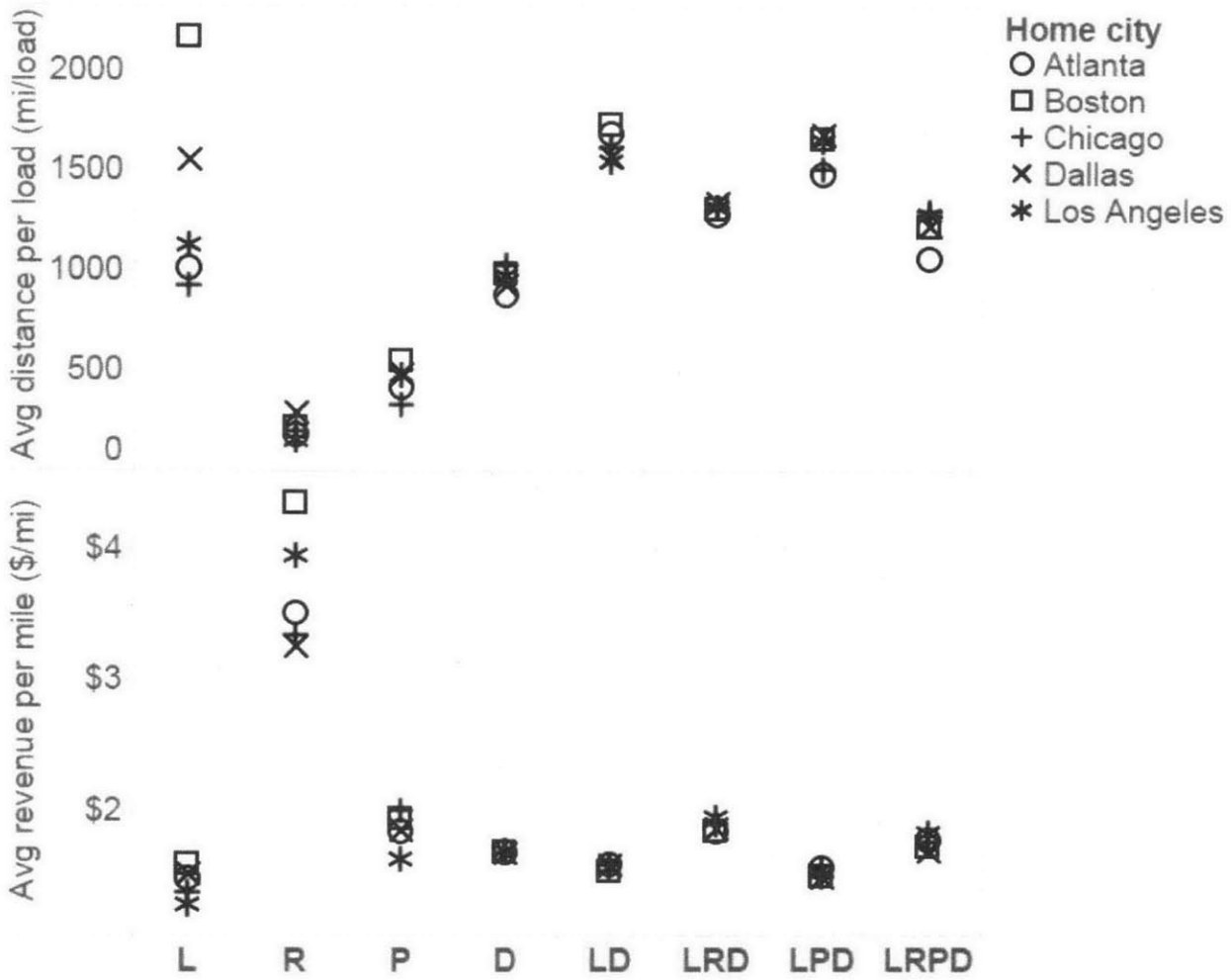


Figure 17 Average length of haul (mi/load) and average revenue per mile (RPM) ($/mile) for each home city resulting from the simulated strategies over five months.

Table 5

Heat Map Showing Length of Haul and Revenue Per Mile (RPM) for Each Home City Resulting from the Simulated Strategies Over Five Months.

		L	R	P	D	LD	LRD	LPD	LRPD
Avg length of haul (mi/load)	Atlanta	1,017	178	419	884	1,689	1,284	1,488	1,064
	Boston	2,171	217	549	988	1,734	1,315	1,661	1,223
	Chicago	927	196	333	1,033	1,619	1,320	1,514	1,305
	Dallas	1,552	293	483	930	1,573	1,342	1,683	1,229
	Los Angeles	1,127	156	482	980	1,548	1,324	1,651	1,290
	Mean	**1,359**	**208**	**453**	**963**	**1,633**	**1,317**	**1,599**	**1,222**
	StDev	**513**	**53**	**81**	**57**	**78**	**21**	**91**	**96**
	CoV	**0.38**	**0.25**	**0.18**	**0.06**	**0.05**	**0.02**	**0.06**	**0.08**
Avg revenue per mile ($/mi)	Atlanta	$1.49	$3.53	$1.85	$1.70	$1.60	$1.85	$1.57	$1.79
	Boston	$1.60	$4.36	$1.95	$1.69	$1.54	$1.85	$1.51	$1.73
	Chicago	$1.38	$3.35	$2.01	$1.69	$1.60	$1.90	$1.51	$1.77
	Dallas	$1.52	$3.26	$1.85	$1.69	$1.59	$1.87	$1.50	$1.70
	Los Angeles	$1.30	$3.96	$1.64	$1.70	$1.58	$1.96	$1.55	$1.83
	Mean	**$1.46**	**$3.69**	**$1.86**	**$1.69**	**$1.58**	**$1.89**	**$1.53**	**$1.76**
	StDev	**$0.12**	**$0.46**	**$0.14**	**$0.01**	**$0.02**	**$0.05**	**$0.03**	**$0.05**
	CoV	**0.08**	**0.13**	**0.07**	**0.00**	**0.02**	**0.02**	**0.02**	**0.03**

Note:
The mean, standard deviation and coefficient of variation were calculated across the five cities for each strategy and metric

The low coefficients of variation (<0.20) for 14 of the 16 groups of results again suggested relatively robust outcomes across cities. Interestingly, strategy L did not always maximize the length of haul, as would be expected by definition. Rather, strategies LD and LPD, which consider a combination of the variables, resulted in longer lengths of haul. The performance of the complex strategies showed a tradeoff in length of haul and RPM

maximization, with LD and LPD scoring higher on the former metric while LRD and LRPD scored higher on the latter metric.

In our interviews, carriers expressed a fixation on achieving a particular RPM (anonymous owner-operators and 3PL employees, personal communication, October 4, 2013). While strategy R yielded by far the highest RPM for all cities as expected, the low average lengths of haul corresponding to this strategy illustrate a tradeoff. Focusing exclusively on RPM produces a pattern of movement characterized by a large number of short-haul, local shipments. To achieve consistently higher lengths of haul and RPMs on average, the driver would do better to apply a more complex strategy in his decision-making.

Among the complex strategies, LRD yielded the highest average RPM of $1.89, and a corresponding average length of haul of 1,317 miles. To determine whether this strategy improved upon the current procedures of actual drivers, we compared this metric to the results achieved by comparable owner-operators in our dataset. We focused on owner-operators who hauled, on average, loads within 15% of 1,317 miles, and who carried at least the owner-operator mean of nine loads during the period. The latter restriction served to exclude those owner-operators that only took a small handful of loads, which skewed the data. It also served to isolate, for the purpose of this comparison, a group that worked at a rate more comparable with our simulated driver, who typically took over 80 loads during the same period. The average RPM of this actual owner-operator group was $1.63. Our best-performing strategy in terms of RPM, strategy LRD, therefore produced a statistically significant ($p = 0.00001$, Welch t-test) RPM increase of 16% over current actual heuristics.

5.5 Chapter Summary

Based on our findings from the data mining exercise, we isolated critical factors that differentiated carriers: origin or home location, current position, and the relative importance a carrier places on the length of haul, the revenue per mile anticipated, and the probability of an onward load. We varied these elements and tested these strategies through computerized simulations of a decision process to select among available loads under well-defined spatial and temporal parameters. The results of initial runs using our load boards and simulation model showed length of haul preference and truck speeds were especially impactful on revenue earned and utilization achieved. Applying these insights, we fixed parameters for our final simulation runs and tested the efficacy of eight strategies. We found that complex strategies considering multiple load characteristics in combination achieved superior results and that, in particular, minimizing deadhead distance was the single most important consideration.

6 Discussion

This section aims to clarify the context of, and provide possible explanations for, observed results. Like Chapters 4 and 5, it first covers data mining, and subsequently, the simulation.

6.1 Discussion of Results and Features of Data Profiling and Mining

The data was probed to understand key similarities and differences between it and the spot trucking market as a whole, and to acknowledge or control departures from industry trends.

6.1.1 Data consistencies and inconsistencies with spot trucking market. The chief objective in exploring the data was to answer whether the data was representative of the transportation market. One metric used to determine this was the percentage of owner-operator carriers relative to total carriers. On the one hand, we would expect owner-operators to represent approximately 70% of all carriers to be consistent with the literature. On the other hand, we expect that the third-party logistic (3PL) provider's data will be less skewed toward owner-operators due to the nature of their business operations and the carrier relationships they maintain. Therefore, the 58% majority position is in line with expectations, but also indicates that this dataset under-represents small carriers. Moreover, several of the largest carriers appeared to have multiple distinct carrier IDs in the dataset, which would have inflated the number of non-owner-operator entries. The linear pattern of revenues versus haul length was expected, though the logistic trend applied to revenues per mile (RPM) versus haul length had an R^2 value indicating only a weak fit. In examining the results, we see this is caused largely by the

tail of the logistic curve approaching zero as haul length increases. A possible explanation for this is that as haul length increases, a minimum RPM applies.

We found that owner-operators in the data set had a shorter average length of haul compared to the average length for the data set as a whole. This was a surprising result as it is a departure from the expected operating model of an owner-operator. This could be a reflection of owner-operators selecting loads beyond a certain mileage threshold or not being approached by the broker for these longer hauls. It was the goal of the simulation to determine whether the assumed operating model favoring long hauls was inferior to an alternative, or whether carriers were not implementing the favored strategy.

We also investigated the average number of loads leaving each state; we found that, as usual, Texas, California, and Georgia had the highest values for λ due to higher traffic originating in these states. Using these and previous findings, we determined that the dataset corresponded with market features except where a departure from the norm was expected.

6.1.2 Limitations of data mining analysis. We acknowledge certain limitations of our analysis despite the robust volume of the dataset. Although we tested the dataset for consistency with the general market, it was supplied by one corporate sponsor, and cannot conclusively be said to be representative of the American spot trucking market. Additionally, as the data covered only a subset of a year, we could not rule out seasonal patterns or annual anomalies. We were also unable to track the overall strategy of any one carrier, as these carriers did not exclusively carry loads brokered by the 3PL. Furthermore, there was no validation of carrier IDs, as we did not have access to the database where carrier profiles were stored; this could and likely did result in duplicate IDs identifying a single customer. Finally, we were unable to distinguish between

cases where the carrier placed a bid at a certain price versus a shipper having asked for a particular price. Without understanding this price dynamic, we could not analyze whether carrier- or shipper-set prices were more favorable to the carrier.

6.2 Discussion of Features and Results of the Simulation

Based on the data mining, we identified key variables to be investigated: month, home city, haul length preference, wait time, truck speed, and deadhead ceiling, c. Initial simulation runs focused on identifying patterns in the resulting success metrics for each variable individually. These insights were then contemplated in the design of strategies for comparison in the final simulation runs.

6.2.1 Discussion of exploratory simulation run results. We found that the performance of the simulated drivers varied from month to month, but the limited horizon of the data from which the load boards were constructed precluded any possibility of tracking broader seasonal patterns. The simulated drivers' performance also varied by home city; the fact that the driver from Chicago achieved the highest revenues in multiple months implied that this variance in performance was not entirely random, but rather, was driven by factors which had to do with the starting location or region. We suggest further investigation of these location effects by testing other home cities or months of data.

Some patterns were consistent across temporal and spatial boundaries. Better performance could be expected from a strategy that preferred long hauls. This result supported the heuristic used by owner-operators to choose a few long hauls as opposed to a large volume of local routes. Performance also improved with truck speed. This was expected; a faster truck

would deliver more loads, thus earning more and staying loaded for a greater portion of the time on the road in general. Thus, the analysis of haul length preference and speed provided generalizable insights into carrier performance.

6.2.2 Discussion of comparison of tested carrier strategies. We found that complex strategies, which considered deadhead in addition to one or more of the other features, were more likely to produce higher revenues and utilization than simple strategies that considered only one factor. Still, among the simple strategies, one stood out as being relatively more successful. Strategy D, which minimized deadhead alone, not only enabled the driver to achieve high utilization by definition, but also kept him in a revenue-earning state for more of the time. Therefore, placing significant weight on deadhead distance to load pickup locations is critical for operational success.

We found one outlier in our final testing: Boston, under strategy L, achieved significantly higher revenues than the other cities. In fact, the revenue figure achieved was higher than any other across the simple strategies. Upon examining the simulated drivers' actual tours, or routes taken, under strategy L, we discovered an interesting pattern: drivers from other cities tended to repeatedly take loads that brought them back to their home city, whereas the Boston driver tended to take cross-country loads that took him from coast to coast (Figure 18), thus earning substantially more during the same time period. One possible explanation for this is the simulation feature that prevents a driver from picking up a load that he cannot get to in time for the scheduled pick-up. This restriction may have prevented drivers originating from locations in the middle of the country, such as Chicago and Dallas, from reaching either coast in time to pick up the lucrative cross-country loads. However, the fact that Boston substantially outperformed

Angeles, also a coastal location, suggests that there were other location and regional effects in play. Again, we suggest further investigation of these effects.

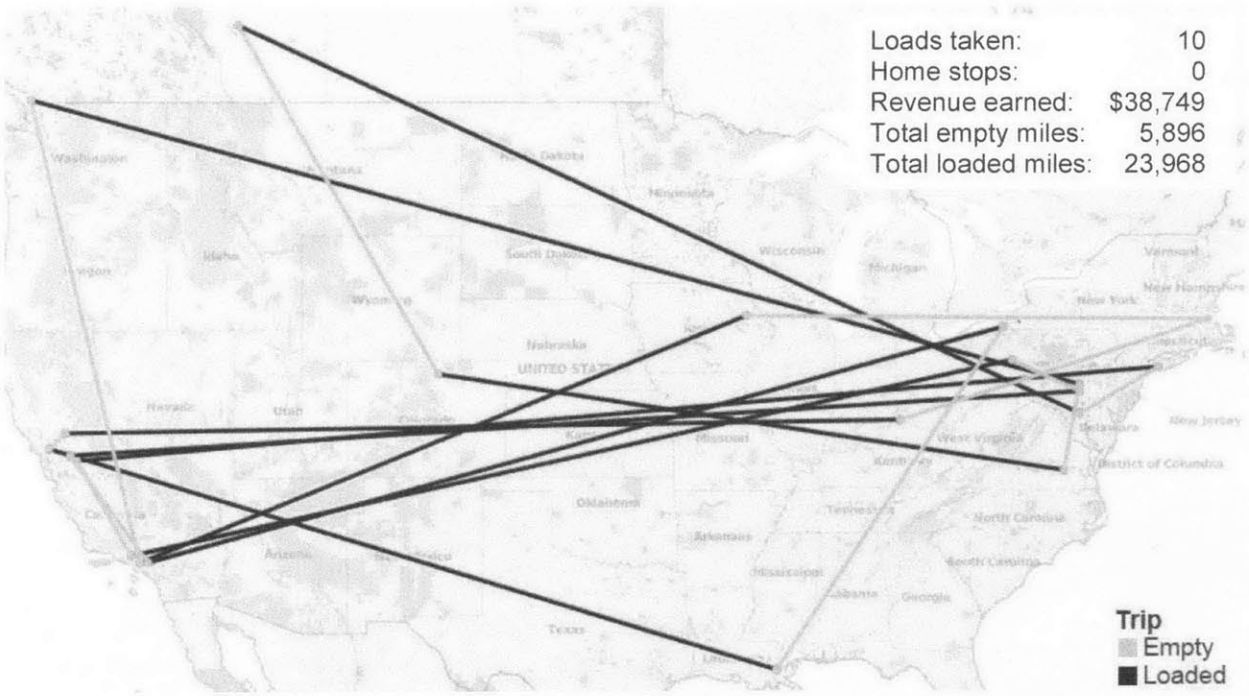


Figure 18 Tour of simulated driver who applies strategy L and originates from and returns to Boston during the month of August.

Overall, we infer that a driver adopting a fairly straightforward strategy of taking long haul loads and minimizing deadhead would already achieve a high degree of success. If earning the highest possible revenue were favored, then he should also consider the RPM of the load under consideration. If maximizing utilization were favored, then he should also consider the probability of an onward load from the destination state. There is an important tradeoff here, as the LRD strategy resulted in more revenues than the LPD strategy, but also a lower utilization

rate. Interestingly, for none of the five home cities was it necessary to consider the full set of variables to maximize either revenue or utilization.

Our other key finding relates to how owner-operators ought to define success; utilization as a success metric provided limited insights. In the strategies considering more than one variable, utilization was already at least 87%. It is interesting to note that carriers outperformed their deadhead ceiling, which was set at 20% for the final simulation runs, supporting to our earlier observation of diminishing returns on increasing deadhead threshold.

Using complex strategies, our simulated carriers achieved substantially longer haul lengths, on average, than actual carriers in the dataset. This was possibly because our simulation only weakly accounted for competition for loads. The simulated driver always had his pick of the absolute best load among all that the 3PL provider had to offer within each review window, and he had no awareness of or need to adjust to the possibility of other competing drivers. In reality, carriers do not always have perfect visibility on, and unobstructed access to, the market. Moreover, the simulation failed to account for a large array of factors that might force a carrier to take a shorter route or fewer jobs, such as restrictions to hours worked and time off the road, weather disturbances, traffic congestion, and the desire to avoid certain states, take breaks, or return home. Essentially, the simulation created drivers that operated robotically, failing to capture some of these more human factors.

6.2.3 Limitations of simulation model. We acknowledge that our simulation model was simplistic, and we encourage future extensions to simulate reality more closely.

Our model did not consider a number of factors that are influential in reality. One such factor was competition. We simulated desirable loads being offered less frequently, but future

iterations should consider other effects of competition on owner-operators' strategies. Another factor was profitability. Our results captured revenues, but did not consider the ramifications of individual carriers' very different operating costs. Personalized activity-based costing can be employed to overcome this limitation. Another key factor overlooked was the drivers' desire to return home at certain intervals, which may be crucial in many owner-operators' routing. This may be addressed by designing the utility function to increase pressure to return home as time on the road increases, or when certain dates of importance to the driver are approached. This design feature is similar to that employed by J.B. Hunt in its "Guaranteed Get Home Program" and its extension program "Operation North Pole" ("Logistics.com's Technology Enables J.B. Hunt's 6,700 Over the Road Drivers To Get Home for the Holidays.," 2000.)

Further, the model simplified the driver's response to the absence of satisfactory loads by sending him home. In reality, he could decide to wait longer, or reposition his truck and analyze loads available in another area. In our case, sending the driver home served as a proxy for Hours of Service (HOS) regulations that restrict consecutive working hours and days. We would recommend that future iterations of the model incorporate decision-making flexibility in situations when no qualifying loads are available, as well as explicit implementation of HOS restrictions (United States Department of Transportation, 2011).

Another weakness was the logic for calculating the probability of a backhaul, which used the likelihood that at least one load left the destination state. This not only failed to capture the more relevant point-to-point detail, but also arbitrarily prioritized political over economic boundaries. Additionally, it did not contemplate the likelihood of getting a backhaul from an adjacent state, or from a location en route to home. We recommend improvements on the region definitions that can accommodate a more granular regional breakdown, and uses a radius or

economic region rather than state boundaries. An effective means of achieving this would be the application of cluster analysis with geometric smoothing (Caplice, 2013a).

Finally, the option to exclude specific states, which was built into the model but not employed, was set up to check only loads that do not originate or terminate in the excluded states. It does not prohibit routes through certain excluded states, nor calculate additional time and distance for a less direct route. As previously mentioned, a restricted state condition was not used in our analysis, though further testing customized to an individual owner-operators' preferences and requirements could show the financial and operational impact of limiting a driver in this fashion.

6.3 Chapter Summary

Our research was designed to answer how carrier decisions affect their success. This study began with a validation of the dataset provided to us by a 3PL brokerage. In general, we found that the data was consistent with industry benchmarks. However, the operating model of the brokerage itself can explain certain departures from the norm, such as the underrepresentation of small carriers, and the haul length of these carriers.

Analysis of the data enabled us to hone in on important characteristics of carriers and loads to inform the design of our simulation. Specifically, we looked at month, home city, haul length preference, wait time, truck speed, and deadhead ceiling as hard-coded variables, and length of haul, RPM, probability of finding an onward load, and deadhead distance in a ranking scale. To this point, early simulation runs revealed that month, wait time, and deadhead ceiling were not particularly important variables, but home city, haul length preference, and truck speed were.

We used these insights to develop eight carrier strategies, and we found that complex strategies that considered a combination of two or more of the scoring parameters, length of haul, RPM, onward load probability, and deadhead distance, outperformed simple strategies that considered only one of the parameters. Additionally, deadhead distance was the single most important factor, and should be accounted for in strategic decisions.

However, both the data mining and simulation exercises have certain limitations. Future study on this topic should address these specifically.

7 Conclusion

With this line of research, we aimed to provide insights on owner-operator behavior and strategies in the spot market. Using data mining to define pertinent features of loads and carrier behavior, we deployed a spot market simulation that tested eight discrete carrier strategies. We suggest that a carrier's strategy may be broken down into decisions on four important characteristics of the available loads themselves: length of haul, revenue per mile, probability of an onward load from the destination, and deadhead distance to the load pickup location. These strategies dictated load choice in the simulation, and the simulation provided objective outputs by which to judge each strategy relative to others. Insights from this study can be used by carriers to improve operational success, as well as by third-party logistics providers to better match freight with carrier capacity. This section provides a summary of the main findings of this inquiry and introduces areas for further research.

7.1 Summary of Main Findings

Carrier strategies in the spot trucking market are dynamic and complex. The approach to choosing which, if any, among a set of available loads to haul at a given moment in time will necessarily vary depending on the carrier's personal situation and requirements as well as the conditions and opportunities available in the marketplace. Still, it is possible to formulate a general framework for such decisions that can meaningfully achieve better results and promote operational success for the carrier.

The best performing strategies were those that considered a combination of length of haul, RPM, probability of an onward load, and deadhead distance. While these characteristics may be evaluated independently of each other, a strategic-minded carrier should consider the

ways in which they interact when formulating the overall operational approach. The complexity of these interactions explain why, in the framework we have developed, there is no single strategy that is universally optimal for all carrier parameters and success metrics. We have seen, for instance, that focusing on RPM tends to be more important for increasing revenue earned whereas focusing instead on the probability of an onward load tends to be more important for maximizing utilization achieved. It should be noted, however, that deadhead distance is the single most important factor, and any strategy ought to consider this parameter. Also important to highlight is that strategies should consider the operating economies of the home location, as we found distinct results by metro area.

Another key finding is that a seemingly straightforward strategy does not always produce the intended results. For example, we found that a strategy focusing on the individual loads' length of haul alone does not produce the highest length of haul in most instances; rather a strategy considering this variable in conjunction with probability of an onward load and deadhead distance actually tends to produce a higher overall length of haul. Other such interaction dynamics should be taken into account when formulating more sophisticated strategies that are tailored to the carriers' unique situation and objectives.

Finally, in the absence of cost data to determine profitability, we found that carrier success ought to be focused on revenues rather than utilization. Although we found a modest tradeoff in revenue and utilization maximization, relatively high utilizations were seen in all cases combining more than one parameter.

7.2 Contributions and Areas for Future Research

By undertaking this research, this study takes a step towards bridging the existing gap in academic literature on owner-operator and carrier strategies in the spot market. We have illuminated the dynamics of the owner-operator's decision on load selection, a topic that has scarcely been covered so far in the literature. By developing the simulation model, we have also advanced a way of examining, testing and comparing general carrier strategies for any kind of sequential decision-making process in the freight market.

Using our load decision framework, we were able to define a strategy that achieved a higher RPM for our simulated driver than that of the comparable owner-operators in our dataset. We infer that there is a margin for improvement in owner-operator's current procedures, and that further research is warranted.

Future research may focus on further expanding the capacities of the simulation model and address its current limitations, which have already been enumerated in Section 5.2.3. To briefly summarize, the simulation logic fails to capture several restrictive market features, such as competition, time on the road and away from home, Hours of Service regulations, and a driver's unwillingness to enter certain states. The decision to define success by revenue earned and utilization, rather than profitability, was a result of the data available; however, future studies should incorporate costs in order to calculate profitability.

Alternatively, the methods of formulating the utility function to model strategies and testing these in a simulation may be applied to larger datasets or other classes of carriers. Finally, further research may adopt a broader framework on carrier strategies, comprising not just a sequence of load selections, but rather a whole array of operating choices, which extend over different time horizons and interact with each other in complex ways. Such research may

take the methods and insights developed by this study and build upon them to formulate an even more comprehensive, nuanced, and effective strategy for bidding on freight in the spot market.

The potential applications of this research extend beyond the academic realm. For the owner-operator, the insights on how different priorities in load selection affect different operational and financial performance may yield a more concrete and evidence-based understanding of how to make better decisions to achieve his objectives. For the freight brokerage company, the research adds new layers to the present understanding of how freight may be matched optimally with the type of carrier capacity provided by owner-operators. It may even help define the logic for an automation of this matching process.

References

Armstrong & Associates, Inc. (2014). *U.S. 3PL Market Size Estimates.* Retrieved from https://www.3plogistics.com/3PLmarket.htm.

Belzer, M. H. (1995). Collective Bargaining After Deregulation: Do the Teamsters Still Count? *Industrial & Labor Relations Review*, *48*(4), 636–655.

C. Barnes & Co. (2014). Barnes Reports: Freight Trucking Long Distance Industry (NAICS 48412). *United States Freight Trucking Long Distance Industry Capital & Expenses Report*, 1–218.

Caldwell, E. R., & Fisher, B. C. (2008). *The impact of lead time on truckload transportation rates* (Master's Thesis). Retrieved from http://hdl.handle.net/1721.1/45252

Cantor, D. E., Celebi, H., Corsi, T. M., & Grimm, C. M. (2013). Do owner–operators pose a safety risk on the nation's highways? *Transportation Research Part E: Logistics and Transportation Review*, *59*, 34–47. doi:10.1016/j.tre.2013.08.002

Caplice, C. G. (1996). *An optimization based bidding process : a new framework for shipper-carrier relationships* (Doctoral Dissertation). Massachusetts Institute of Technology. Retrieved from http://dspace.mit.edu/handle/1721.1/10846

Caplice, C. (2007). Electronic Markets for Truckload Transportation. *Production & Operations Management*, *16*(4), 423–436.

Caplice, C. (2013a). *Carrier Fundamentals I: Linehaul Operations.* Personal Collection of C. Caplice, Massachusetts Institute of Technology, Cambridge, MA.

Caplice, C. (2013b). *Transportation Analysis: Local Routing using Approximation Methods.* Personal Collection of C. Caplice, Massachusetts Institute of Technology, Cambridge, MA.

Caplice, C., & Sheffi, Y. (2003). Optimization-Based Procurement for Transportation Services. *Journal of Business Logistics*, *24*(2), 109–128. doi:10.1002/j.2158-1592.2003.tb00048.x

Cherry, C. R., & Adelakun, A. A. (2012). Truck driver perceptions and preferences: Congestion and conflict, managed lanes, and tolls. *Transport Policy*, *24*, 1–9. doi:10.1016/j.tranpol.2012.07.012

Costello, B. (2013). The Trucking Industry: The Lynchpin of the U.S. Economy. *Business Economics*, *48*(3), 195–201. doi:http://dx.doi.org.libproxy.mit.edu/10.1057/be.2013.16

Garrido, R. A. (2007). Procurement of transportation services in spot markets under a double-auction scheme with elastic demand. *Transportation Research Part B: Methodological*, *41*(9), 1067–1078. doi:10.1016/j.trb.2007.04.001

Goddard, L. (2014, January). Freight Forwarding Brokerages & Agencies in the U.S. *IBIS World Industry Report.*

Goes, P. B., Karuga, G. G., & Tripathi, A. K. (2012). Bidding Behavior Evolution in Sequential Auctions: Characterization and Analysis. *MIS Quarterly*, *36*(4), 1021–1042.

Kochar, A., Monahan, S., Peters, J., Ward, J. (2012, July 12). Multi-Modal Management: It's a Group Effort. *Supply & Demand Chain Executive*. Retrieved April 16, 2014, from http://www.sdcexec.com/article/10741605/multi-modal-management-its-a-group-effort

Kafarski, L., & Caruso, D. A. (2012). *Effects of truckload freight assignment methods on carrier capacity and pricing* (Master's Thesis). Massachusetts Institute of Technology. Retrieved from http://dspace.mit.edu/handle/1721.1/77464

Kirkeby, K. (2014, February). Transportation: Commercial. *Standard & Poor's Industry Surveys.* Retrieved from Standard & Poor's NetAdvantage database.

Kvidera, M. (2009). Extended Stay. *Overdrive*, *49*(7), 22–24.

Ledyard, J. O., Olson, M., Porter, D., Swanson, J. A., & Torma, D. P. (2000). The first use of a combined value auction for transportation services. Social Science Working Paper 1093, Division of the Humanities and Social Sciences, California Institute of Technology. Retrieved from http://www.hss.caltech.edu/SSPapers/wp1093.pdf

Ledyard, J. O., Olson, M., Porter, D., Swanson, J. A., & Torma, D. P. (2002). The First Use of a Combined-Value Auction for Transportation Services. *Interfaces*, *32*(5), 4–12.

Logistics.com's Technology Enables J.B. Hunt's 6,700 Over the Road Drivers To Get Home for the Holidays. (2000, December 22). Retrieved from http://www.thefreelibrary.com/Logistics.com%27s%20Technology%20Enables%20J.B.%20Hunt%27s%206,700%20Over%20the%20Road...-a068360009

Marcus, A. A. (1987). From Market Dominance to Credible Commitments: Shippers9 Strategies in a Deregulated Trucking Environment. *Transportation Journal (American Society of Transportation & Logistics Inc)*, *26*(4), 4–15.

Mergent, Inc. (2014). *Truckload Carriers.* Retrieved November 23, 2013, from Mergent Online database.

Moore Jr., E. W., Warmke, J. M., & Gorban, L. R. (1991). The Indispensable Role of Management Science in Centralizing Freight Operations at Reynolds Metals Company. *Interfaces*, *21*(1), 107–129.

Neumann, J.v. & Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Nickerson, J. A., & Silverman, B. S. (2003). Why Aren't All Truck Drivers Owner-Operators? Asset Ownership and the Employment Relation in Interstate for-Hire Trucking. *Journal of Economics & Management Strategy*, *12*(1), 91–118. doi:10.1162/105864003321220742

Peoples, J., & Peteraf, M. (1995). Deregulation and the Competitive Fringe: Owner-Operators in the Trucking Industry. *Journal of Regulatory Economics*, *7*(1), 27–42.

Powell, W. B., Sheffi, Y., Nickerson, K. S., Butterbaugh, K., & Atherton, S. (1988). Maximizing Profits for North American Van Lines' Truckload Division: A New Framework for Pricing and Operations. *Interfaces*, *18*(1), 21–41.

Rivera, E. (2014, February). Long Distance Freight Trucking in the U.S. *IBIS World Industry Report.*

Stephenson, F. J., & Stank, T. P. (1994). Truckload Motor Carrier Profitability Strategies. *Transportation Journal (American Society of Transportation & Logistics Inc)*, *34*(2), 5–17.

TransCore (2011a). *Carrier Benchmark Survey: TransCore 2011.* Auburn, MA.

TransCore (2011b). *The Spot Freight Market: A Primer for Shippers.* Auburn, MA.

Tsai, M.-T., Saphores, J.-D., & Regan, A. (2011). Valuation of freight transportation contracts under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, *47*(6), 920–932. doi:10.1016/j.tre.2011.03.005

United States Department of Transportation. Federal Motor Carrier Safety Administration. (2011, December 27). *Hours of Service of Drivers: Final Rule.* Washington DC: Government Printing Office. (Federal Register 76(248)). Retrieved 27 February 2014 from the Federal Motor Carrier Safety Administration website: http://www.gpo.gov/fdsys/pkg/FR-2011-12-27/pdf/2011-32696.pdf

United States Department of Transportation. Federal Highway Administration. (2014, February). *Traffic Volume Trends.* Retrieved 4 March 2014 from the Federal Highway Administration website: http://www.fhwa.dot.gov/policyinformation/travel_monitoring/14febtvt/

Wyckoff, D. D., & Maister, D. H. (1975). *The Owner-Operator: Independent Trucker.* Lexington, MA: Lexington Books: D.C. Heath and Company.

Zeithammer, R. (2003). *Forward-looking bidders in sequential auctions* (Thesis). Massachusetts Institute of Technology. Retrieved from http://dspace.mit.edu/handle/1721.1/29645