Traditional Routing Guide Performance and Segmentation to Improve Compliance with Contracted Budget

by

Tala Muneer Alnajdawi
B.S. in Supply Chain Management
Arizona State University – W.P. Carey School of Business

and

Israel Lopez Jimenez

MSc in Mechanical Engineering
Lawrence Technological University

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE OR MASTER OF ENGINEERING IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2020

Signature of Author: ____
Department of Supply Chain Management
May 8, 2020

Signature of Author: ____
Department of Supply Chain Management
May 8, 2020

Certified by: ____
Dr. Chris Caplice
Executive Director, Center for Transportation and Logistics
Capstone Advisor

Certified by: ____
Angela Acocella
PhD Candidate, Research Assistant
Capstone Co-Advisor

Accepted by: ____
Prof. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

Traditional Routing Guide Performance and Segmentation to Improve Compliance with Contracted Budget

by

Tala Alnajdawi

and

Israel Lopez Jimenez

Submitted to the Program in Supply Chain Management
on May 8, 2020 in Partial Fulfillment of the
Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT:

A strong seller's market in 2017 and 2018 led to dramatically increased costs in transportation due to demand surpassing supply, government regulations, and a shortage of truck drivers. As a result, the carrier rejection rate by primary carriers in the routing guide increased. This research examines the performance of routing guides to segment freight to help shippers identify and characterize where and how budget overruns occur. Using data characterization and regression modeling, we examine the plan data (carrier/lane/volume) and analyze how the transactions performed against it. We analyze one year's worth of shipper data from March 2019, when the plan was made, to March 2020 for three shipper sizes. We classify how lanes perform relative to the planned budget to determine the underlying factors that contribute to budget overruns by creating a freight categorization framework. A linear regression model was built to quantify the impact of independent variables such as distance, lane volume, origin/destination, and lane freight types (dry/refrigerated/frozen) on spend, volume, and total cost contribution to deviations from planned budget. The research found that frozen lane freight loads contribute to higher budget deviations, while dry van loads tend to contribute to lower budget deviations. Furthermore, specific origins and destinations impact budget deviations depending on the shipper. While volume deviations contribute to budget overruns more than price deviations. Finally, we provide insights to determine better segmentation strategies for procurement and management of transportation bids in the future.

Capstone Advisor: Dr. Chris Caplice
Title: Executive Director, Center for Transportation and Logistics
Capstone Co-Advisor: Angela Acocella
Title: PhD Candidate, Research Assistant

ACKNOWLEDGEMENTS

We would like express our very great appreciation to our advisor, Dr. Chris Caplice, for his valuable and constructive feedback during this capstone project. We are grateful to him for offering his time and expert knowledge throughout the year.

We are particularly appreciative to our co-advisor, Angela Acocella, for her extensive time and effort related to this capstone, especially on the methodology and data aspects. We would like to express our gratitude for her continuous encouragement and positive attitude.

Next, we would like to thank our sponsor company, especially Joe and Sara, for their valuable insights on the industry and analyzing the data of this project.

Finally, we would like to thank Pamela Siska for her comprehensive feedback during the writing of this capstone.

TABLE OF CONTENTS

LIST OF FIGURES:

LIST OF TABLES:

1. INTRODUCTION

In 2017 and 2018, the logistics industry faced a challenging freight environment driven by demand surpassing supply, U.S. government regulations that impacted driver efficiency, and other factors. The combination of newly-enforced Electronic Logging Device (ELD) regulations and a strong economy negatively affected routing guides and transportation budgets across truckload transportation, causing unprecedented volatility (Soskin, 2018). Companies have placed additional emphasis on logistics costs due to the impact their supply chains have on earnings.

This study identifies the factors influencing the costs associated with shipping and how carrier performance influences the bottom line of the logistics industry as a whole. The trucking industry in the U.S. bases its business competitiveness on the level of confidence with which it assigns primary carriers to specific lanes. This research seeks to improve the process by which shippers procure lane contracts with carriers and to identify the gaps in the current state of truckload procurement. Although our research utilizes truckload data from the sponsor company, the findings of this research are applicable to the truckload industry as a whole.

1.1 Problem Statement

The challenge the transportation industry, including our sponsor company, faces is that while transportation costs are negotiated up front with the shipper on the basis of the primary carrier's lane pricing, the volume's capacity commitments are non-binding. Since competitive pricing is generally negotiated with the primary carrier, that leaves companies responsible for absorbing additional costs in transportation if the primary carrier acceptance rate is subpar. In general, truckload contractual agreements do not force a shipper to honor a volume commitment, or a carrier to honor a capacity commitment. The nature of these contracts does not protect against carriers rejecting loads or abandoning a specific lane altogether if volumes do not materialize or if better opportunities arise elsewhere. Routing guide performance from the previous year should be scrutinized to identify patterns, correlations, and (if possible) root causes of such deviations while quantifying the economic impacts to the company's bottom line.

This capstone project analyzes the performance of our sponsor company's 2019 routing guides to determine their advantages and disadvantages. Initially, we identify and examine the underlying factors and develop strategies and recommendations on how to improve overall performance. The methodologies used for this capstone examine the plan data (carrier/lane/volume) and analyze how the transaction data performed against it. Data characterization methods include segmenting lanes that are included in the annual bid (planned) versus those not included (unplanned). We classify the independent lane attributes including origin/destination, lane freight types, distance, and lane volume that result in a deviation from the shipper's budget. Finally, we develop a segmentation framework that could help the sponsor company determine better segmentation strategies for procurement and management. This capstone project utilized substantial data analysis and visualization to capture and document current performance.

The capstone project requires considerable understanding of database analysis frameworks, BI visualization development, and machine learning techniques. This paper uses these methods to improve the process by which shippers procure lane contracts with carriers and to identify the gaps in the current state of truckload procurement. In general, failures in the routing guide and unplanned freight contribute to budget overruns but our research proposes strategies that can mitigate the costly discrepancies. The remainder of this report is organized as follows: Chapter 2 presents a literature review of the problem and current research conducted in this area. Chapter 3 discusses the methodologies used including data collection, cleaning, characterization, freight segmentation, and regression models. Chapter 4 describes the regression models and their results. Finally, Chapter 5 provides a summary and offers management insights and recommendations for future research.

2. LITERATURE REVIEW

2.1 Truckload Transportation Market

The U.S. trucking market is very large. The American Trucking Association (ATA) highlights that the trucking market was valued at $767 billion from gross freight revenue in 2018. With 71% of all the freight tonnage moved in the United States moving on trucks (ATA, 2018). Commercial full truckload transportation accounted for $296 billion in revenue, less-than-truckload generated $72 billion, and the remaining $300 billion was from private fleets. Truckload carriers mostly transport large shipments from point of origin to destination with no intermediate stops or handling. Truckload operations follow a physical process which involves a shipper, a carrier, and a receiver. Hundreds of thousands of manufacturers, wholesalers, and retailers make up the demand side of the truckload industry that serve their end markets. These shippers determine what to make and at what quantities. While the truckload transportation (carriers) market supplies the ability to meet the demand set by the shippers. The availability of the supply to meet this demand is what determines market pricing (Pickett, 2018).

2.2 Truckload Procurement

Since the 1980s, the dominant method for shippers to procure for-hire carriers across the transportation industry is to run annual reverse auctions to secure carrier capacity on all lanes (origin to destination route) within a network based on forecasted volumes over the next year. These rates and commitments are fed into a routing guide, an electronic catalog used by shippers. Routing guides contain the “rules of engagement” for shipping product from suppliers to customers (Ruriani, 2007). Carriers are sequenced in the routing guides by the most competitive bid prices, carrier capacity, and other criteria determined critical by the shipper. An alternative method to procure transportation is called the spot market. Spot markets are where a shipper tenders a single load to a carrier at a rate determined at the time of tender. The spot market method is usually undesirable for shippers that require regular load commitments on a lane as it is more volatile in price. Typically, but not always, spot market rates are higher than those of contracted rates (LaGore, 2019).

Several components make up transportation costs, to include line-haul costs, fuel surcharges, accessorial, and service fees. This research focuses on the line-haul charges for each lane (origin to destination route) because these are the rates that can be controlled versus fuel, which is subject to the market index pricing at a point in time. Line-haul costs are calculated using the fixed cost per mile (distance), processing time to load and unload the truck (dwell time), and the cost of the equipment (truck). Carriers usually quote the line-haul charge as the cost per load or per mile.

Our sponsor is an industry leading third-party logistics provider that manages the truckload procurement and logistics on behalf of large shippers in the food service/restaurant industry. They deliver food and supplies to the regional distribution centers for their customers. As shown in Figure 1, the truckload procurement process includes the five major tasks: carrier screening, information exchange, carrier assignment, load tendering, and performance review (Caplice, 1996). Aemireddy and Yuan (2019) summarized the following functions of each step in the process.

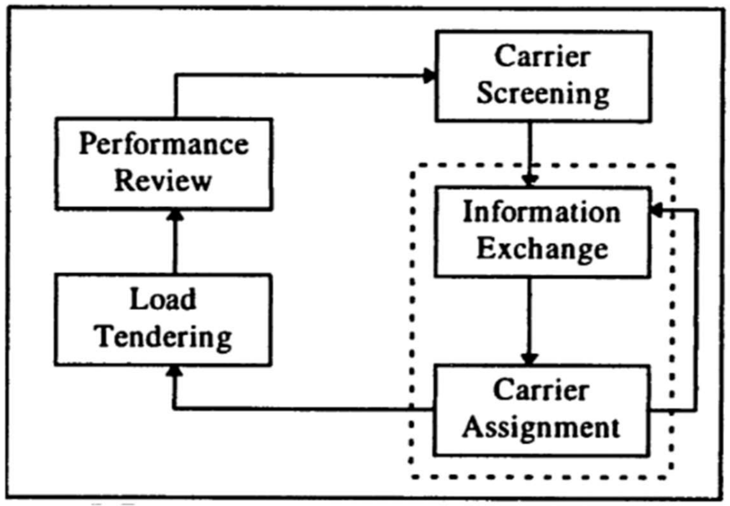


Figure 1: Five-Step Truckload Procurement Process. Source: Caplice (1996)

1. Carrier Screening - identifies and reduces the number of potential carriers. This process ensures that candidates meet the minimum capacity requirements for the shipper.
2. Information Exchange - Potential carriers participate in an annual reverse auction to submit quote/bid on each lane.
3. Carrier Assignment - The shipper selects the carriers with the most competitive bids and award capacity on each lane to primary carriers. The primary carrier(s) costs are used to budget for transportation costs for the following year. The contracted bids are loaded in an electronic catalog called "routing guides" or "transportation management system (TMS)" (Caplice, 2007). In addition, alternative carrier bids are also logged into the routing guide in the case that the primary carrier rejects the load/tender shipment. However, these "waterfall" carriers generally do not have as competitive pricing as the primary.
4. Load Tendering - a real-time process of assigning shipments (loads) to the primary carrier taking place in the routing guide. However, the challenge is that contracts are non-

binding, meaning the carrier has no contractual obligation to accept the load assignment. This requires the shipper to use alternative carriers to move the shipment on time if the primary refuses.

5. Performance Review - carrier performance based on percentage of tender acceptance (how often shipment is picked up) and on-time delivery is how the shipper determine where to allocate next year's capacity to. Although carriers are not contractually obligated to accept a shipment/load, the key performance indicators of tender acceptance will impact future business.

Although contracts are non-binding for carriers in terms of tender acceptance and volume commitment, the costs negotiated and loaded into the routing guide is fixed until the following year’s reverse auction. That means carriers cannot increase the cost per load on that lane for the duration of that year’s bid. Furthermore, shippers are not obligated to meet the forecasted volumes, in that shippers are not directly penalized if volumes are lower than expected by the carrier. The only firm aspect of the contract between shippers and carriers is the fixed line-haul costs. However, if performance is poor from either partner over the course of a year, they might restrict business in the future.

The negotiated and agreed upon carriers and rates are assigned and loaded into the routing guide. The routing guide automatically tenders pending loads to the primary carrier that will either be accepted or rejected. If rejected by the primary carrier, the load gets tendered to the subsequent carrier. Aemireddy and Yuan (2019) illustrated the automated process of load tendering as shown in Figure 2. The tender escalation process is usually conducted using a TMS. Our sponsor company usually completes the annual competitive bid process for their network in the first quarter of each year. Once the contracted bids are finalized, and analyzed, the primary carriers selected and awarded a contract for specific lanes. These rates are then loaded into the routing guide of the TMS. Once a load is dropped into the routing guide, the primary carrier is expected to accept all loads tendered. However, if the primary carrier rejects that specific load, it will be automatically routed to the backup carrier in the routing guide. This process continues through the routing guide until one of the backup carriers accepts that load (Caplice, 2009). Furthermore, if the load has not been accepted by a carrier in the routing guide, the shipper than sends it to the spot market. Our sponsor company follows a similar procurement process illustrated in Figure 2.

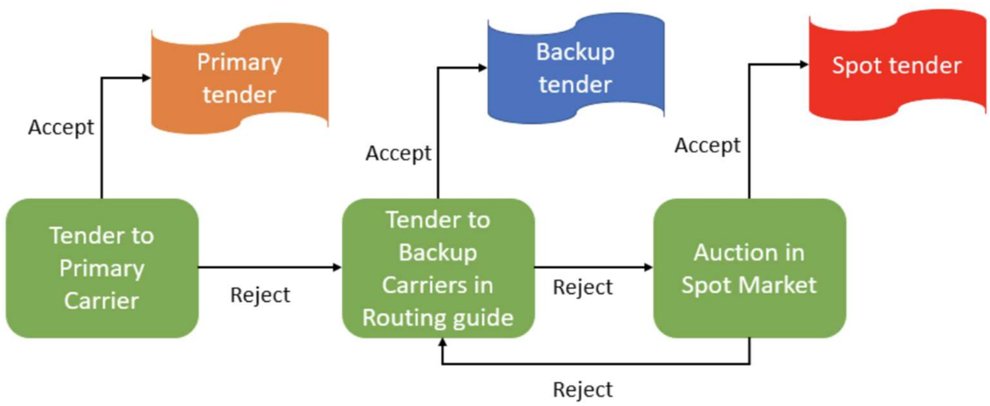


Figure 2: Tender Escalation Process. Source: Aemireddy and Yuan (2019)

2.3 Carrier Performance

Substantial research has examined the drivers of carrier acceptance level based on factors such as price, geographical patterns, and length of haul. Caplice and Sheffi (2003) note that economies of scale do not typically apply in the truckload industry, i.e. allocating more volume to a specific carrier does not always result in lower prices. Research conducted by Yoo Joon Kim (2013) suggested that weekly volume volatility is correlated with tender rejection rates. In 2017, the effects of increased shipper demand while carrier capacity was limited fostered an environment that incentivized primary carriers to reject the loads off the routing guide. Soskin (2018) suggested that “carriers will opportunistically reject a portion of their primary freight to free up capacity to take advantage of an inflated spot market.” This led the industry into a “vicious” cycle where the increased demand generated for spot market capacity accelerated the increase in spot prices.

2.4 Conclusion

Several studies explore the relationship of price and performance in truckload transportation. However, none explored the performance of routing guides, and specifically to understand where they succeed and where they fall short. This paper examines the underlying factors and develops strategies and recommendations for how to improve overall performance. Particularly, this capstone will identify whether carrier rejections can be mitigated in the future based on analyzing actual transaction data provided by the sponsor company.

3. METHODOLOGY

In this chapter, we examine one full year (March 2019–March 2020) of truckload shipment data, which includes master and transaction data files, provided by our sponsor company.

The data set includes the following information:

1. Master Award File: contains the information for the awarded primary carrier, negotiated lane pricing, origin to destination route (lane ID), and bid dates necessary to evaluate the annual budget for 2019.
2. Transaction File: includes the actual rates incurred per truckload during 2019.

3.1 Scope

The research has limited the scope to evaluating four shippers in terms of spend and volume. The shipper data is characterized by annual volume and represented as "large", "medium", and "small". Figure 3 breaks down the percentage of volume by shipper segment (large, medium, and small). There are two small category shippers.

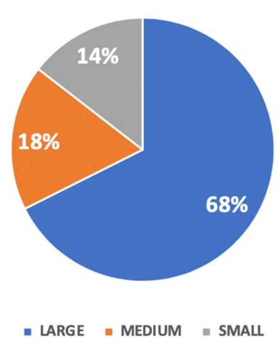


Figure 3: Percent Volume (loads) per Shipper

This capstone evaluates the costs of four shippers, the transaction data that lists the actual shipments made in the specified time period shows that the largest shipper makes up more than 65% of our sponsor company's spend. Figure 4 shows the degree of difference between the large shipper and the medium and small shippers in terms of volume (load transactions) per week. This graph illustrates that the largest shipper dominates the total number of loads executed overall, while the rest of the shippers are in a distant second and third place. We examine large and medium sized shippers to understand the differences in how their routing guides perform.

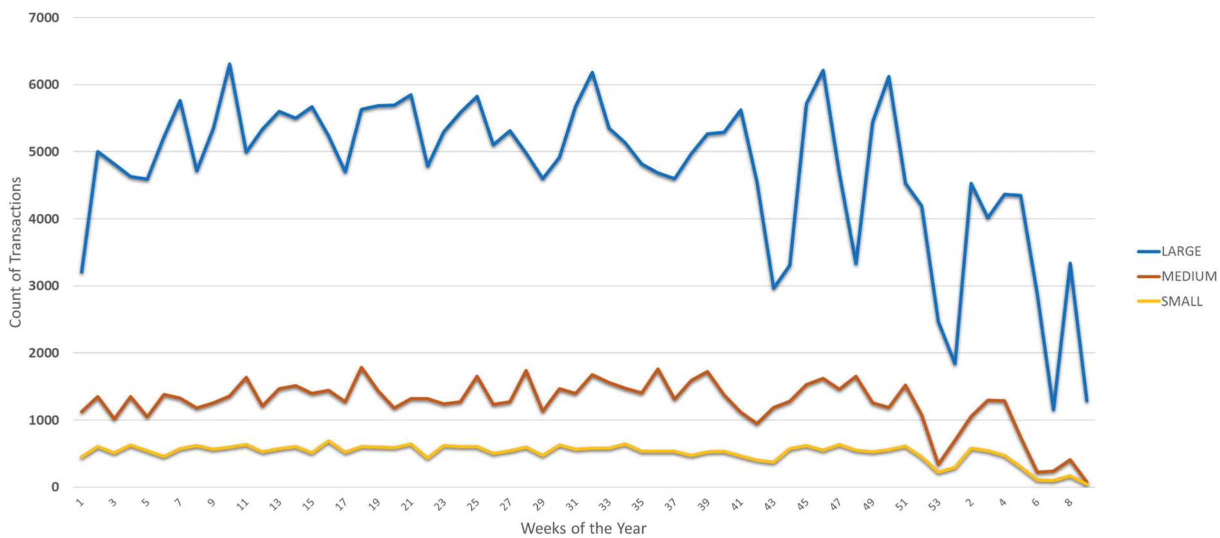


Figure 4: Total Number of Loads per Shipper per Week

3.2 Data Collection

The master award data consists of all the attributes related to the shipper, origin, destination, bid date, distance of lane, award type (primary or waterfall carriers), carrier name, quoted quantity (volume per lane), actual rate (line-haul cost), lane ID, and transit lead time. The load transaction data provides actual line-haul rates incurred per lane, shipper, pick up dates, drop off dates, carrier selected, lane ID (origin to destination route). The sponsor company also provides the truck type, number of pallets per truckload, and other tertiary pricing that will not be used in this capstone. Figure 5 illustrates how the master award and load transaction data are built and how this data architecture can be joined by lane ID to characterize the data and generate freight segmentation categories.

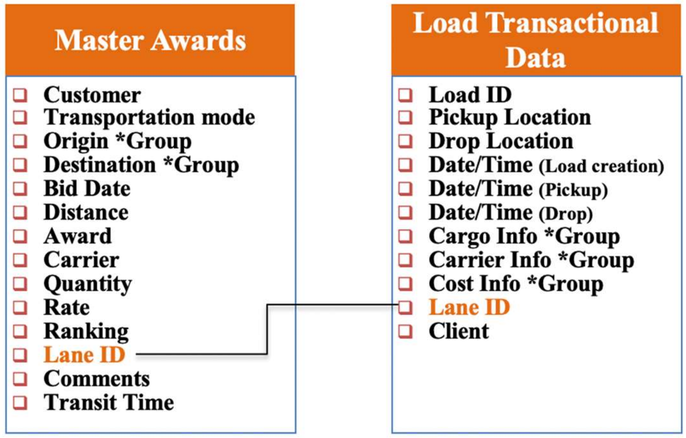


Figure 5: Database Architecture

Prior to analysis, the master award file was combined with the transaction file by the unique lane ID and primary carrier. Filtered only for full truckload shipments, and loads that are missing a line-haul rate entry from the master data. Additionally, we filtered out loads that have missing or erroneous data in origin and destination fields for routes within the U.S. We removed a total of 1,453 lanes. Finally, we evaluate if there is missing information: actual truckloads that do not have a master award cost or master award inputs that do not have actual shipments.

3.3 Lane Characterization

In order to analyze the routing guide performance, we examined each lane over the course of a year. Based on how the actual price and volume compared to the planned levels, we categorized each lane into one of five categories described in Table 1.

Table 1 Freight Categories. Source: Angela Acocella (2019)

Freight Category	Lane	Volume	Price
(1) Planned, On Budget	Bid	= Committed	= Contract
(2) Planned, Over Budget	Bid	> Committed Committed > Committed	Contract > Contract > Contract
(3) Planned, Under Budget	Bid	Committed	< Contract

		< Committed < Committed	Contract < Contract
(4) Unplanned (Over Budget)	Not Bid	> 0	> 0
(5) Ghost (Planned, Under Budget)	Bid	0	> 0

3.4 Freight Category Segmentation

The methodology focuses on segmenting the freight transaction data into five categories according to where they fall in terms of contract type (bid/not bid), volume, and contract price. This approach will determine the distribution of loads that fall into each category and narrow the scope of the project to areas showing routing guide degradation. Table 1 shows the variables for each segment category. The lane column shown in Table 1 reflects “bid” if the load was on a lane that was bid in the award file, meaning it had a contract price and volume. The volume column reflects whether the load met the committed volume per year or whether it was below or above what was in the award file. The price column reflects whether the contract price awarded to the primary carrier met the “contract” price or whether it was below or above. The price column indicates “contract” if the actual cost is within +/- 2% threshold of the contracted price.

Freight Category Definitions:

1. Planned, On Budget – lanes in this category represent the ideal scenario where it was executed according to the contract price (+/- 2%) and meets volume commitments exactly.
2. Planned, Over Budget –lanes that were awarded but were over the expected volume and/or contract price awarded in the routing guide.
3. Planned, Under Budget – lanes that were awarded but were under the expected volume and/or under the contract price awarded in routing guide.

4. Unplanned (Over Budget) – lanes that were never awarded or "not bid," so there is no committed volume or contract price to compare to.
5. Ghost (Planned, Under Budget) – lanes that were in the annual bid but no volume materialized.

The process of segmenting the freight into these distinct categories allows us to generate our hypothesis for data modeling which will be tested using a regression model. Figure 6 illustrates the flow chart determining the lane categories. Note it is difficult to capture ghost lane impact on budget deviations as these are lanes that were bid but volume did not materialize. Additionally, ghost lane volume could have been eliminated if the lane was no longer supported by the shipper. Our data analysis did not result in any ghost lane volume so we will focus on four categories.

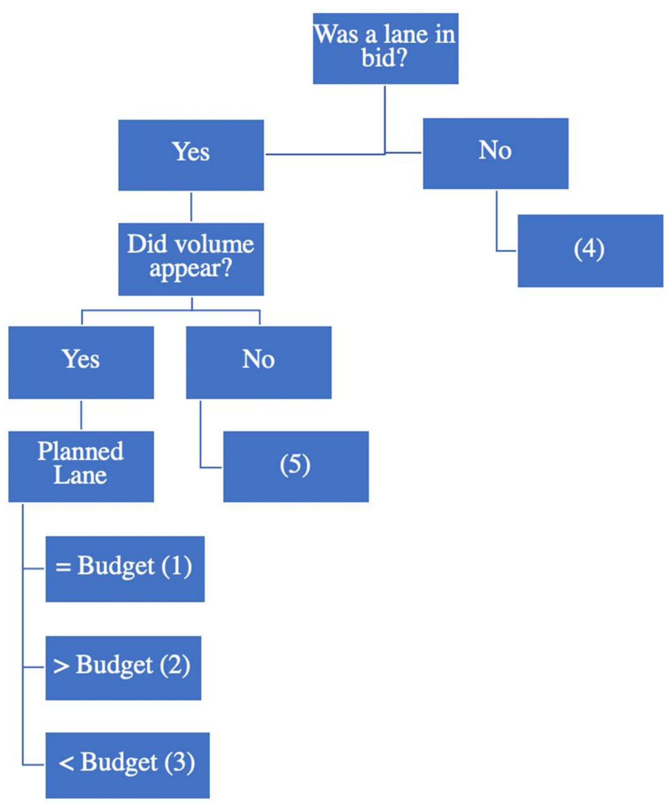


Figure 6: Freight Category Flow Chart

These freight categories allow us to visualize how a company's routing guide performs over time. For example, Figure 7 illustrates a hypothetical routing guide where the total tendered volume in each freight category across a shipper's network is expected to change over the contract lifetime, which will be one-year increments for this analysis. Figure 7 is an unpublished illustration of a hypothesis for the tendered volume deviations created by Angela Acocella.

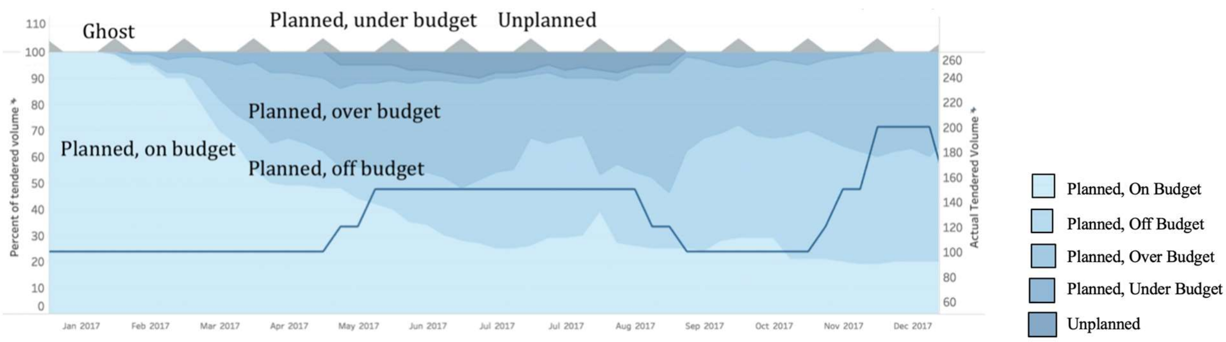


Figure 7: Tendered Volume Deviations (illustration). Source: Angela Acocella (2019)

3.5 Data Modeling

Upon completion of the initial data characterization and freight segmentation of the load transaction data, we used statistical Ordinary Least Square (OLS) multiple linear regression to identify and quantity the impact of input variables on the percent deviation from the shipper's budget per lane. The response variable (Y) is the result of how the dependent variable changes with independent variables (X). The method estimates the relationship by minimizing the sum of the squared errors, or the difference between the observed and predicted values of the dependent variable configured as a straight line (Shmueli, Patel, Bruce, & Torgo, 2017). In this case, OLS multiple linear regression will be discussed in the context of a model in which there are multiple independent variables (X) predicting a dependent variable (Y). The equation below illustrates how the model is used to fit a relationship between a numeric outcome variable (Y) and a set of predictors (X1, X2, .. Xn):

$$Y = \beta 0 + \beta 1 * X1 + \beta 2 * X2 + \ldots + \beta n * Xn + \epsilon,$$

where $\beta 0$ is the intercept, βi is the coefficient of each independent variable and ϵ is the error term, or the variability in the data that is not explained by the model. The quality of the model is measured by the R-squared. The adjusted R-squared is modified for the number of predictors (input variables) in the model. In addition to capturing which independent variables can affect deviations from a shipper's budget, this approach will also determine how future bid processes can take these attributes into account. Sections 3.6–3.9 provide an overview of the four models we test.

The four main characteristics of a lane that contribute to changes in expenditure are shown in Table 2. These variables are distance, lane volume, origin/destination, and lane freight type. The distance variable measures the number of miles traveled per load on a specific lane. The origin/destination aggregated at the state level uses binary variables to represent each origin and each destination, allowing us to quantify budget deviations for any origin to destination combination. Each corresponding origin and destination binary variable equals "1." The lane freight type refers to the type of truck dedicated for that specific lane.

Lane volume is segmented into high, medium, and low volume lanes based on how our sponsor company characterizes lane operations. The criteria for lane volume is described in Table 2. Lane volume is segmented into ABC categories because our sponsor company places emphasis on high and medium volume lanes during the bid process. While low volume lanes tend to be more volatile, high volume lanes can contribute to a higher deviation from budget.

Table 2 Independent Variables for Regression Models

Variable	**Details**	**Regression Columns**	**Hypothesis**
Distance	Measured per mile	Number of miles traveled per load on a specific lane	Budget deviation is positively correlated with distance
Lane Volume per Year (ABC lanes)	Percentage of load volume per lane. Segmentation: A lanes (high volume): >= 200 loads/year or at least 4 loads/week B lanes (medium volume): 12 < x < 200 loads/year or at least 1 load/month C lanes (low volume): < 12 loads/year or less than 1 load/month	Categorized as A, B, or C lane to indicate lane volume	Low volume lanes contribute to higher budget deviations
Origin/Destination (by State)	42 distinct origin states 25 distinct destination states	1 = equals specific origin/destination 0 = does not equal a specific origin/destination	Specific regions contribute to higher budget deviations
Lane Freight Type	Dedicated lanes by truck type (dry/refrigerated/frozen)	Categorized by lane freight type	Refrigerated and frozen lanes contribute to higher budget deviations compared to dry lanes

3.6 Model 1: Spend Ratio

We used OLS multiple linear regression to model the impact of lane volume, lane freight type, origin/destination, and distance on the percent deviation of spend. The equation to derive spend deviation is shown below:

$$Percent\ Deviation\ (Spend) = ((Actual\ Spend - Planned\ Spend) / Planned\ Spend)$$

$$Spend\ Ratio = Percent\ Deviation\ (Spend) + 1$$

The spend ratio captures the deviation of spend per lane that is independent of the lane volume. For example, if the spend deviation is a negative value (-0.056), it means that on average the total spend for loads on a specific lane was 5.6% lower than the contracted price. Our model uses the below equation to test the hypothesis that the four independent variables can predict the percent deviation on expenditure. In other words, given the combination of distance, lane freight types, lane volume (ABC), and origin/destination, how much over or under plan is our sponsor company spending per variable per lane?

$$Percent\ Deviation\ (Spend) = \beta 0 + \beta 1 * (X\_Dist) + \beta 2 * (X\_Frozen) + \beta 3 * (X\_Dryvan) + \beta 4 * (X\_A) + \beta 5 * (X\_C) + \beta 6 * (X\_O1) + \beta 7 * (X\_O2) + \ldots + \beta 45 * (X\_O41) + \beta 46 * (X\_D1) + \ldots + \beta 69 * (X\_D24)$$

3.7 Model 2: Volume Ratio

We used OLS multiple linear regression to model the impact of lane volume on the independent variables by capturing the deviation from budget. The equation to derive volume deviation is shown below:

$$Percent\ Deviation\ (Volume) = ((Actual\ Volume - Planned\ Volume) / Planned\ Volume)$$

$$Volume\ Ratio = Percent\ Deviation\ (Volume) + 1$$

Volume deviations can lead to budget overruns even though spend per lane is as planned. Similarly, if volume deviation is a negative value, that means that average volume per load on a specific lane is lower than contracted.

$$Percent\ Deviation\ (Volume) = \beta 0 + \beta 1*(X\_Dist) + \beta 2*(X\_Frozen) + \beta 3*(X\_Dryvan) + \beta 4*(X\_A)+ \beta 5*(X\_C) + \beta 6*(X\_O1) + \beta 7*(X\_O2)+\ldots+ \beta 45*(X\_O41)+ \beta 46*(X\_D1)+\ldots+ \beta 69*(X\_D24)$$

3.8 Model 3: Total Cost Contribution (Spend x Volume) Ratio

We used OLS multiple linear regression to model the impact of the total cost contribution spend and volume model captures the total deviation from budget per lane. The total cost contribution ratio is calculated by multiplying the spend ratio and volume ratio to account for the total effect of the two. For instance, the percent deviation from budget could be significantly higher because the rate that loads went for were higher than contracted; however, the volume was under the contracted amount. The opposite outcome could be that the percent deviation from budget could appear to be higher because the volume was higher than contracted but the rate was lower than contracted. Combining these two variables captures the total effect. Our model uses the below equation to test the hypothesis that the four independent variables can predict the percent deviation on annual volume and spend. In other words, given the combination of distance, lane freight types, lane volume (ABC), and origin/destination, how much over or under is the combined deviation on our sponsor company's budget per lane?

$$Percent\ Deviation\ (Total\ Cost\ Contribution) = \beta 0 + \beta 1*(X\_Dist) + \beta 2*(X\_Frozen) + \beta 3*(X\_Dryvan) + \beta 4*(X\_A)+ \beta 5*(X\_C) + \beta 6*(X\_O1) + \beta 7*(X\_O2)+\ldots+ \beta 45*(X\_O41)+ \beta 46*(X\_D1)+\ldots+ \beta 69*(X\_D24)$$

3.9 Model 4: Unplanned Lanes

The fourth and final model used a binary logistic regression model to test if unplanned lanes tend to have distinguishing characteristics. Unplanned lanes are those that were never awarded in a formal bid but volume did move on it over the course of a year, so there is no committed volume or contract price to compare to. Unplanned lanes are important to capture because they represent unanticipated spending and may be difficult to manage. The model is predicting given the input variables, how likely is a lane to be unplanned versus planned. The dependent variable (Y) is the binary variable which equals "1" for unplanned lanes and "0" for planned lanes. We implement a binary logistic regression for this model, which tells us the probability a lane is going to be unplanned (i.e. unexpectedly show up and cost a shipper money during a budgeted time period) based on the lane's characteristics. This model tests the independent variables distance, lane freight type, origin/destination but differs by including the sum of spend on an unplanned lane. This is because an unplanned lane would not have a planned cost or volume to compare the actual numbers to.

$$\text{Probability Lane is Unplanned } (Y=1) = P$$

$$Log(P/1-P) = \beta 0 + \beta 1*(X\_Dist) + \beta 2*(X\_Frozen) + \beta 3*(X\_Dryvan) + \beta 4*(X\_A) + \beta 5*(X\_C) + \beta 6*(X\_O1) + \beta 7*(X\_O2) + \ldots + \beta 45*(X\_O41) + \beta 46*(X\_D1) + \ldots + \beta 69*(X\_D24)$$

4. DATA ANALYSIS RESULTS & DISCUSSION

In this chapter, we discuss the data and present the models developed using the freight category segmentation methodology described in section 3.4. The data summary is presented in two sections. The first section summarizes the data findings; the second section discusses the results of our models.

4.1 Freight Categorization Findings

The figures in this section illustrate the distribution of the loads across the four categories by size of shipper. The Shipper's Freight Categorization figures in this section show how 2019 load transactions fall into the categories by showing percentage of loads transactions over the course of a year. We evaluate the routing guide's performance starting from March (week 9) 2019 through March (week 9) 2020 of the following year as this is when the annual bid rates take effect for the duration of the year.

Overall, the graphs show that the routing guide's performance improved over the course of the year, contrary to our assumption that we would see degradation over time. The percentage of loads that were planned and on budget improved over that time frame. Table 3 shows 2019 averaged 80% of the loads tendered were planned and on budget. Further analysis is conducted to understand the attributes that influence the likelihood of a load being unplanned, meaning that the lane was not accounted for in the annual bidding process, which leads to a budget overrun. In addition, we analyze the deviations in expenditure and volume from plan at the lane and load level to understand where the budget deviations are originating.

Table 3: Total Freight Categorization Breakdown (2019)

Freight Category	Price	% of Total Loads	% of Lanes
Planned, On Budget	Contract (+/-2%)	80%	78%
Planned, Over Budget	> Contract	4%	3%
Planned, Under Budget	< Contract	5%	5%
Unplanned	N/A	12%	14%
		100%	100%

Table 3 displays the total number of loads that fall into the four freight categories of the large, medium, and small shippers combined in 2019. Ghost lanes did not appear in our sponsor company's data. Table 3 shows that the percentage of loads that were planned and on budget is 80% on average during the year. Additionally, 78% of lanes that were planned and on budget accounted for 80% of loads tendered during the year. Next, we characterize the loads that were unplanned and seek to understand whether specific attributes are the reason these lanes are not accounted for during the annual bid auction.

Based on discussions with our sponsor company, we classified the three shippers into large, medium, and small based on annual volume and spend. The large shipper is referred to as Company A, medium shipper as Company B, and the two small shippers are examined jointly and are referred to as Company C. Figure 8 illustrates the routing guides performance for Company A, the large shipper. About 80% of loads were planned and on budget throughout the year but started to drop into the low 70% range by the end of the year. Furthermore, the routing guide's performance did not degrade over time as initially expected; it maintained a steady planned and on budget performance. Company A maintained a steady percentage of loads on unplanned lanes compared to the smaller shippers that were more volatile.

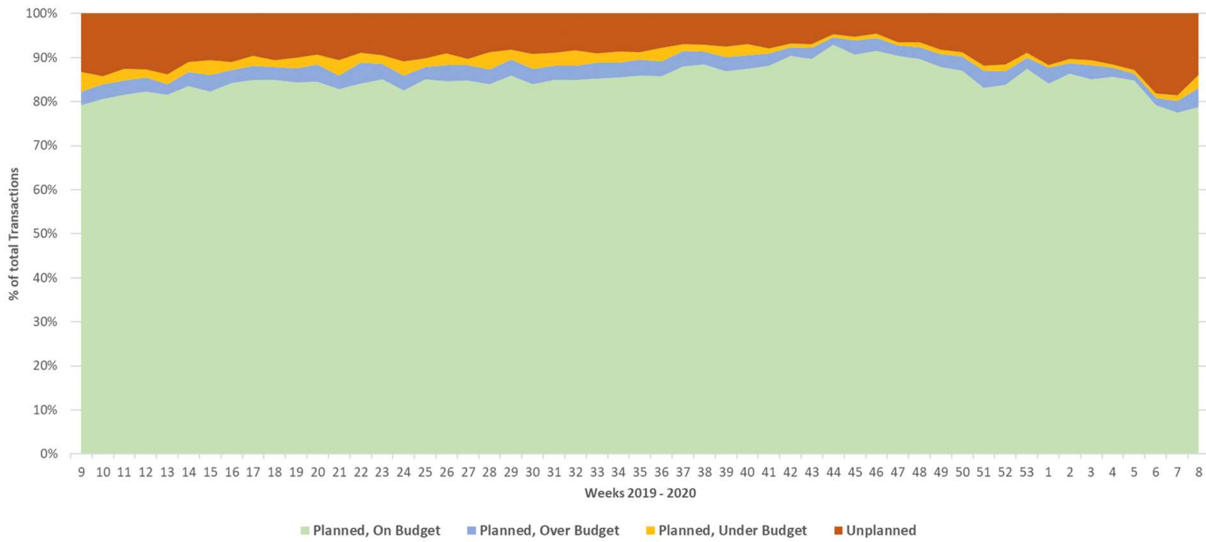


Figure 8: Company A's Freight Categorization

Figure 9 illustrates Company A's deviation on contract price with the y-axis representing the number of lanes and the x–axis representing the percent spend deviation. The y-axis is cut-off at 400 lanes to show greater detail. 91% of lanes fall within +/- 6% from contract price. The histogram of lanes is represented in each line in the graph, which is the count of lanes that went at different spend deviations. For example, the line at 54% is a lane that went at 54% above contract price on average. Figure 9 provides insight as to the extent to which Company A's lanes deviate from contract price. This graph provides a deeper insight into the distribution of lanes that fall into planned, over budget and planned, and under budget. Company A has a longer right tail than the other shipper sizes shown, which indicates that a greater number of the lanes were carried out at a higher cost than anticipated as compared to the budget.

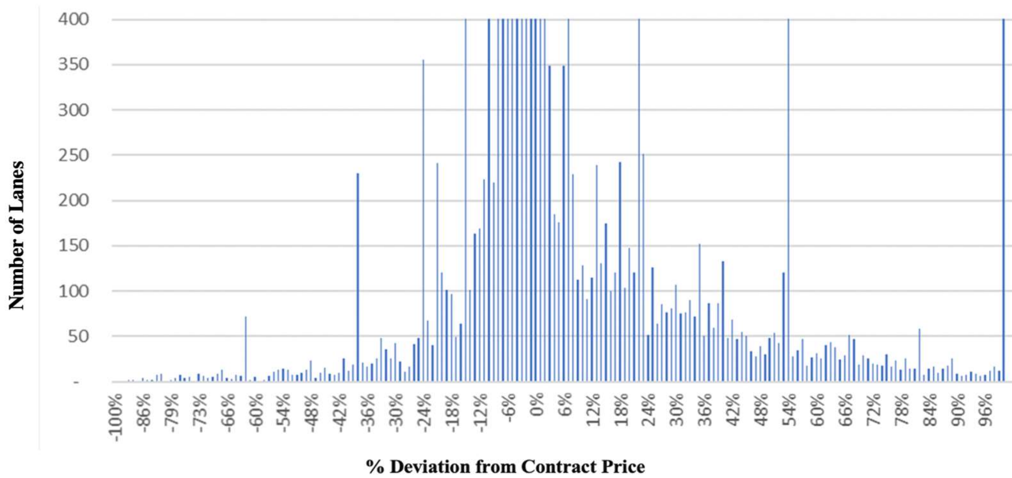


Figure 9: Company A's Deviation on Contract Price (2019)

Each shipper's distribution of volume per lane is different. Using the ABC analysis, which is a comprehensive way to segment each shipper's lane by high, medium, and low volume, we want to determine where shippers should focus their resources. Our sponsor company classifies as high volume lanes (A) those that have at least 200 loads per year or operationally at least 4 loads per week. These lanes contribute heavily to the annual spend. Medium volume (B) lanes are categorized as having between 12 to 200 loads per year or at least 1 per month. Finally, low volume (C) lanes are those that have under 12 loads per year or less than 1 per month. Breaking down lanes into three distinct segments allows us to test the impact of volume on the annual budget.

Table 4 shows the ABC segmentation for Company A. The table shows that high volume lanes make up 55% of total volume per year. While the majority of lanes are either medium or low volume lanes making up 45% of the total volume for Company A. In addition, the total number of lanes bid out are 3,600, which also has a higher concentration of loads on a lane as compared to Company B.

Table 4: ABC Segmentation for Company A

Lane Segments	Number of Lanes	% of Lanes	% of Volume	Label
A	384	11%	55%	High Volume
B	2,166	60%	44%	Medium Volume
C	1,050	29%	1%	Low Volume
Total	**3,600**	**100%**	**100%**	

Figure 10 shows Company B's freight categorization that had the second largest number of loads supported by our sponsor company. However, the average percentage of loads that are planned and on budget fluctuated between 60-90% throughout the year with greater volatility earlier in the year. Company B has a greater percentage of unplanned loads, especially when the 2019 bid rates took effect in week 9 but improved throughout the year. Overall, Company B's routing guides performance is less stable than that of Company's A and C.

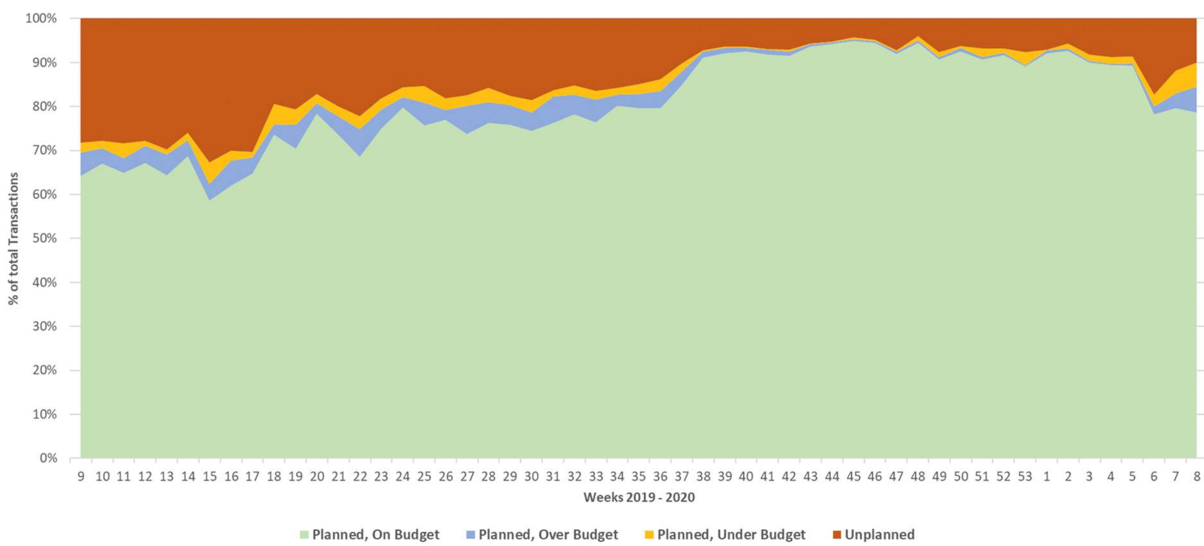


Figure 10: Company B's Freight Categorization

Figures 11 illustrates Company B's deviation on contract price with the y-axis representing the number of lanes and the x-axis representing the percent spend deviation. The y-axis is cut-off at 400 lanes to show greater detail. 89% of lanes fall within +/- 6% from contract price. Although the Company B has the second largest percentage of volume supported by our sponsor company, the lanes are less dense. Figure 11 distribution has a left tail, which indicates that a majority of loads on a specific lane averaged a lower price than contracted. However, one particular lane averaged a 100% increase in contract price per load.

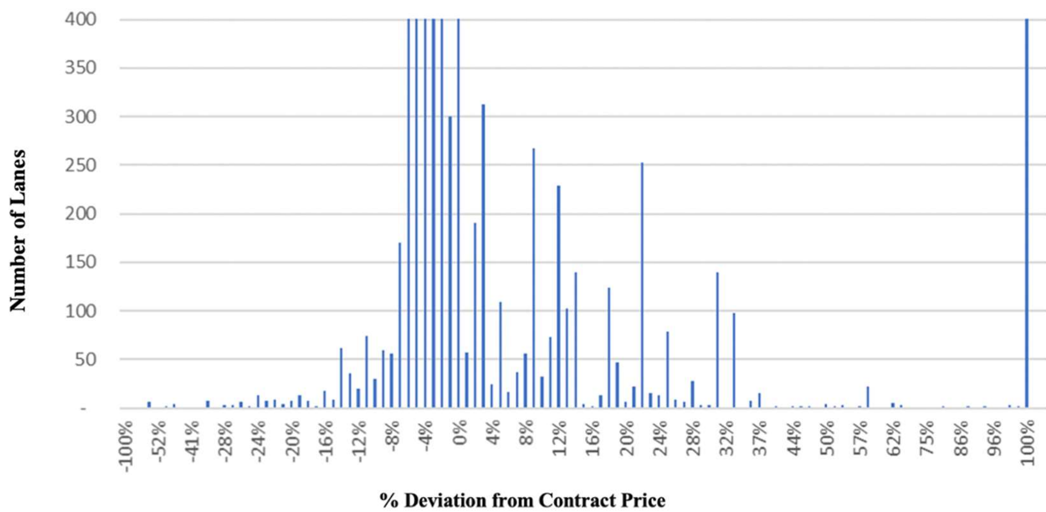


Figure 11: Company B's Deviation Contract Price

Table 5 shows the ABC segmentation for Company B. The table shows that 7% of lanes are high volume making up 37% of volume per year. Medium volume lanes make up 60% of the total lanes for Company B with 60% of total volume. Finally, low volume lanes make up 3% of total volume. While the percentage of lanes that fall into ABC are similar for Company A and Company B, the percentage of volume on A and B lanes are significantly different. Less than a third of lanes are high volume for company B.

Table 5: ABC Segmentation for Company B

Lane Segments	**Number of Lanes**	**% of Lanes**	**% of Volume**	**Label**
A	124	7%	37%	High Volume
B	1,102	60%	60%	Medium Volume
C	620	34%	3%	Low Volume
Total	**1,846**	**100%**	**100%**	

Figure 12 illustrates Company C's planned and on budget loads are slightly more volatile than the overall distribution for 2019. Additionally, on average about 20% of loads were unplanned. This shipper represents a different segment in the restaurant industry that makes its freight characteristics unique. In general, the distribution and percentage of planned and on budget loads remain consistent throughout the year.

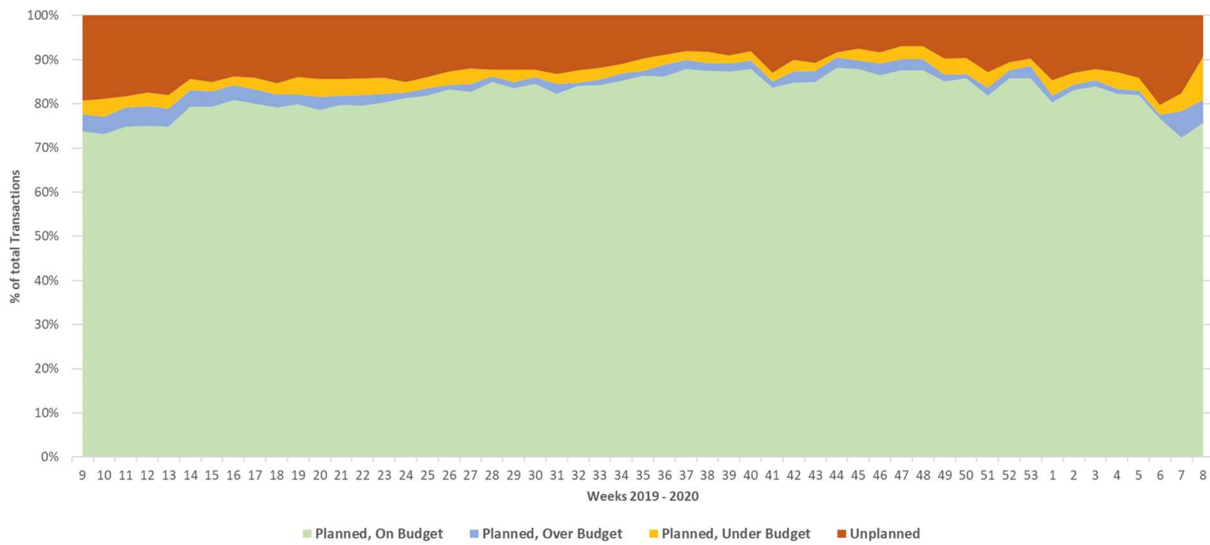


Figure 12: Company C's Freight Categorization

Figures 13 illustrates Company C's deviation on contract price with the y-axis representing the number of lanes and the right x-axis representing the percent deviation in spend. The y-axis is cut-off at 400 lanes to show greater detail. 83% of lanes fall within +/- 6% from contract price. Company C appears to have a clear left tail with a majority of lanes moving at rates lower than contract price. Figure 13, further shows that Company C does not have as much volatility in contract price compared to the medium shipper but does have one lane that averaged a 96% increase in contract price per load.

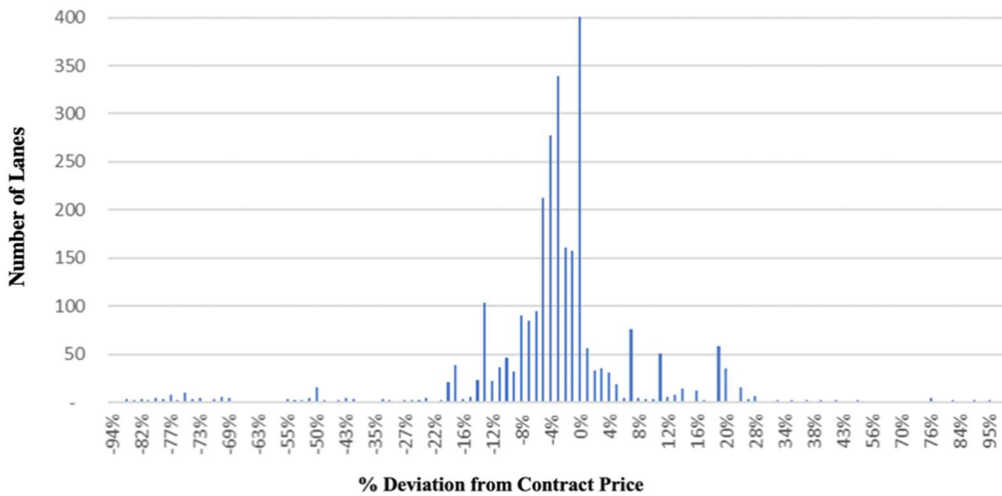


Figure 13: Company C's Deviation Contract Price (2019)

Table 6 shows the ABC segmentation for Company C. The table shows that only 5% of lanes are high volume with 26% of total volume on those lanes. Company C bid 1,704 lanes in 2019. The vast majority of lanes are medium volume. The distribution of volume per lane segment is significantly different than the other two companies. This is because the percentage of volume on each lane segment, particularly for high volume lanes, is much lower for Company C.

Table 6: ABC Segmentation for Company C

Lane Segments	Number of Lanes	% of Lanes	% of Volume	Label
A	89	5%	26%	High Volume
B	1,036	61%	70%	Medium Volume
C	579	34%	4%	Low Volume
Total	**1,704**	**100%**	**100%**	

Analyzing the overall performance and contract price deviations per shipper provides a clearer understanding of their routing guide's performance over the year, the following section will analyze how the volume contributes to the percent deviation in a budget overrun. Figure 14-16 show the percentage of loads on a specific lane that deviate from the annual bid per shipper size. The x-axis represents the spend deviation of expenditure, which is calculated by evaluating the actual cost over the contracted cost. The y-axis represents the volume deviation, which is calculated by evaluating the actual volume per year over the contracted volume per year. Each dot on the graph represents a lane. For example, the dot on the far-right side of Figure 13 positioned at 0.06 (x-axis) and 2.00 (y-axis) indicates that the loads that occurred on that specific lane averaged 6% higher cost and volume was 2.00 times more than in the bid. By graphing volume and expenditure (price) ratios per lane, we can better understand the total cost contribution on the budget by differentiating between planned, over budget and planned, under budget loads. For example, loads on a specific lane could average a higher contract price but the annual volume was lower than committed so that lane would be categorized as planned, under budget. The results of each lane's performance on volume and expenditure deviations is included in the regression modeling.

Figure 14 shows the percentage of loads on a specific lane that deviate from the annual bid for Company A. The majority of lanes are centered on the y-axis, meaning that they met the contract price per load but depending on the position deviated from volume. Furthermore, Company A has a higher percentage of lanes that have extreme values. For instance, the blue dot on the right side of the graph hovering to the left of 0.05 (x-axis) and 3.45 (y-axis) means that the lane averaged a 5% contract price increase while the volume is 3.45 times the volume committed. Regression modeling tells us the specific attributes on that lane that cause it to have higher expenditure and volume depending on the input variables. Company A appears to have more clustering in the center of the graph indicating that they majority of lanes tendered are on budget. The distribution aligns with the freight categorization findings that 78% of lanes are planned and on budget. Company A has more lanes which is the reason that there is a higher number of lanes on extreme ends of the graph in terms of volume and expenditure.

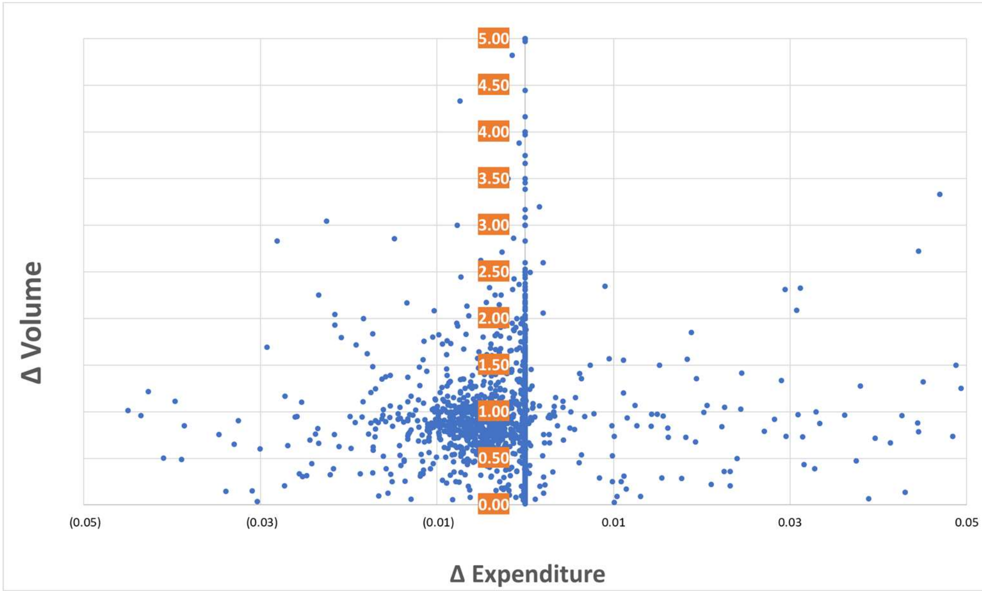


Figure 14: Company A's Percent Deviation in Volume & Expenditure

Figure 15 shows the percentage of loads on a specific lane that deviate from the annual bid for Company B. Similar to Company A, the majority of lanes are centered on the y-axis, meaning that they met the contracted price per load but deviate from the contracted yearly volume. However, the distribution of lanes that averaged extreme values for expenditure and volume deviations are more volatile. Additionally, there is greater volatility in terms of expenditure versus volume deviations. Relative to its size, comparing Company A and B shows that Company B tends to have more lanes on average that deviate from both volume and expenditure. Overall, the results of these graphs shows a trend across all three shipper sizes, that volume tends to be a larger contributor to budget deviation than expenditure.

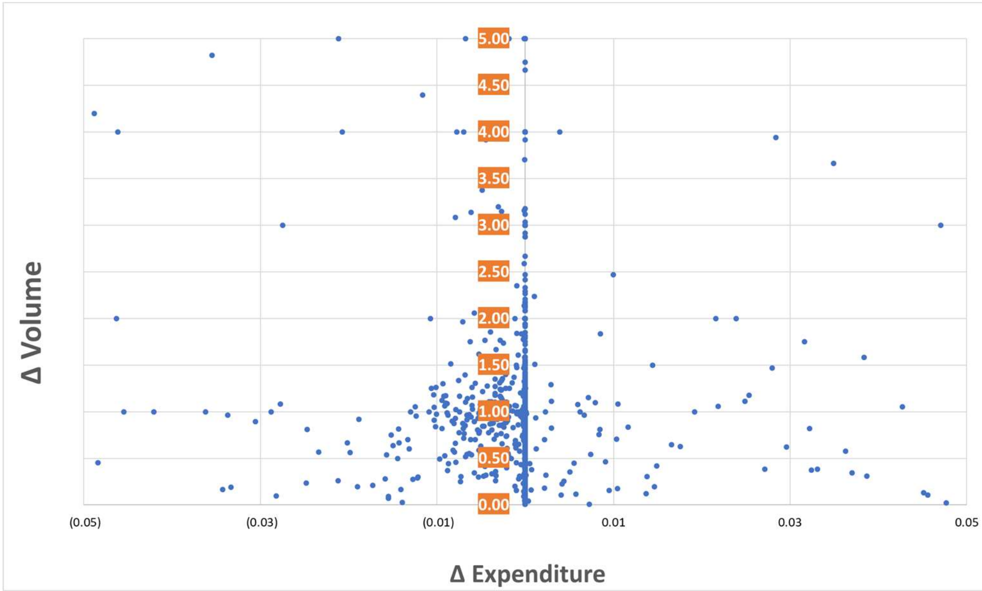


Figure 15: Company B's Percent Deviation in Volume & Expenditure

Figure 16 shows the percentage of loads on a specific lane that deviate from the annual bid for Company C. Company C allows us to compare if a different business model from the other two companies impacts the performance positively or negatively. Similarly, the majority of lanes are centered on the y-axis, meaning that they met the contracted price per load but deviate from contracted yearly volume. It appears that relative to the total number of lanes tendered by Company C, there are more lanes that averaged a lower expenditure but met committed volume. However, there appears to be more extreme deviations from expenditure compared to Company A and B. The main difference is the number of lanes that resulted in extreme deviations for expenditure and volume is the lowest of the three companies. The spread of lanes is clustered to the left side of the graph indicating that expenditure tends to be lower than budgeted.

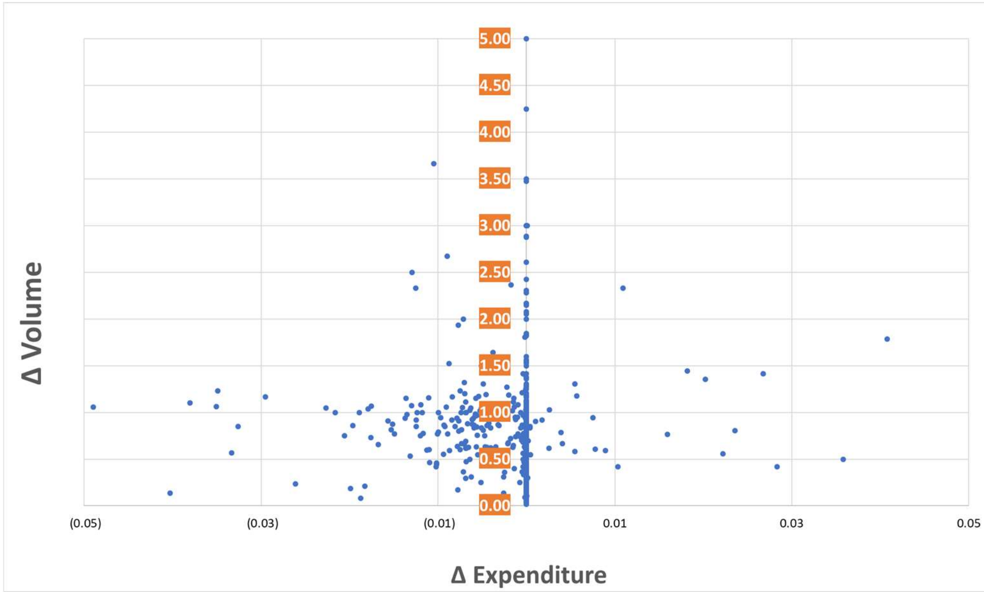


Figure 16: Company C's Percent Deviation in Volume & Expenditure

4.2 Cadence, CV, and Lane Volume Impact

In the regression analysis, the independent variable lane volume (ABC) is categorized as high, medium, and low volume. The response (dependent) variable is total cost deviation from budget per shipper. The models did not provide degree of high confidence (low adjusted R^2). Our hypothesis that low volume (C) lanes contribute to higher budget deviations did not prove to be the case. We evaluate this hypothesis in an alternative method by graphing each lane in a scatter plot by the number of loads (x-axis) and coefficient of variation (CV) (y-axis). The CV is defined as the ratio of the standard deviation to the mean of loads per week. This allows us to compare lanes of varying volume to evaluate the variability in the percentage of weeks that have a load on a lane. In addition, we evaluate cadence as a measure of percentage of weeks with loads on a lane per shipper. Figure 17 is the variation of loads with lanes representing the dots in the graph. For example, the dot on the far right at 2,500 loads on the x-axis and 0.1 on the y-axis means that for this high-volume lane, the variation of loads per week is low. The vertical lines represent the cut-off of lane volume categories (ABC). For example, all data points to the right of the orange vertical line are high volume (A) lanes (at least 200 loads per year). The data points between the blue and orange vertical lines are medium volume (B) lanes (between 12-200 loads per year). In Company A's case, only 1% of volume is on low volume (C) lanes which is the reason the lines are skewed to the left of the graph. A lower CV has values that are close to the mean and has a lower standard deviation of loads on a lane. The results of this graph indicate that overall, higher volume lanes tend to have less variability while low volume lanes have a larger spread in terms of CV values and thus, variability in weekly load volumes.

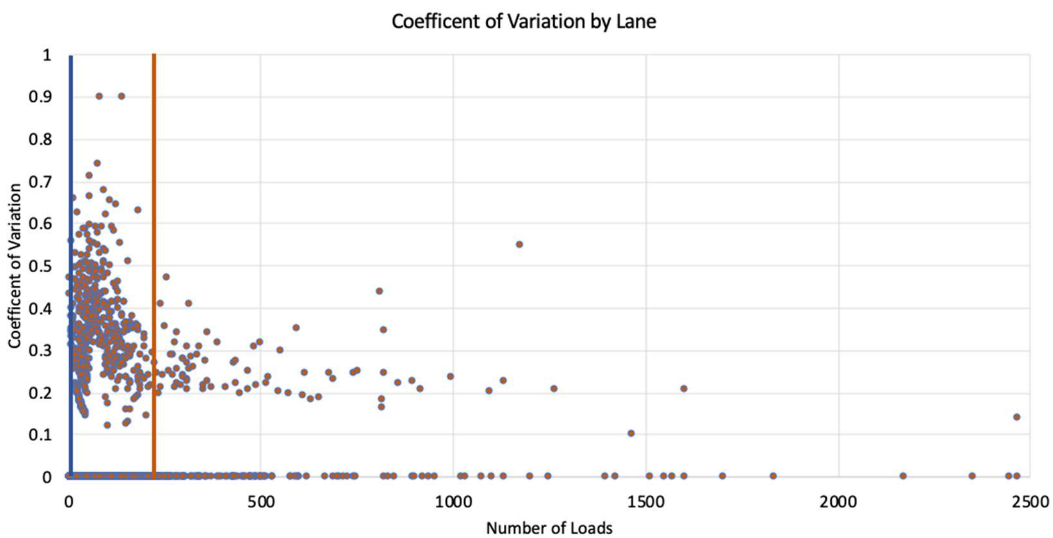


Figure 17: Coefficient of Variation by Lane (Company A)

Figure 18 shows the cadence of loads or percent of weeks that have loads on a lane. The majority of Company A's lanes has loads that occur on 25% of the weeks in the year or less (i.e., low cadence). Company A has almost an even split of lanes with loads that have low and high cadence. A considerable portion of volume tendered is likely moved in 25% of the week in the year or less.

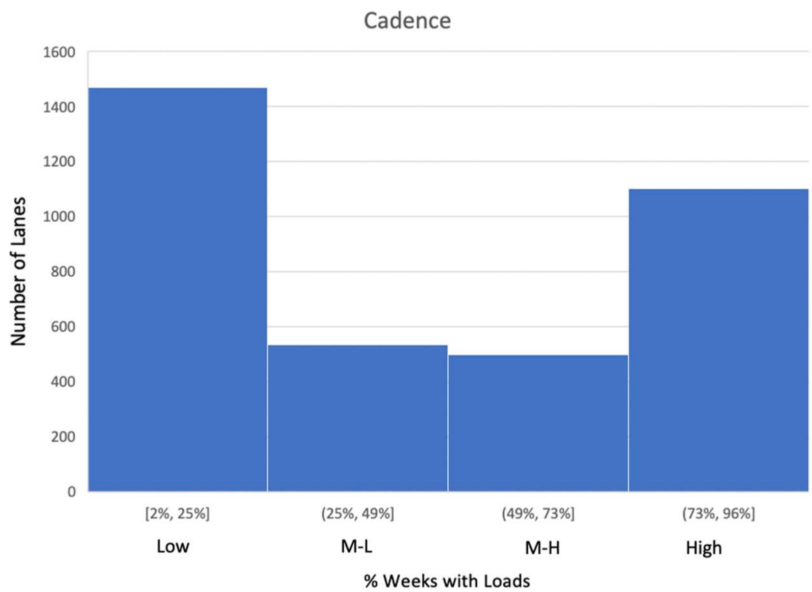


Figure 18: Cadence (Company A)

Table 7 compares the behavior of cadence, CV, and lane volume on Company A's budget. We segment the performance of lanes into four categories: planned, on budget (P=B), planned, over budget (P>B), planned, under budget (P<B), and unplanned (UP). These segments take volume into account when comparing to the budget. For example, a lane could be above budget due to volume deviations rather than spend. Planned, on budget transactions are aggregated at the lane level where deviation +/- 10% is considered on budget. While previously we categorized planned, on budget load transactions to be within +/- 2% of budget.

This is because the difference between lane volume and number of lanes will cause challenges when comparing the results. Table 7 summarizes how different lane characteristics correlate to budget performance. The results show that unplanned lanes tend to be low cadence, high variability (CV), and low volume. These results highlight how volume tends to be the dominating contributor for budget overruns. Company A has the largest number of total lanes and volume per year that has a greater percentage of lanes in all categories that are on budget.

Table 7: Cadence, CV, Lane Volume (ABC) Comparison for Company A

	% of Lanes	% of Volume	Avg Yearly Volume/Lane	P=B	P>B	P<B	UP
Cadence							
Low	41%	6%	12	21%	1%	1%	77%
M-L	14%	7%	44	72%	3%	1%	24%
M-H	14%	10%	61	86%	2%	2%	10%
High	31%	77%	219	92%	2%	2%	4%
Coefficient of Variation (CV)							
Low	67%	79%	103	67%	2%	1%	30%
Med	21%	20%	82	63%	2%	3%	32%
High	12%	1%	9	9%	1%	1%	89%
Lane Volume (ABC)							
Low (C)	29%	1%	4	13%	1%	1%	85%
Med (B)	60%	44%	63	77%	2%	1%	20%
High (A)	11%	55%	449	85%	3%	2%	10%

Figure 19 shows the variation of loads by lane for Company B. Similar to Company A, the graph indicates that higher volume lanes tend to have less variability while low volume lanes vary number of loads per week. However, Company B has fewer lanes with very high CV above 0.6.

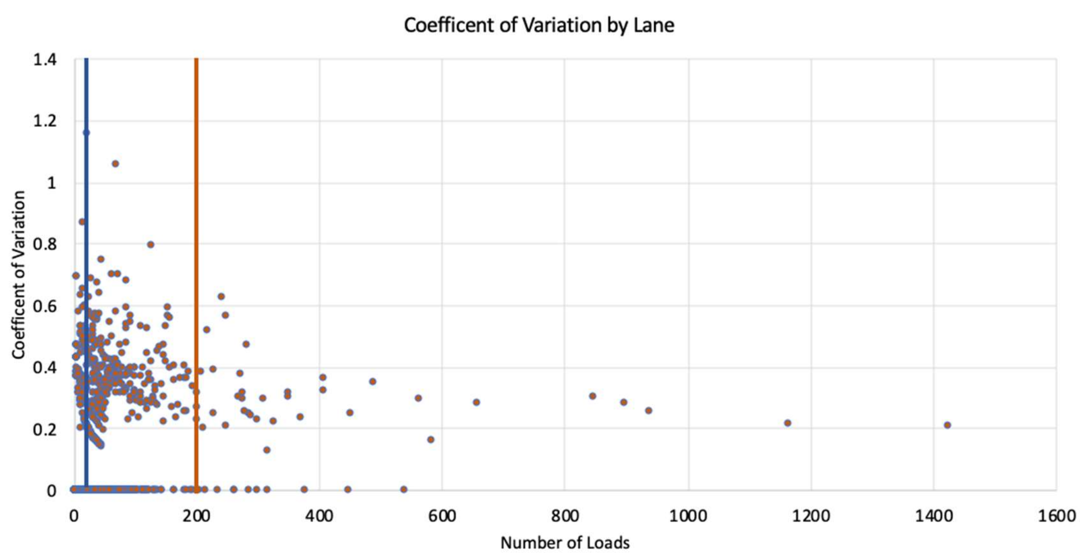


Figure 19: Coefficient of Variation by Lane (Company B)

Figure 20 is the cadence of loads or percent of weeks that have loads on a lane. The Company B shows a distribution of lanes per percentage of weeks that is skewed more to low cadence. The majority of the lanes have loads that occur on 25% of weeks in a year or less.

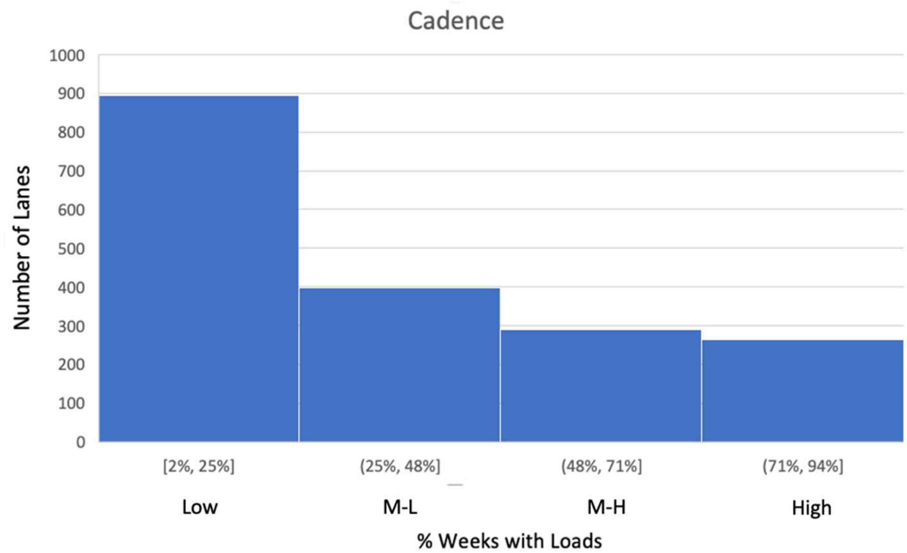


Figure 20: Cadence (Company B)

Table 8 compares the behavior of cadence, CV, and lane volume on Company B's budget. Similar trends for unplanned lane characteristics are apparent in the results. In general, lanes that are low cadence, high variability (CV), and low volume tend to be unplanned. Overall, the percentage of lanes and volume that is on budget aligns is the findings in this research that our sponsor company's routing guide is strong. The main difference between Company A and B's results is that the majority of lanes in Table 8 are medium volume (B) with volume clustered on low cadence and low variability lanes. However, high volume (A) lanes accounting for 37% of volume in the year included 18% of lanes that were unplanned. Our sponsor company could evaluate these lanes and determine if they are one-time occurrence or if they should be included in the next year's bid.

Table 8: Cadence, CV, Lane Volume (ABC) Comparison for Company B

	% of Lanes	% of Volume	Avg Yearly Volume/Lane	P=B	P>B	P<B	UP
Cadence							
Low	48%	10%	10	44%	4%	2%	50%
M-L	24%	20%	40	77%	1%	1%	20%
M-H	14%	20%	71	82%	2%	2%	13%
High	14%	50%	179	85%	6%	2%	8%
Coefficient of Variation (CV)							
Low	65%	63%	47	66%	3%	2%	29%
Med	22%	20%	45	59%	3%	2%	35%
High	13%	17%	64	53%	5%	3%	39%
Lane Volume (ABC)							
Low (C)	34%	3%	5	38%	3%	3%	57%
Med (B)	60%	60%	49	77%	2%	1%	19%
High (A)	7%	37%	263	67%	3%	3%	18%

Figure 21 is the variation of loads by lane for Company C. Similar to Company A, the graph indicates that higher volume lanes tend to have less variability while low volume lanes vary number of loads per week. However, Company C has a more random distribution of variability on lane which may be attributed to the different restaurant model as compared to Companies A and B.

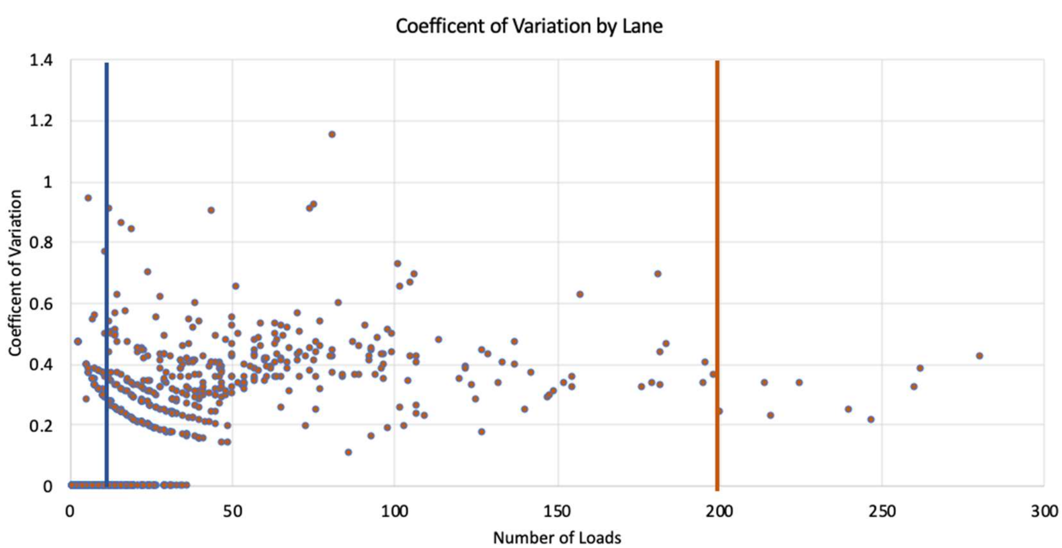


Figure 21: Coefficient of Variation by Lane (Company C)

Figure 22 is the cadence of loads or percent of weeks that have loads on a lane. Company C shows a distribution of lanes per percentage of weeks with loads with more of a left tail indicating that majority of lanes have loads that occur on 25% of weeks in a year or less. This aligns with the average load per week being on the lower end that allows for less shipments in the year.

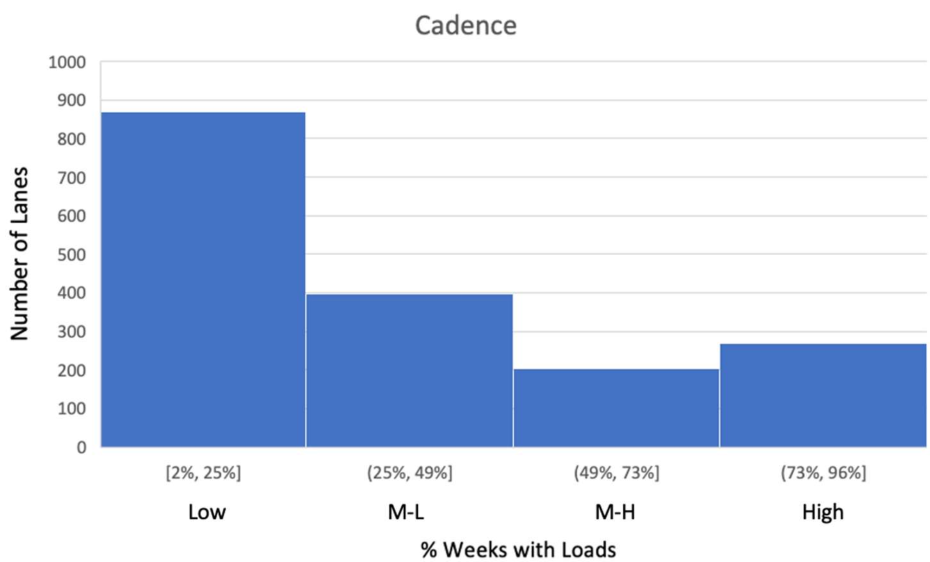


Figure 22: Cadence (Company C)

Table 9 displays the behavior of cadence, CV, and lane volume on Company C's budget. These results confirm the trend that regardless of shipper size, unplanned lanes tend to be on low cadence, high CV, and low volume lanes. The main difference is that volume concentration is higher for the Company A than the medium and small sized companies. This is likely due to the difference in the number of distribution centers that are serviced by each shipper.

Table 9: Cadence, CV, Lane Volume (ABC) Comparison for Company C

	% of Lanes	% of Volume	Avg Yearly Volume/Lane	P=B	P>B	P<B	UP
Cadence							
Low	50%	45%	37	56%	3%	2%	39%
M-L	23%	17%	31	58%	2%	2%	38%
M-H	11%	13%	48	72%	1%	4%	23%
High	16%	25%	64	78%	3%	2%	17%
Coefficient of Variation (CV)							
Low	70%	64%	38	63%	2%	3%	32%
Med	15%	19%	51	65%	3%	2%	30%
High	15%	17%	47	53%	3%	3%	41%
Lane Volume (ABC)							
Low (C)	34%	4%	4	25%	2%	2%	71%
Med (B)	61%	70%	48	81%	3%	2%	14%
High (A)	5%	26%	206	87%	2%	1%	10%

Cadence, CV, and lane volume contribute to volume deviations from budget in a similar trend across all three shipper sizes. In particular, CV categories have more of a uniform distribution of unplanned lanes that fall into either low, medium, and high. However, Company A has more unplanned lanes skewed to high CV or variability in weekly loads. A reason for this could be that that 99% of the total volume per year has low or medium variability (CV). To further analyze how the lane categories contribute to price deviations, we evaluated performance based on lane pricing independent of volume. Table 10 displays the average price deviation per lane across all categories for the three shipper sizes. The values are calculated by averaging the percent deviation of each load in a lane from the expected price. Then taking the average of all the lanes deviations for each category to highlight trends in the behavior across shippers. For example, low cadence lanes for Company A average 1.71% higher cost than budgeted. Low cadence, high CV, and low volume lanes tend to deviate the largest from budget in terms of price. Company C, the smallest shipper, has the highest price deviations on average across most of the categories, particularly high CV lanes. The analysis in Table 10 further supports that volume deviations are dwarfing missed budget compared to price deviations.

Table 10: Average Percent Price Deviation (All Shippers)

	Company A (Large)	Company B (Medium)	Company C (Small)
Cadence			
Low	1.71%	0.65%	1.45%
M-L	1.25%	0.12%	1.10%
M-H	0.77%	0.86%	-0.41%
High	0.34%	0.44%	0.82%
Coefficient of Variation (CV)			
Low	0.9%	0.2%	0.41%
Med	0.7%	1.4%	0.78%
High	1.0%	0.6%	4.51%
Lane Volume (ABC)			
Low (C)	1.8%	0.71%	1.79%
Med (B)	0.7%	0.46%	0.94%
High (A)	0.7%	0.18%	-0.06%

4.3 Regression Modeling Results

Linear regression models were built to quantify the impact of different variables on the deviation from budget. We built four models that model the impact on spend, volume, total cost contribution deviation, and unplanned lanes. The independent variables tested include lane volume, lane freight type, origin/destination, and distance. The regression model results for all three shipper sizes are shown in Tables 7–9. Results for the independent variables are considered statistically significant if p-value < 0.1. The full results for the 71 key variables are shown in Appendix B–D for each shipper.

In this section, we discuss the results of the regression models impact on the independent variables percent deviations from budget for Company A. Full model results for Company B and C on spend, volume, and total cost contribution deviations are shown in Appendix B; however, this section discusses the results from the total cost contribution model in greater detail. The total cost contribution model captures the total deviation from budget per lane. We remove one origin and one destination, and one lane volume (ABC), in order to avoid multicollinearity. These eliminated variables are our base case and are partially represented by the intercept. Each independent variable's coefficient has an associated p-value. The p-value is the measure of how meaningful the independent variable is to the response variable (Y). An independent variable with a low p-value is likely to be a meaningful addition to the model because changes in that independent variable's value are related to changes in the response variable. We will consider p-values that are less than 0.1 to be a meaningful variable in our model. Figure 23 illustrates the distribution of total cost ratio ranges for Company A. While the majority of lanes deviate above budget by 2 or more times, the concentration of total volume is in lanes that are within 0.9–1.1 of budget.

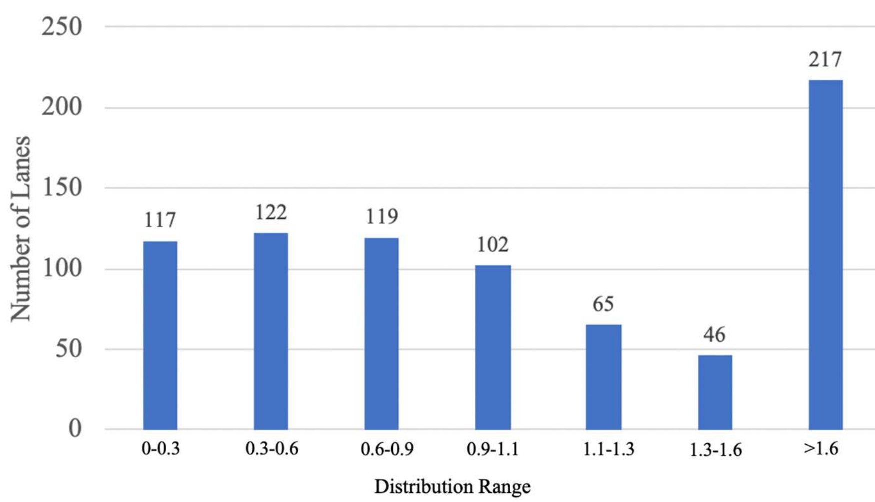


Figure 23: Total Cost Distribution for Company A

Table 11 displays the total cost deviations corresponding to each variable compared to the base case. The base case in this model is a refrigerated, medium volume lane (B), that has an origin in California and destination in Georgia. Although the adjusted R^2 is very low (0.01), suggesting that the variables selected are not sufficient to explain the deviation in the response variable, we can see that lane freight types have an impact on percent deviation from total cost of a load. As shown in Table 11, dry truck loads contribute to 0.88 percentage points lower in deviation from budget compared to frozen trucks that account for an average of 0.86 percentage points more in deviation from total cost. However, contrary to our hypothesis that budget deviations are positively correlated with distance, the model suggests that is not the case. Distance and lane volume (ABC) variables not shown to be significant. The results suggest that the origin Georgia (GA) and destination Pennsylvania (denoted DPA) have an impact with an increase from total budget of 6.5 percentage points and 5.20 percentage points, respectively.

Table 11: Total Cost Contribution Deviation Results (Company A)

Adjusted R^2	**0.01**			
Term	**Estimate**	**Std Error**	**t Ratio**	**P-value**
Intercept	1.26	0.78	1.63	0.1041
Distance	0.00	0.00	1.21	0.2257
Temperature [DRY]	-0.88	0.55	-1.59	0.1111
Temperature [FROZEN]	0.86	0.64	1.34	0.1808
ABC_A	-0.93	1.24	-0.75	0.4508
ABC_C	-1.78	1.45	-1.22	0.2213
GA	6.50	1.59	4.1	<.0001
DPA	5.20	1.62	3.2	0.0014

4.4 Regression Results Summary

Developing the freight categorization figures to illustrate the distribution of loads across four categories for three shipper sizes (large, medium, and small) provides an answer to the first part of our research question: how well do routing guides perform? Overall, the findings show that the routing guide's performance improved or remained stable over the course of the year, contrary to our assumption that we would see degradation over time. Across all three shipper sizes on average 80% of total loads were planned and on budget. This indicates that our sponsor company's routing guide performed as expected for the majority of loads with just 4 % of lanes over budget or 5% below. The freight categories do not provide insight as to the reason lanes are over or under budget because this could be a result of spend and/or volume deviations. The remaining 12% of loads were unplanned, meaning that the lane was never bid.

The second research question looks to provide insights to our sponsor company to improve routing guide compliance for the future. After developing the regression models as described in Chapter 4, we analyzed how independent variables such as distance, lane volume, freight lane type, and origin/destination impact the spend, volume, and total cost deviations from budget. Although the models did not produce a high adjusted R^2, we summarize the impact of the independent variables on the response variables and highlight whether our hypotheses in Table 2 are confirmed. Table 12 reflects "No Impact" if the independent variables were not significant (p-value < 0.1) in impacting the model.

Table 12: Regression Summary of Independent & Dependent Variables

Variable	**Spend Ratio (Model 1)**	**Volume Ratio (Model 2)**	**Total Cost Ratio (Model 3)**
Distance	No Impact	No Impact	No Impact
Lane Volume per Year (ABC lanes)	No Impact	Low volume lanes (C) tend to decrease volume deviation from budget	No Impact
Origin/Destination (by State)	Specific states contribute deviation from budget	Specific states contribute deviation from budget	Specific states contribute deviation from budget
Lane Freight Type	Dry freight tends to decrease deviation while frozen freight	Dry freight tends to decrease deviation while frozen freight	Dry freight tends to decrease deviation while frozen freight

	tends to increase deviation from budget.	tends to increase deviation from budget.	tends to increase deviation from budget.

The impact of distance on all three models is counterintuitive to our initial hypothesis that longer lanes increase deviation of spend, volume, and total cost from budget. Distance has no impact in the results which may be because it is factored into the original bid contracts. The lane volume (ABC) categorized by high, medium, and low volume lanes was also the opposite of our hypothesis that the lower a volume lane, the higher the deviation in spend, volume, and total cost. However, low volume lanes (C) tend to decrease volume deviation from budget. This may be attributed to the fact that this is compared to the majority of lanes being medium volume (B) lanes, which is the base case for all three models. The impact of origin/destination states is aligned with our hypothesis that some regions result in a higher or lower deviation across all three response variables (spend, volume, and total cost contribution) from budget. However, the impact of individual states is different for each shipper size. These individual results per shipper are shown in Figures 24-29 below. Finally, lane freight types strongly align with our hypothesis that dry truck loads tend to have lower deviation while frozen truck loads tend to see higher deviation from budget across all response variables.

In order to develop a segmentation strategy that takes into account the variability of these independent variables, we evaluate the spend, volume, and total cost contribution by state for each shipper. In this section, we discuss the average total cost contribution deviation per origin and destination states. The maps for spend and volume deviation per origin and destination are in Appendix E. Figure 24 shows each origin state’s average total cost contribution deviation for Company A referenced by the distribution legend. The total cost contribution captures the total deviation from budget per lane. The total cost contribution ratio is calculated by multiplying the spend ratio and volume ratio to account for the total effect of the two. Combining these two variables captures the total effect. The total cost contribution deviation refers to percentage over or under budget. For example, states in red are those that average the highest total deviation from budget. As compared to the base case, the majority of Company A’s origin states average exact or below total cost deviation from budget. In addition, the East Coast region tends to exceed budget in terms of total cost by both origin and destinations.

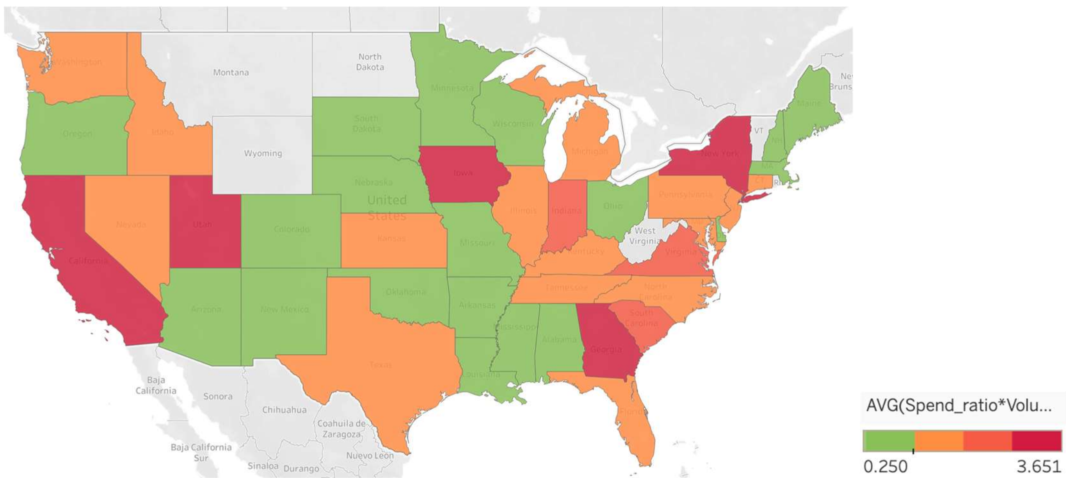


Figure 24: Total Cost Deviation by Origin (Company A)

Figure 25 shows each destination state’s average total cost contribution deviation for Company A. There are 42 origin to 25 destination states for the largest shipper. Some states, like California and Texas, average higher deviation from budget, while other states have improved results depending on where the loads originate. Overall, the sponsor company can focus on destinations in order to have the greatest potential to improve deviations from budget as performance from destinations is lower than that of origin states.

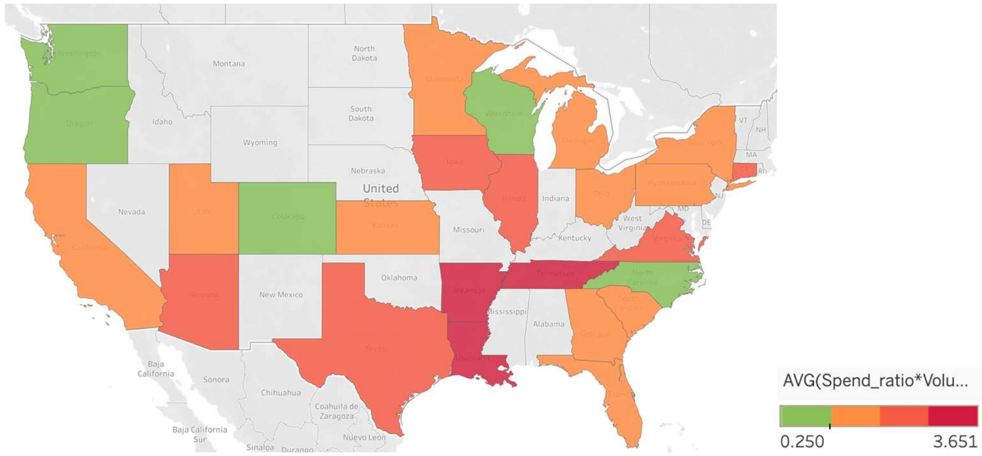


Figure 25: Total Cost Deviation by Destination (Company A)

Figure 26 illustrates the origin states' average total cost contribution deviation for Company B. Compared to Company A, there are far fewer origin states, but the overall deviation from budget is lower. Pennsylvania and Illinois have the highest deviations for origin states.

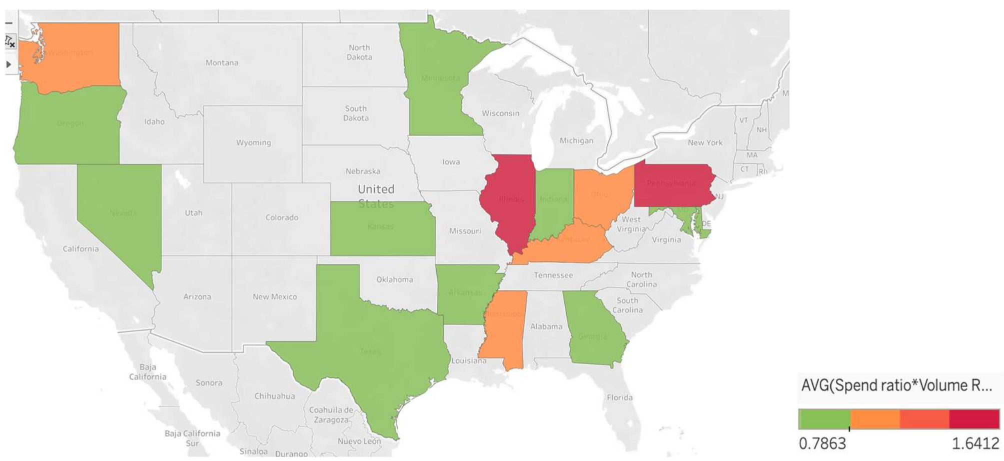


Figure 26: Total Cost Deviation by Origin (Company B)

Figure 27 illustrates the destination states' average total cost contribution deviation for Company B. Similar to the origin state map, Pennsylvania and Illinois still have the highest deviations for destination states. These states could be further explored and validated to ensure better routing guide compliance for future bids.

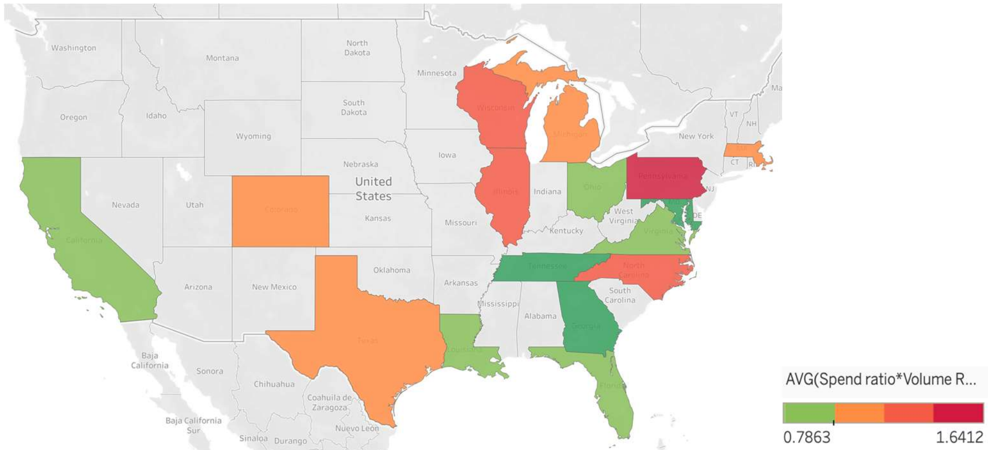


Figure 27: Total Cost Deviation by Destination (Company B)

Figure 28 illustrates the origin states' average total cost contribution deviation for Company C. All states except for California and Illinois average a higher total cost deviation from budget. This is the opposite of the results shown for Company A and B. This could be attributed to the significantly lower volume of the small shipper relative to the other shippers.

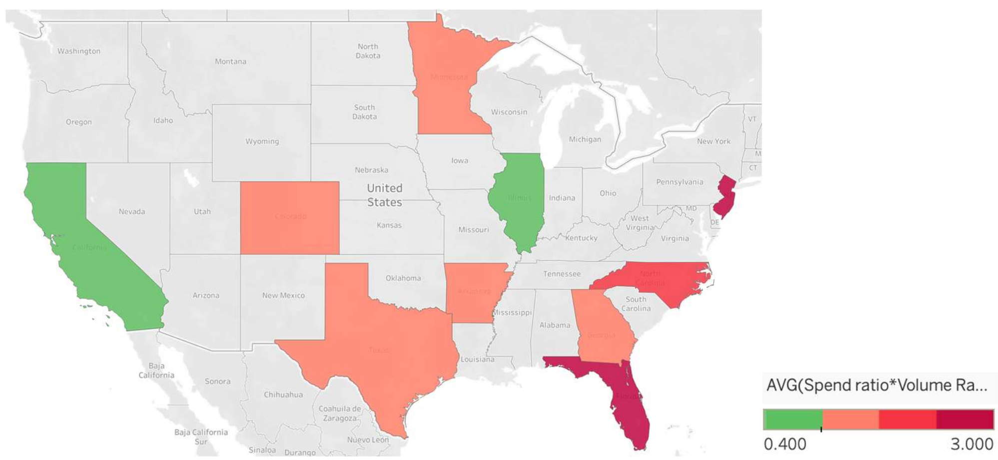


Figure 28: Total Cost Deviation by Origin (Company C)

Figure 29 illustrates the destination states' average total cost contribution deviation for Company C. New York appears to be the largest contributor to total deviation from budget for the small shipper's destination routes.

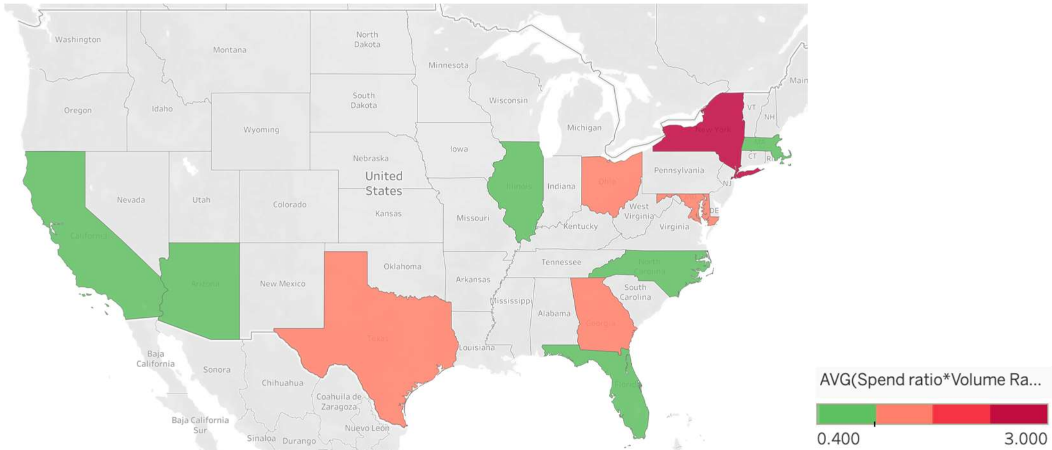


Figure 29: Total Cost Deviation by Destination (Company C)

5. CONCLUSION

This capstone evaluates the performance of our sponsor company's routing guide and compares performance by customer size. By segmenting each load into four categories -- planned and on budget, planned and above budget, planned and under budget, and unplanned -- we evaluated whether the routing guide was successful over the course of the year. The results show that for each of the three shippers, the percentage of planned and on budget loads was stable at around 80%. Overall, our sponsor company's routing guide is strong.

Next, we determine the impact of independent variables such as distance, lane volume, lane freight types, and origin/destination on spend, volume, and total cost contribution deviation from budget. As expected, factors such as dry and frozen loads either decrease or increase deviation, respectively. Therefore, shippers can incorporate contingencies in the annual bid process. For example, a shipper could mark up frozen lanes in terms of price to account for the increase deviation from budget. Furthermore, origin and destination variables do play a role in budget deviations depending on the shipper size. The average cost deviation from budget for origin and destinations are highlighted in Chapter 4 by mapping the regions with highest budget overruns.

Finally, we gained a better understanding of how volume contributes to budget overruns. Our research found that volume tends to be the dominating contributor for budget overruns. Volume on a lane correlates to the variability in loads per week. Low volume lanes have higher variability in loads per week. This could potentially result in a higher percentage of loads in the routing guide being rejected by the primary carrier because of inconsistent demand, which would lead to a budget overrun when secondary or spot market carriers execute on that specific load.

5.1 Management Insights

In this section, we provide insights our sponsor company and their shippers could proactively take to reduce their costs and improve routing guide compliance based on our research findings.

5.1.1 Incorporate Budget Deviations into Refrigerated & Frozen Lanes Bid

The research found that refrigerated and frozen lanes result in higher spend, volume, and total cost deviations from budget. Although the fixed line-haul rate includes the cost of the specific truck configuration, these lanes typically led to higher costs than budgeted. Hence, our sponsor company could evaluate whether a premium should be added to the contract rate set with the shippers to cover for this deviation above budget for future bids.

5.1.2 Regional Deviations by State

Another finding is that key origin and destination states have different impacts on total cost deviations from budget. Specific origins average higher deviations from the budget, in particular California and a number of East Coast states. In addition, origin and destination impacts differ even for the same shipper. Our sponsor company could incorporate our findings for these specific states and ensure contracted line-haul rates include price mark ups depending on the route originates or ends in a high deviation region.

5.1.3 Load Variability by Lane Volume

Our research analyzed the impact of cadence, which is the percentage of weeks that loads occur on a lane, and coefficient of variation, which plots the relationship between lane volume and average variability of loads per week. While high volume lanes tend to have lower variability because of consistent demand, low volume lanes suffer from higher variability. Since carriers tend to accept loads that are convenient and are a good network fit, higher variability lanes may result in lower primary carrier acceptance which leads to budget overruns. The impact of lane variability is worthy to be explored and segmented further to evaluate if loads can be standardized to reduce variability and randomness. The data analysis in Chapter 4 plots the distribution of lanes with varying impact by shipper that could be addressed for future bids referred to as "Coefficient of Variation by Lane" (Figures 22, 24, 26).

5.1.4 Volume Contributes to Higher Budget Deviations than Price

The research found that volume is a larger contributor to missed budgets than price. We dissected the performance of our sponsor company's routing guide by analyzing spend, volume, and total cost deviations from budget by impact of independent variables. We evaluate how cadence, CV, and lane volume per shipper size behave relative to the planned budget. The results concluded that unplanned lanes tend to be low cadence and volume with high variability that led to budget overruns. Our research has not determined causation of unplanned lane volume but we are able to characterize its behavior. While our sponsor company only has control of spend, they can highlight volume deviations to their shippers in case of budget overruns in the future. Our recommendation is that our sponsor company analyze the historical performance in order to improve their routing guide performance in the next cycle.

5.2 Further Research

Our research provides insight into how our sponsor company can improve routing guide compliance. However, further research is warranted in a few areas such as predicting unplanned lanes based on load characteristics. Although our logistic regression model was inconclusive, this is likely related to the independent variables chosen for this capstone project. A critical factor to improve routing guide compliance is to predict and incorporate unplanned lanes (i.e. those not bid but materialize in the transactions) to avoid unexpected costs throughout the year.

Additionally, analysis of lane characteristics in terms of cadence, CV, and volume provided insights that unplanned lanes tend to be low cadence and volume with high variability. However, an interesting insight in the future would be to measure is unplanned lane costs compared to market rate.

Finally, our research concluded that factors such as origin and destination and route combinations impact the total deviation from budget which shows some regional sensitivity. Based on the results of this capstone, optimization models could be created to identify if there are alternatives to specific route combinations to reduce budget deviations. The research also evaluated how cadence and variability per load is correlated with lane volume. We concluded that low volume lanes tend to have higher variability. Future analysis could be conducted to standardize low volume loads per week. Potentially optimizing the number of loads carried out per week to reduce the variability and maintain a high primary carrier acceptance.

References

Aemireddy, N., & Yuan, X. (2019). *Root Cause Analysis and Impact of Unplanned Procurement on Truckload Transportation Costs*. Retrieved from https://dspace.mit.edu/handle/1721.1/121317

American Trucking Association (2018). American Trucking Trends 2018. Washington DC: American Trucking Association.

Caplice, C. G. (1996). An optimization based bidding process: a new framework for shipper-carrier relationships (Thesis). Massachusetts Institute of Technology.

Caplice, C. and Sheffi, Y. (2003). Optimization-based Procurement for Transportation Services. Journal of Business Logistics, Vol 24, No. 2.

Caplice, C. (2009). Electronic Markets for Truckload Transportation. *Production and Operations Management*, *16*(4), 423–436. https://doi.org/10.1111/j.1937-5956.2007.tb00270.x

Caplice, C. (2009). Electronic Markets for Truckload Transportation. Production and Operations Management, 16(4), 423–436.

Caplice, C., & Sheffi, Y. (2013). *Combinatorial Auctions for Truckload Transportation*. MIT Press Scholarship Online. doi: https://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262033428.001.0001/upso-9780262033428-chapter-22

Kim, Yoo. J. (2013). Analysis of truckload prices and rejection rates (Thesis, Massachusetts Institute of Technology).

LaGore, R. (2019). Freight Contract Rates vs Spot Rates - Comprehensive Guide. Retrieved November 10, 2019, from https://blog.intekfreight-logistics.com/freight-contract-rates-vs-spot-rates-comprehensive-guide.

Pickett, C. (2018, May 4). Navigating the US truckload capacity cycle: Where are freight rates headed and why? Retrieved February 14, 2020, from https://www.henrystewartpublications.com/sites/default/files/Pickett.pdf

Ruriani, D. (2007, January 1). Creating an Inbound Routing Guide. Retrieved April 15, 2020, from https://www.inboundlogistics.com/cms/article/creating-an-inbound-routing-guide/

Shmueli, G., Patel, N. R., Bruce, P. C., & Torgo, L. (2017). Data Mining for Business Intelligence: Concepts, Techniques, and Applications in R.

Soskin, R. (2018). The Vicious Cycle: Spot Market Volatility and Primary Tender Acceptance. Retrieved November 12, 2019, from https://convoy.com/blog/the-vicious-cycle-spot-market-volatility-and-primary-tender-acceptance/.

APPENDIX A:

Table 1: Deleted Columns from Award File

Column Name	Description
Go Live Date	Date/time contract awards take effect
Requirements	Specification on truck/loading

Table 2: Deleted Columns from Transaction File

Column Name	Description
Load Status	Confirmation of completion
Load Created Date	Date/time load was entered into system
Drop Actual Date	Date/time load was dropped off by carrier
Pick Actual Date	Date/time load was picked up by carrier
Carrier Closed Date	Date/time load was closed in system
Tender Date	Date carrier scheduled to pickup
Weight	Weight of pallet
Pieces	Number of pieces in load
Date Changed	Date/time load was changed
Fuel Surcharge	Cost of fuel per load
Detention Origin Charge	Charge for wait time at origin
Detention Destination Charge	Charge for wait time at destination
Detention Charge	Charge to hold shipment
Misc. Charge	Miscellaneous charges
Other Charges	Other related charges
Currency Exchange	Currency exchange charge
Dunnage	Cost of materials (packaging)
DC Shuttle Stop	Distribution center stops
Layover	Time stopped
Drop Trailer	Number of drop trailers
TONU	Tonnage
Deadhead	Truck with empty trailer
Stopoff	Number of stops
Redelivery	Redelivery required
Team Service	Number of team members
Overweight	Pounds over weight
Loading	Loading time
Unloading	Unloading time

Pallet Exchange	Pallet exchanges
Washout	N/A
Out of Route Miles	Number of miles off route
HazmatFr	Hazmat required
Detention NC	Detention at NC
Trailer Detention	Number of trailer layovers
Trailer Detention Origin	Trailer layover at origin
Trailer Detention Destination	Trailer layover at destination
Incremental Freight	Incremental freight costs
Excess Transit Costs	Excess stops costs
Intermodal Detention	Number of intermodal detentions

APPENDIX B: SPEND DEVIATION MODEL FULL RESULTS

Company A Results:

Term	Estimate	Std Error	t Ratio	Prob>|t|
DPA	12.4265	5.605275	2.22	0.0267
Temperature [DRY]	-1.93861	0.922989	-2.1	0.0358
Distance	0.002561	0.001314	1.95	0.0514
IL	-5.70788	2.95352	-1.93	0.0534
MI	-7.03061	3.860663	-1.82	0.0687
IN	-6.27004	3.709271	-1.69	0.0911
DIl	8.69294	5.422403	1.6	0.109
PA	-6.26608	3.954213	-1.58	0.1132
OH	-6.28464	3.97333	-1.58	0.1139
MN	-7.10436	4.492472	-1.58	0.1139
KY	-6.27958	3.982217	-1.58	0.115
OK	-5.96741	3.843293	-1.55	0.1206
WI	-6.84671	4.537823	-1.51	0.1315
FL	-6.25228	4.338702	-1.44	0.1497
IA	-5.68362	3.945146	-1.44	0.1498
TX	-4.77802	3.629393	-1.32	0.1882
TN	-5.3652	4.079349	-1.32	0.1886
NC	-4.58116	3.557165	-1.29	0.1979
Temperature[FROZEN]	1.296697	1.072099	1.21	0.2266
NY	-5.8393	4.990724	-1.17	0.2421
MO	-6.37325	5.534576	-1.15	0.2496
SC	-6.48802	5.893531	-1.1	0.2711

MA	-7.4739	7.167543	-1.04	0.2972
NE	-8.46581	8.132098	-1.04	0.298
DOH	5.524931	5.700744	0.97	0.3326
NJ	-6.08248	6.683693	-0.91	0.3629
VA	-5.66016	6.343214	-0.89	0.3723
DMI	5.467667	6.150111	0.89	0.3741
SD	-5.17652	5.828721	-0.89	0.3746
DNY	4.967394	5.76232	0.86	0.3888
DCT	5.023167	5.886669	0.85	0.3936
DIA	5.109694	6.251113	0.82	0.4138
ABC_C	-1.776	2.204727	-0.81	0.4206
DWI	4.83991	6.137943	0.79	0.4305
AR	-5.58666	7.331044	-0.76	0.4461
DVA	4.649551	6.17096	0.75	0.4513
GA	2.596857	3.474469	0.75	0.4549
MD	-12.3964	16.82494	-0.74	0.4613
KS	-3.94445	5.444974	-0.72	0.4689
DGA	3.975225	5.496326	0.72	0.4696
AL	-4.83729	6.705892	-0.72	0.4708
DKS	4.404048	6.147022	0.72	0.4738
CT	-6.24436	8.781955	-0.71	0.4771
DMN	4.566416	6.44175	0.71	0.4785
DTN	4.183218	6.120005	0.68	0.4943
WA	-4.98505	7.665157	-0.65	0.5155
DNC	3.893239	6.165331	0.63	0.5278
ABC_A	-1.15167	1.947889	-0.59	0.5544
DAR	3.755415	6.412939	0.59	0.5582
DTX	3.268077	5.594801	0.58	0.5592
DSC	3.450074	6.121178	0.56	0.5731
OR	-4.35722	9.325102	-0.47	0.6404
DFL	2.590079	5.699397	0.45	0.6496
UT	-3.53903	7.944523	-0.45	0.656
MS	-3.94554	9.034323	-0.44	0.6624
DLA	2.725964	6.246292	0.44	0.6626
DCO	2.516002	6.229812	0.4	0.6864
CO	-5.14153	13.09128	-0.39	0.6945
NH	-6.06413	16.86244	-0.36	0.7192
ID	-2.90879	8.506256	-0.34	0.7324
ME	-8.95075	28.70139	-0.31	0.7552

DCA	1.505198	5.259331	0.29	0.7748
DE	-3.70424	14.56987	-0.25	0.7993
LA	-2.17147	9.479473	-0.23	0.8188
DAZ	1.143218	6.368251	0.18	0.8575
Intercept	0.879061	5.101832	0.17	0.8632
DUT	1.046646	6.433222	0.16	0.8708
NM	-1.58462	10.29976	-0.15	0.8777
NV	-0.4834	8.536365	-0.06	0.9548
AZ	-0.55461	13.02706	-0.04	0.966
DWA	0.139725	6.040421	0.02	0.9815

Company B Results:

Term	Estimate	Std Error	t Ratio	P-value
AR	0.079417	0.022032	3.6	0.0036
MS	0.080258	0.026056	3.08	0.0095
D_MD	0.03268	0.02503	1.31	0.2162
D_FL	0.074549	0.065772	1.13	0.2792
D_IL	0.048153	0.043989	1.09	0.2952
D_NC	0.067061	0.062045	1.08	0.301
PA	0.071486	0.067155	1.06	0.3081
D_GA	0.064485	0.060911	1.06	0.3106
D_CA	0.108813	0.104291	1.04	0.3173
D_TX	0.07106	0.068263	1.04	0.3184
WA	0.127408	0.123308	1.03	0.3219
Distance	0.000	0.000	-1.02	0.3279
D_CO	0.098617	0.099232	0.99	0.3399
NV	-0.03397	0.038766	-0.88	0.3981
KS	0.032639	0.037552	0.87	0.4018
OH	2.74E-02	3.40E-02	0.8	0.4366
MN	0.031862	0.040648	0.78	0.4483
KY	0.015176	0.021441	0.71	0.4926
MD	0.0389	0.055424	0.7	0.4962
IL	0.02287	0.03968	0.58	0.575
IN	0.017916	0.034488	0.52	0.6129
D_MA	0.012547	0.02702	0.46	0.6507
GA	0.007002	0.0151	0.46	0.6511
D_WI	-0.00983	0.023806	-0.41	0.6869
ABC_A	0.004642	0.012831	0.36	0.7238
D_VA	0.007304	0.021585	0.34	0.7409

OR	-0.00613	0.020646	-0.3	0.7717
Temperature [DRY]	-0.00092	0.005055	-0.18	0.8583
D_TN	-0.00866	0.075864	-0.11	0.911
D_OH	-0.00207	0.020411	-0.1	0.9208
Temperature [FROZEN]	-0.00034	0.005841	-0.06	0.9547
Intercept	0.92556	0.075523	12.26	<.0001

Company C Results:

Term	**Estimate**	**Std Error**	**t Ratio**	**P-value**
ABC_A	0.06023	0.01650	3.65	0.0053
CA	0.05700	0.02819	2.02	0.0739
NJ	0.03155	0.02208	1.43	0.1869
DCA	-0.03405	0.02490	-1.37	0.2046
DTX	0.03239	0.02508	1.29	0.2287
DAZ	-0.03651	0.02970	-1.23	0.2501
DIL	0.02524	0.02067	1.22	0.2532
DMD	0.01650	0.01524	1.08	0.307
Distance	0.00002	0.00002	1.08	0.3088
DGA	0.01644	0.01700	0.97	0.3589
GA	0.00718	0.00828	0.87	0.4085
CO	-0.01737	0.02063	-0.84	0.4217
MN	-0.01444	0.01839	-0.79	0.4524
Temperature [FROZEN]	0.00473	0.00616	0.77	0.4617
Temperature [DRY]	-0.00585	0.00948	-0.62	0.5523
TX	-0.01100	0.02341	-0.47	0.6497
DFL	-0.00897	0.02380	-0.38	0.715
AR	0.00551	0.02063	0.27	0.7955
DOH	-0.00230	0.01663	-0.14	0.8932
FL	0.00108	0.01626	0.07	0.9484
NC	0.00087	0.01314	0.07	0.9489
DNY	-0.00086	0.01526	-0.06	0.9565
DMA	0.00080	0.01704	0.05	0.9636

APPENDIX C: VOLUME RATIO MODEL FULL RESULTS

Company A Results:

Term	**Estimate**	**Std Error**	**t Ratio**	**Prob>|t|**

DAR	2.718537	1.085708	2.5	0.0124
DLA	2.579825	1.057497	2.44	0.0148
NY	1.751748	0.845562	2.07	0.0384
IA	1.37139	0.668334	2.05	0.0403
GA	1.106304	0.589209	1.88	0.0606
UT	2.496254	1.345188	1.86	0.0636
DTN	1.894144	1.036128	1.83	0.0677
ABC_C	-0.66334	0.377087	-1.76	0.0787
DTX	1.403964	0.947205	1.48	0.1384
DAZ	1.524966	1.078088	1.41	0.1574
OH	-0.91168	0.67359	-1.35	0.176
OK	-0.85854	0.651083	-1.32	0.1874
DCT	1.282388	0.996766	1.29	0.1984
DVA	1.283531	1.044788	1.23	0.2194
DIA	1.24562	1.058342	1.18	0.2393
MS	-1.78142	1.529733	-1.16	0.2443
Intercept	0.955939	0.864332	1.11	0.2689
Temperature [FROZEN]	0.196863	0.181521	1.08	0.2783
DOH	0.994778	0.96575	1.03	0.3031
DNY	0.960643	0.975637	0.98	0.3249
MN	-0.73566	0.761097	-0.97	0.3339
MA	-1.1661	1.213902	-0.96	0.3368
LA	-1.47907	1.605192	-0.92	0.3569
IN	0.577623	0.628827	0.92	0.3584
TX	-0.54445	0.616453	-0.88	0.3772
WI	-0.6418	0.768914	-0.83	0.404
AR	-0.97173	1.241232	-0.78	0.4338
KS	-0.71457	0.922179	-0.77	0.4385
FL	-0.56677	0.735103	-0.77	0.4408
DMI	0.800649	1.041254	0.77	0.442
DMN	0.80721	1.090565	0.74	0.4593
SD	-0.72287	0.98703	-0.73	0.464
DUT	0.795572	1.089087	0.73	0.4652
PA	-0.48524	0.670617	-0.72	0.4694
DPA	0.625753	0.94912	0.66	0.5098
NE	-0.87125	1.376935	-0.63	0.527
DSC	0.651057	1.036315	0.63	0.5299
DFL	0.601023	0.964982	0.62	0.5335
IL	-0.31007	0.501775	-0.62	0.5367

AL	-0.69114	1.135838	-0.61	0.5429
NM	-0.95832	1.743758	-0.55	0.5827
DCA	0.481766	0.891878	0.54	0.5891
NH	-1.52884	2.854981	-0.54	0.5924
MO	-0.49581	0.937386	-0.53	0.5969
AZ	-1.1271	2.205616	-0.51	0.6094
ABC_A	-0.16779	0.329804	-0.51	0.611
OR	-0.7886	1.578773	-0.5	0.6175
NC	-0.29311	0.603106	-0.49	0.627
DE	-1.19096	2.46678	-0.48	0.6293
SC	0.475013	0.998369	0.48	0.6343
DKS	0.485962	1.040682	0.47	0.6406
CT	-0.54734	1.487237	-0.37	0.7129
TN	0.237998	0.691373	0.34	0.7307
NV	-0.48937	1.44537	-0.34	0.735
MI	-0.22119	0.65439	-0.34	0.7354
DGA	0.313999	0.930581	0.34	0.7358
Temperature [DRY]	0.051018	0.156308	0.33	0.7442
VA	0.310252	1.074423	0.29	0.7728
Distance	5.89E-05	0.000223	0.26	0.792
KY	-0.14155	0.68058	-0.21	0.8353
ID	-0.27054	1.440324	-0.19	0.851
WA	0.240297	1.29775	0.19	0.8531
DNC	0.176061	1.043841	0.17	0.8661
DWA	0.166534	1.0226	0.16	0.8706
DIl	0.147191	0.91839	0.16	0.8727
CO	-0.32609	2.21647	-0.15	0.8831
ME	-0.47718	4.858904	-0.1	0.9218
DWI	0.063286	1.039165	0.06	0.9514
DCO	0.040472	1.05468	0.04	0.9694
NJ	0.019717	1.132096	0.02	0.9861
MD	-0.03987	2.848701	-0.01	0.9888

Company B Results:

Term	Estimate	Std Error	t Ratio	P-value
ABC_A	0.6787	0.4174	1.63	0.1299
AR	-0.9972	0.7166	-1.39	0.1894
D_OH	-0.9140	0.6639	-1.38	0.1937

OR	-0.8925	0.6715	-1.33	0.2086
Temperature [FROZEN]	0.2029	0.1900	1.07	0.3066
D_IL	1.3739	1.4308	0.96	0.3559
PA	1.9119	2.1844	0.88	0.3986
IL	1.0136	1.2907	0.79	0.4475
D_TN	1.5684	2.4676	0.64	0.537
Distance	0.0005	0.0009	0.57	0.5765
D_LA	1.5094	2.6748	0.56	0.5829
D_NC	1.0770	2.0182	0.53	0.6033
D_TX	1.0107	2.2204	0.46	0.6571
D_CO	1.3979	3.2277	0.43	0.6726
MS	-0.3354	0.8475	-0.4	0.6993
GA	0.1849	0.4911	0.38	0.7131
MD	0.6238	1.8028	0.35	0.7353
Temperature [DRY]	-0.0545	0.1644	-0.33	0.7459
Intercept	-0.8001	2.4565	-0.33	0.7503
KY	0.2186	0.6974	0.31	0.7594
D_GA	0.6202	1.9813	0.31	0.7596
D_CA	1.0408	3.3923	0.31	0.7642
D_WI	-0.2134	0.7743	-0.28	0.7876
IN	-0.3006	1.1218	-0.27	0.7933
MN	-0.3313	1.3222	-0.25	0.8064
D_MA	0.2094	0.8789	0.24	0.8157
D_FL	0.4867	2.1394	0.23	0.8239
OH	0.2225	1.1059	0.2	0.8439
WA	0.5494	4.0108	0.14	0.8933
D_VA	0.0699	0.7021	0.1	0.9224
KS	-0.0586	1.2214	-0.05	0.9625
NV	0.0359	1.2609	0.03	0.9777
D_MD	0.0083	0.8142	0.01	0.992

Company C Results:

Term	**Estimate**	**Std Error**	**t Ratio**	**P-value**
Distance	0.00171	0.00074	2.31	0.046
FL	1.06375	0.55084	1.93	0.0855
CO	-1.25039	0.69907	-1.79	0.1073
DFL	-1.41457	0.80650	-1.75	0.1133
DGA	0.97897	0.57596	1.7	0.1234
DTX	1.40553	0.84976	1.65	0.1325

NC	0.73012	0.44517	1.64	0.1354
DIL	0.96548	0.70039	1.38	0.2013
CA	1.18735	0.95511	1.24	0.2452
DOH	0.60949	0.56339	1.08	0.3075
AR	-0.73140	0.69901	-1.05	0.3227
Temperature [FROZEN]	0.21438	0.20861	1.03	0.3309
MN	-0.55617	0.62316	-0.89	0.3954
DMA	-0.46607	0.57718	-0.81	0.4402
ABC_A	0.37760	0.55918	0.68	0.5165
Intercept	-0.43663	0.78240	-0.56	0.5904
DCA	-0.42560	0.84349	-0.5	0.626
DAZ	-0.38943	1.00615	-0.39	0.7077
GA	0.10450	0.28057	0.37	0.7182
DNY	-0.18616	0.51692	-0.36	0.7271
Temperature [DRY]	0.03219	0.32127	0.1	0.9224
DMD	0.04270	0.51624	0.08	0.9359
TX	0.00998	0.79317	0.01	0.9902
NJ	15.72	0.75	21.01	<.0001

APPENDIX D: TOTAL COST CONTRIBUTION RATIO MODEL FULL RESULTS

Company A Results:

Term	**Estimate**	**Std Error**	**t Ratio**	**Prob>|t|**
DPA	8.501207	3.791699	2.24	0.0251
Temperature [DRY]	-1.19411	0.624446	-1.91	0.056
Temperature [FROZEN]	1.220481	0.72517	1.68	0.0925
GA	3.944526	2.353869	1.68	0.0939
IL	-3.19806	2.004575	-1.6	0.1108
OH	-4.25432	2.690968	-1.58	0.114
Distance	0.001391	0.000891	1.56	0.1187
OK	-4.03723	2.601052	-1.55	0.1208
MN	-4.57029	3.040556	-1.5	0.133
MI	-3.90881	2.614266	-1.5	0.135
PA	-3.83807	2.679089	-1.43	0.1521
WI	-4.31296	3.071784	-1.4	0.1604
KY	-3.43304	2.718891	-1.26	0.2068
DIl	4.631066	3.668936	1.26	0.207
FL	-3.70163	2.936711	-1.26	0.2076

ABC_C	-1.89871	1.50645	-1.26	0.2077
TX	-2.84154	2.462708	-1.15	0.2487
MA	-5.33544	4.849492	-1.1	0.2714
IN	-2.64191	2.512141	-1.05	0.2931
DAR	4.487751	4.337365	1.03	0.3009
NE	-5.66321	5.500805	-1.03	0.3033
DCT	4.07E+00	3.982042	1.02	0.307
NC	-2.44969	2.409386	-1.02	0.3094
DOH	3.849646	3.858137	1	0.3185
MO	-3.58648	3.744822	-0.96	0.3383
DTN	3.949375	4.139294	0.95	0.3401
DIA	3.940924	4.228036	0.93	0.3514
TN	-2.45118	2.762009	-0.89	0.3749
DMI	3.687583	4.15977	0.89	0.3755
DLA	3.725912	4.22466	0.88	0.3779
DNY	3.430321	3.897633	0.88	0.3789
DVA	3.597368	4.173892	0.86	0.3889
AR	-4.08668	4.958676	-0.82	0.4099
SD	-3.22816	3.94315	-0.82	0.4131
DTX	2.931986	3.784049	0.77	0.4385
SC	-3.05708	3.98845	-0.77	0.4435
ABC_A	-0.97269	1.317554	-0.74	0.4604
DMN	3.19454	4.356769	0.73	0.4635
NJ	-3.19579	4.522684	-0.71	0.4799
MD	-8.02805	11.38046	-0.71	0.4806
KS	-2.49696	3.684073	-0.68	0.498
DKS	2.739768	4.157488	0.66	0.51
AL	-2.95106	4.537632	-0.65	0.5155
CT	-3.78625	5.941457	-0.64	0.524
DWI	2.588809	4.151427	0.62	0.533
MS	-3.76809	6.111227	-0.62	0.5376
DGA	2.222077	3.717635	0.6	0.5501
IA	-1.52695	2.669969	-0.57	0.5674
DSC	2.144857	4.140042	0.52	0.6045
OR	-3.22159	6.30714	-0.51	0.6096
VA	-2.18159	4.29228	-0.51	0.6113
DAZ	2.153729	4.306923	0.5	0.6171
WA	-2.49316	5.184462	-0.48	0.6306
DNC	1.959169	4.170106	0.47	0.6385

NY	-1.55026	3.37799	-0.46	0.6463
NH	-5.07596	11.40555	-0.45	0.6563
DFL	1.636501	3.855068	0.42	0.6712
DCA	1.324249	3.56302	0.37	0.7102
LA	-2.224	6.412685	-0.35	0.7288
CO	-2.83978	8.854718	-0.32	0.7485
DCO	1.305988	4.213406	0.31	0.7566
ID	-1.7764	5.754042	-0.31	0.7576
DUT	1.280644	4.350864	0.29	0.7685
DE	-2.75392	9.854697	-0.28	0.7799
ME	-5.36595	19.41114	-0.28	0.7822
Intercept	0.936157	3.452976	0.27	0.7863
NM	-1.55106	6.966251	-0.22	0.8238
AZ	-1.297	8.811354	-0.15	0.883
UT	0.715563	5.373978	0.13	0.8941
NV	-0.62547	5.774199	-0.11	0.9138
DWA	0.278496	4.085249	0.07	0.9457

Company B Results:

Term	**Estimate**	**Std Error**	**t Ratio**	**P-value**
ABC_A	0.68245	0.41648	1.64	0.1272
D_OH	-0.91426	0.66249	-1.38	0.1927
OR	-0.89598	0.67013	-1.34	0.206
AR	-0.92702	0.71512	-1.3	0.2192
Temperature [FROZEN]	0.20159	0.18959	1.06	0.3086
D_IL	1.40902	1.42780	0.99	0.3432
PA	1.96737	2.17973	0.9	0.3845
IL	1.03156	1.28795	0.8	0.4387
D_TN	1.55441	2.46240	0.63	0.5397
D_LA	1.50089	2.66914	0.56	0.5843
D_NC	1.12992	2.01388	0.56	0.5851
Distance	0.00047	0.00086	0.55	0.5949
D_TX	1.06590	2.21568	0.48	0.6391
D_CO	1.47728	3.22088	0.46	0.6547
GA	0.19068	0.49010	0.39	0.7041
MD	0.65300	1.79897	0.36	0.7229
Intercept	-0.85819	2.45132	-0.35	0.7323
D_GA	0.67091	1.97706	0.34	0.7402

Temperature [DRY]	-0.05496	0.16409	-0.33	0.7435
D_CA	1.13015	3.38510	0.33	0.7442
KY	0.23069	0.69593	0.33	0.746
MS	-0.26646	0.84572	-0.32	0.7581
D_WI	-0.22241	0.77270	-0.29	0.7784
D_FL	0.54693	2.13483	0.26	0.8021
IN	-0.28502	1.11942	-0.25	0.8033
D_MA	0.21416	0.87703	0.24	0.8112
MN	-0.30371	1.31937	-0.23	0.8218
OH	0.24735	1.10358	0.22	0.8264
WA	0.65691	4.00234	0.16	0.8724
D_VA	0.07594	0.70062	0.11	0.9155
D_MD	0.03605	0.81244	0.04	0.9653
KS	-0.03010	1.21886	-0.02	0.9807
NV	0.00503	1.25827	0	0.9969

Company C Results:

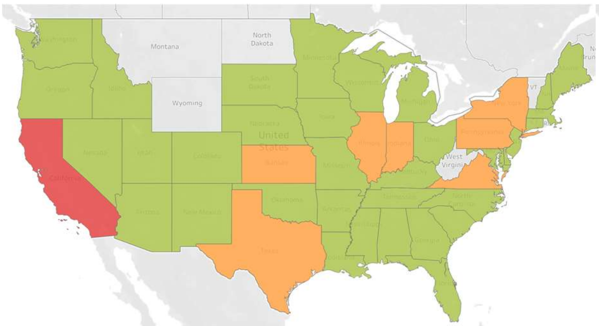
Term	**Estimate**	**Std Error**	**t Ratio**	**P-value**
Intercept	-0.08413	0.02825	-2.98	0.0155
ABC_A	0.05840	0.02019	2.89	0.0178
CA	0.05816	0.03448	1.69	0.1259
DCA	-0.04084	0.03045	-1.34	0.2128
NJ	0.03311	0.02701	1.23	0.2514
DAZ	-0.04226	0.03633	-1.16	0.2746
DMD	0.01816	0.01864	0.97	0.3554
Temperature [DRY]	-0.01032	0.01160	-0.89	0.3966
DTX	0.02514	0.03068	0.82	0.4337
DIL	0.02006	0.02529	0.79	0.4479
GA	0.00777	0.01013	0.77	0.4628
DGA	0.01454	0.02079	0.7	0.502
AR	0.01672	0.02524	0.66	0.5243
Temperature [FROZEN]	0.00480	0.00753	0.64	0.5398
Distance	0.00001	0.00003	0.51	0.6215
NC	-0.00797	0.01607	-0.5	0.632
MN	-0.00994	0.02250	-0.44	0.6691
DOH	-0.00890	0.02034	-0.44	0.672
CO	-0.00881	0.02524	-0.35	0.7352
TX	-0.00840	0.02864	-0.29	0.7759
FL	-0.00348	0.01989	-0.17	0.865

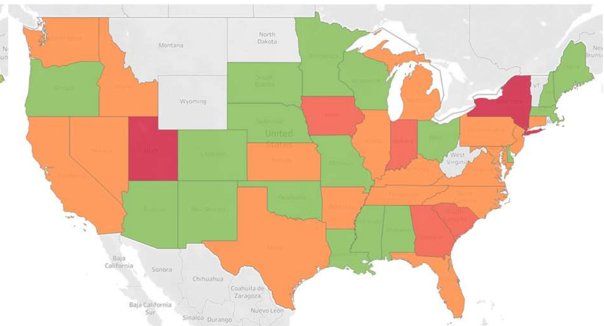
DMA	0.00268	0.02084	0.13	0.9004
DNY	0.00051	0.01866	0.03	0.9789
DFL	0.00068	0.02912	0.02	0.982
Intercept	-0.08413	0.02825	-2.98	0.0155
ABC_A	0.05840	0.02019	2.89	0.0178
CA	0.05816	0.03448	1.69	0.1259
DCA	-0.04084	0.03045	-1.34	0.2128
NJ	0.03311	0.02701	1.23	0.2514
DAZ	-0.04226	0.03633	-1.16	0.2746
DMD	0.01816	0.01864	0.97	0.3554
Temperature [DRY]	-0.01032	0.01160	-0.89	0.3966
DTX	0.02514	0.03068	0.82	0.4337
DIL	0.02006	0.02529	0.79	0.4479
GA	0.00777	0.01013	0.77	0.4628
DGA	0.01454	0.02079	0.7	0.502
AR	0.01672	0.02524	0.66	0.5243
Temperature [FROZEN]	0.00480	0.00753	0.64	0.5398
Distance	0.00001	0.00003	0.51	0.6215
NC	-0.00797	0.01607	-0.5	0.632
MN	-0.00994	0.02250	-0.44	0.6691
DOH	-0.00890	0.02034	-0.44	0.672
CO	-0.00881	0.02524	-0.35	0.7352
TX	-0.00840	0.02864	-0.29	0.7759
FL	-0.00348	0.01989	-0.17	0.865
DMA	0.00268	0.02084	0.13	0.9004
DNY	0.00051	0.01866	0.03	0.9789
DFL	0.00068	0.02912	0.02	0.982

APPENDIX E: Spend & Volume Deviations by Origin & Destination States

Company A Deviation by Origin State

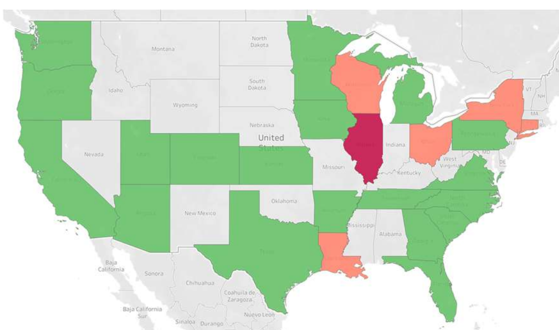
Spend: **Volume:**

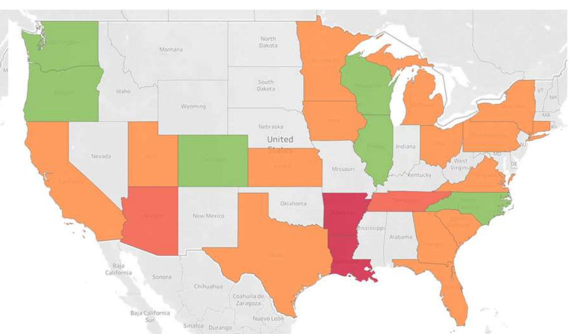




Company A Deviation by Destination State

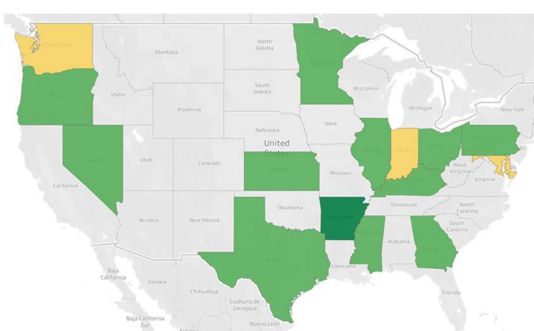
Spend: **Volume:**

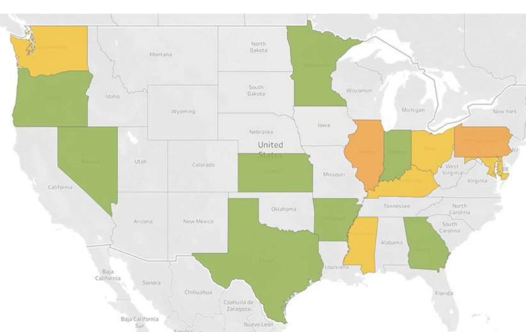




Company B Deviation by Origin State

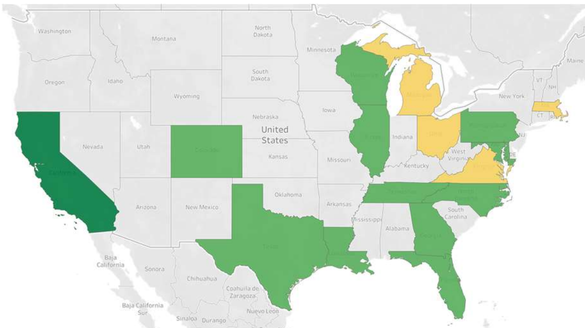
Spend: **Volume:**



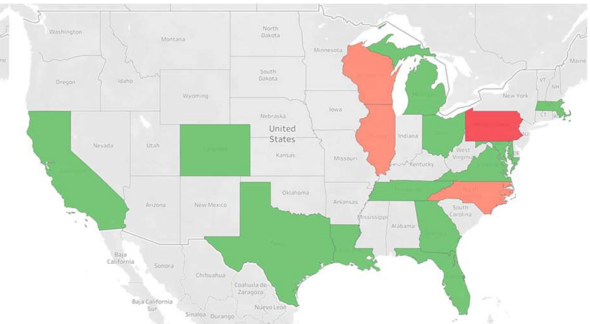


Company B Deviation by Destination State

Spend:

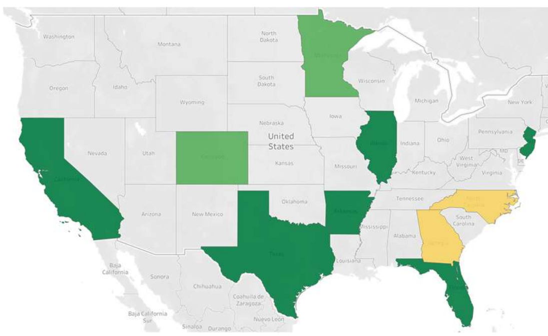


Volume:

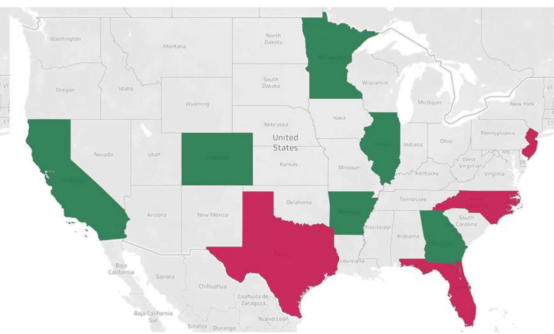


Company C Deviation by Origin State

Spend:

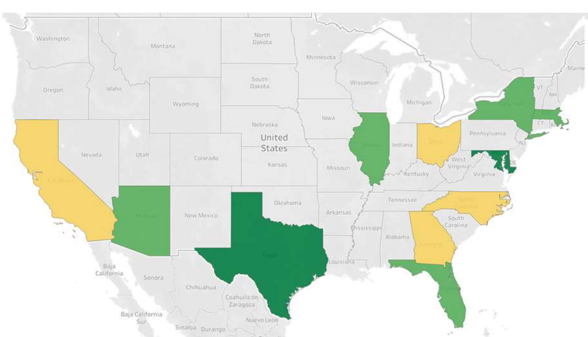


Volume:



Company C Deviation by Destination State

Spend:



Volume: