Alternative Freight Contracts: Data-driven Design Under Uncertainty

by

Angela Josephine Acocella

B.S., Rensselaer Polytechnic Institute (2012)

S.M., Massachusetts Institute of Technology (2015)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2022

Author . . . .

Department of Civil and Environmental Engineering

December 22, 2021

Certified by . . . .

Yossi Sheffi

Elisha Gray II Professor, Engineering Systems

Professor, Civil and Environmental Engineering, MIT

Thesis Supervisor

Certified by . . . .

Chris Caplice

Senior Research Scientist

Executive Director, Center for Transportation & Logistics, MIT

Thesis Supervisor

Accepted by . . . .

Colette L. Heald

Professor of Civil and Environmental Engineering

Chair, Graduate Program Committee

Alternative Freight Contracts: Data-driven Design Under Uncertainty

by

Angela Josephine Acocella

Submitted to the Department of Civil and Environmental Engineering
on December 22, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We demonstrate how firms (shippers) should incorporate market uncertainty into their strategic procurement of truckload (TL) transportation services from suppliers (carriers). These direct point-to-point TL movements are a major segment of total freight volumes for shippers. As such, firms emphasize the need to form strong relationships with their carriers. This becomes difficult due to the two-sided, non-binding nature of TL contacts; shippers are not required to offer the stated contract volumes - causing uncertainty and cost escalations for carriers - and carriers are not required to accept the offered freight - forcing shippers to rely on higher priced backup providers. These costs intensify when freight markets cycle between periods of over and under supply; soft and tight markets, respectively.

We propose an empirical modeling approach utilizing large, detailed microeconomic data sets to help shippers and carriers form better contractual relationships given present uncertainties. Previous research on TL procurement and operations has predominantly taken analytical approaches due to limited availability of real-world industry data. Thus, it has been limited in addressing the three sources of uncertainty: (1) demand from shippers, (2) capacity supplied by carriers, and (3) the fluctuations in the freight market that shift the power back and forth between parties.

This thesis provides six main contributions. First, we explicitly incorporate the three sources of uncertainty into shippers' TL transportation contracting decisions by developing empirical behavioral models. Second, we confirm when the underlying structure of the freight markets change and impact carrier behaviors. Third, we identify which actions shippers can take to encourage contracted carriers to maintain high freight acceptance rates during tight markets. Fourth, we quantify carriers' contract price stickiness and identify which segments of a shipper's network and carrier base are most promising for a market-based contract to mitigate the negative effects of market fluctuations on cost and performance levels. Fifth, we determine the market-based contract designs that result in a Pareto improvement for both shippers and carriers over the traditional long-term, fixed-price contract and quantify the expected benefit to both sides. Finally, we measure the causal effect of index-based contracts on the service levels and costs shippers experience.

Thesis Supervisor: Yossi Sheffi
Title: Elisha Gray II Professor, Engineering Systems
Professor, Civil and Environmental Engineering, MIT

Thesis Supervisor: Chris Caplice
Title: Senior Research Scientist
Executive Director, Center for Transportation & Logistics, MIT

Acknowledgments

I am eternally grateful to the village of people whose support made all of this possible. First and foremost, I cannot thank my advisors, Prof. Yossi Sheffi and Dr. Chris Caplice, enough for everything they have done for me during my PhD. Yossi, for even taking a chance on me; for showing patience while still pushing me as I grew. From the very beginning you have been my biggest supporter in the room - even more than myself some days. You allowed me the independence to find myself in an academic and research setting and have become one of my most valued mentors. Chris, for every opportunity to share my research and start to make a name for myself. You've taught me everything I know about transportation and I'm still learning. More than that though, you've shown me how to be a great teacher and advisor myself. I look forward to continuing to learn from you.

I also want to thank my other committee members. Prof. Jan Fransoo for taking interest in my research from our first meeting, sharing your time and thoughts that helped me think about things from a different perspective. And Prof. Cathy Wu, for your sharp questions that ensured I could clearly articulate my research contributions.

My time at MIT has been shaped by a few amazing groups of people that have become family to me. First, I am so thankful for my TPP Master's cohort turned lifetime friends and family. Second, CTL has become home during my PhD. I cherish each one of my relationships with the CTL researchers, postdocs, other PhD and Master's students, and staff. From feedback on research, to conversation by the coffee machine, beers at the Muddy's, pick-up soccer and basketball, hiking in the White Mountains, and just our daily lunches together have been some of the most fulfilling experiences of my life. And finally, a fellow PhD student in the CEE department who became one of my most important sounding boards, friends, and hopefully future colleagues.

And last but not least, I am forever thankful for my family. To my parents, your support, understanding, encouragement, and advice through this process and the many years it took to get here have made all of this possible. I'm so proud to have been able to share this moment with you both. And to my sister, for being my life-long best friend, keeping me in line when I needed it, and showing me how to be unapologetic in pursuit of your dreams.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The freight transportation function is a major source of cost and uncertainty for most firms. In fact, at $1.1 trillion, transportation represent 68% of total logistics costs in the US, according to the Council of Supply Chain Management Professionals [13]. However, mounting availability of data and advanced analytics capabilities have helped shippers (the firms buying transportation) make more informed decisions on how to cover their transportation needs from suppliers (carriers). Still, due to the complexities of a highly fragmented transportation market and price and demand volatility, these decisions remain significant challenges.

This holds particularly true in the over-the-road trucking sector – an industry that totaled $794 billion in 2020 and accounts for 73% of freight hauled in the US [13]. The relationships between shippers and their motor carriers take on many forms; from in-house, private trucking fleets, to single-transaction exchanges, and many varieties of long- and short-term contractual relationships in between.

These relationships have developed since deregulation of the trucking industry in 1980. Following the Motor Carrier Act of that year, direct point-to-point - otherwise known as truckload (TL) - shipping lanes across the country would no longer be priced according to government-stipulated rates. Instead, carriers could each set a price they were willing to accept to provide their transportation services. This greatly opened up competition to the industry.

Shipper-carrier relationships have further evolved as demands on the transportation function have intensified. Shippers' distribution networks are growing in size and complexity. Drives toward resiliency and flexibility along supply chains in manufacturing processes and inventory policies with reduced lead times, end customer expectations for shortening transport times, and calls for greater visibility and track and trace functions have pushed shippers and carriers alike to consider alternative relational forms to meet these growing demands.

A unique aspect of the shipper-carrier relationship in the TL sector that has endured over the last 40 years is the two-sided non-binding contract – that is, while the price of the service is committed to by both sides, neither the supply of transportation capacity from the carrier nor demand for the services by the shipper are obligatory of the signing parties. Both sides agree to this capacity and volume flexibility because neither wants to be penalized for the uncertainties they face that they cannot control.

On one hand, the shipper may request more than, less than, or even none of the contracted volume, with no direct financial penalty. This is because forecasting the precise time and location of future demand for trucking capacity is not realistic, especially when these projections are done often months or even a year in advance. Moreover, forecasting errors are amplified when shippers' end customer demand changes over the course of the transportation contract.

On the other hand, carriers are not obligated to accept 100% of the volume that is offered or tendered to them. Responding to many shippers' demand makes committing a truck to a precise future time and location almost impossible, especially when commitments are expected to be in effect for a year or more in most cases. As a result, a carrier can enter a contract with a shipper for an expected amount of freight but variability from other shipper customers' demand, general market dynamics, or changes in the carrier's own business require the flexibility to reject loads in real-time capacity allocation decisions.

The non-binding agreements further complicate planning long-term budgets and services accurately for both parties. The problems are compounded as shippers procure transportation services from many carriers across their distribution networks and carriers serve many shippers [42, 43]. Shippers must balance trade-offs between pricing, expected transportation service levels, and their own customer demand uncertainty, while carriers face a dynamic pickup and delivery problem in which they face random demand and short lead times as their trucks move from point to point [116, 22][1].

In addition to the supply and demand uncertainty, a third source of uncertainty adds to the challenges: fluctuations in the freight market. As the overall economy cycles through expansions and contractions, so too does the transportation industry. Strong demand for goods, and resultantly transportation services, combined with limited capacity forces prices upward in what is referred to as a tight freight market. During a soft freight market, overall demand for transportation of goods goes down as do prices.

These macroeconomic factors are widely known by practitioners to impact shipper-carrier relationships, shippers' transportation costs and carriers' service levels and profitability. However, the dynamics have received almost no attention from the research community to help shippers and carriers make better decisions that mitigate the effects of the three main sources of uncertainty: demand volatility, supply availability, and market fluctuations.

1.1 Research Objective

In this thesis, we explore how shippers should incorporate the dynamics of freight market fluctuation into their transportation procurement strategies and the relationships they form with TL motor carriers. In particular, we consider (1) what shippers should do to maintain existing carrier relationships and ensure high carrier service levels during tight freight market conditions in order to reduce cost escalations, (2) when (i.e., during what market conditions) it pays off to implement these actions, (3) where within the shippers' distribution networks and carrier bases to consider novel contractual relationship forms, and (4) how to design these alternative contractual relationships.

We address these questions from an empirical perspective, utilizing uniquely detailed and comprehensive data representing strategic and operational decisions and pricing of both shippers and carriers. This is a fundamental piece of the practical and academic contributions of this thesis as we can capture the microeconomic behaviors and decisions

[1]See [161, 160] for a detailed discussion of the trade-offs and implications of contract economics and relational contracts and [39], which focuses on these topics explicitly in the trucking industry.

across time and geographic regions in ways that have not been done before.

In the remainder of this section, we expand on the key processes and participants. We motivate the importance of studying alternative shipper-carrier contractual forms to mitigate the challenges from the compounding supply, demand, and market uncertainties.

1.2 Truckload Transportation Industry Overview

The two trucking options available to shippers are less-than-truckload (LTL) and full truckload (FTL or TL). With LTL, the carrier operates a hub and spoke distribution model: smaller shipments that cannot fill an entire trailer typically from more than one shipper and multiple pick-up locations are consolidated together at a central hub, perhaps transshipped in a point-to-point movement to another consolidation terminal, and finally the carrier re-distributes them to multiple drop-off locations.

Alternatively, TL consists of direct point-to-point movements. A single carrier moves a single shipper's large shipments - typically over 10,000 pounds that fill up to an entire standardized 53-foot truck - on a single origin-destination lane. For-hire TL transportation spend totaled $306 billion in 2020, or 44% of over-the-road freight transportation [13].

The key to carriers operating a successful TL network is balancing the inflow and outflow of trucks at each node. That it, the carrier is primarily concerned with finding follow-on loads for inbound trucks. For example, a truck picking up a load from point A to point B must be able to find another subsequent loads out of B to point C (or alternatively, back to point A), where it has its next scheduled load. Otherwise, the carrier will have to reposition the truck by moving empty, or unloadad to point C. The carrier wants to minimize these "deadhead" moves because it is not paid by a shipper for the distance traveled.

Certain inbound or outbound regions of the country are more or less attractive to carriers. Areas such as the Port of Los Angeles/Long Beach and the industrial and agricultural belts of the Midwest function as freight sources; high volumes of freight originate in these regions and carriers know they can easily find follow-on loads when dropping a previous load nearby. Conversely, regions such as New England and the Florida peninsula are freight sinks; it is typically difficult for carriers to find follow-on loads. Importantly, these regional differences factor into carriers' pricing decisions. We discuss the procurement and pricing process further in subsection 1.4.1.

1.3 Shipper-Carrier Relationship Forms

Shipper-carrier relationships sit along a spectrum of forms (see Figure 1-1). Most shippers and carriers engage in a combination of these relationships for different segments of their businesses.

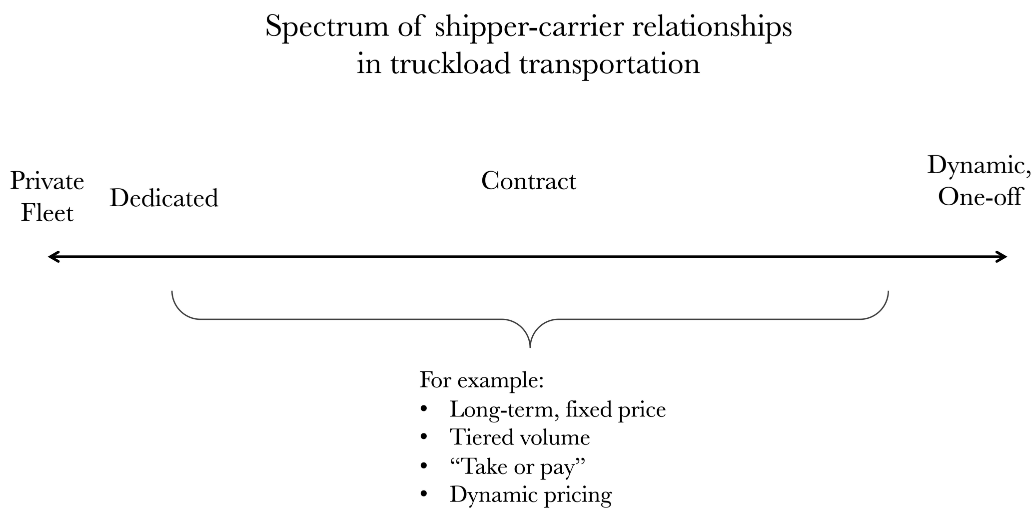


Figure 1-1: Shipper-Carrier Relationship Forms

1.3.1 Integrated Services

At one extreme is the private fleet. In this setting, the shipper vertically integrates the transportation function into its business. It owns the assets (tractors and trailers) and manages its transportation needs in-house. This includes hiring and training company drivers, scheduling and routing shipments, and equipment maintenance. All processes are controlled by the shipper.

Shippers with very high volume density may opt for private carriage. Owning the process offers extremely high guaranteed service such as on-time pickup and delivery and availability of capacity when needed. This reduces potential unexpected costs that often result in the for-hire relationship form. Moreover, company drivers can garner better relationships at customer facilities and improve the shipper's reputation in this way. However, the operational efficiency is worse for private fleets than their outsourced alternatives (e.g., lower truck utilization and higher empty miles driven and cost per mile).

Outsourcing some or all aspects can be beneficial. A very closely related relationship form, and sitting just next to the vertical integration form on the spectrum, is the dedicated fleet relationship. Under this model, the shipper leases some fixed amount of an external carriers' capacity, which is at the shipper's full disposal. The shipper still owns the operations and expects the same performance benefits as it does when it vertically integrates the transportation function with a private fleet, but the shipper does not own the physical assets and its required long-term capital investment.

Shippers must weigh the cost, efficiency, and performance tradeoffs when considering private or dedicated relationships. They require internal resources and capital investments that are not present with an outsourced relationship.

1.3.2 Spot Market

At the other end of the spectrum are the dynamic, one-off transactions on the spot market. A spot transactions occurs in a highly competitive market: there are many firms (i.e., carriers) with very low barriers to entry (only the cost of a truck and trailer at about $100k, and a commercial driver's license, or CDL) supplying a homogeneous product (an empty, standardized truck), and no single firm is large enough to have an effect on market prices (all firms are price takers).

Exchanges here are for a single load transaction. They are typically, but not always, used as a backup option when the shipper cannot find capacity otherwise. If a shipper has a load it wants moved on the spot market, it posts the load's information (pick-up and drop-off location information, equipment requirements, schedule needs, etc.) on a load board. Carriers with available trucks can see these postings and call to make arrangements. Similarly, carriers looking to fill their trucks, often to reposition to a next scheduled pickup load and avoid a deadhead move, can post their current location, desired destination location, and the price they require for the move on a load board.

These load board are now almost exclusively online but historically had been physical bulletin boards at truck stops. The first online load board was introduced in 1995 by Truckstop.com (originally InternetTruckstop.com) where shippers and carriers could post their information to an electronic load board hosted on the internet with access to all the information at their fingertips. Many electronic load board options exist today; their use has evolved since the mid-1990s. The matching process grew from the manual process described to newer automated freight matching. With these "digital freight matching" services, third party firms provide the load board platform and apply optimization algorithms to find the most attractive, highest priced freight for carriers on the supply side, and lowest-priced capacity for the shippers on the demand side.

Regardless of the method, the spot market price a shipper pays is determined by many factors. These include the the lead time (or the amount of time the shipper can wait to negotiate or see other carriers post capacity and lower prices), the current availability of capacity and general supply and demand balance in the market, the available carriers' internal cost structure and as a result, reservation price (that is, the lowest price the seller of a good or service is willing to accept to provide), and so on. In fact, on any given day, two loads could be moved on the same lane via the spot market but at very different prices by two different carriers (discussed in detail in Chapter 4).

Moreover, the spot market price functions closely to those of commodities such as energy markets and bulk agricultural products. TL spot market prices tend to be similarly volatile in their fluctuations (see [113]). Thus, not only is the shipper-carrier relationship in the spot market in effect for just a single transaction, but the price and performance uncertainty levels for the shipper are high. This is one of the reasons shippers tend to rely on the spot market as a backup capacity option - typically only about 5-10% of a shipper's total transactions [40]. However, some shippers do incorporate a set, perhaps higher amount of spot transactions into their strategic transportation plan.

1.3.3 Contractual Relationships

Occupying the middle of the spectrum are the for-hire contractual agreements. A shipper and a carrier agree to a set price on a lane-by-lane basis established during the shipper's strategic procurement event (discussed in subsection 1.4.1). Within the contracted for-hire relationship, shippers and carriers use a portfolio of contract types, some listed in the Figure and described further in Section 1.7. The most prevalent contract, however, is the long-term, fixed-price contract which is in place for typically 1-2 years.

Shippers and carriers typically engage in more than one relationship form. For example, shippers with their own large private fleets such as PepsiCo and Walmart Stores also supplement their TL demand with for-hire and spot capacity. Alternatively, they may offer their excess capacity to other shippers on backhaul lanes in order to balance their network and reposition assets. Further, most shippers engage in both contract and spot relationships. At the same time, carriers may provide services along the whole spectrum as well: dedicated capacity for some customers, contracts for others, and spot capacity for one-off moves when they need to reposition trucks for their next shipments.

1.4 Two Stages of Truckload Transportation

Within the contractual relationship, the forms and mechanisms implemented vary. To understand the challenges that come up between shippers and their outsourced, for-hire TL carriers, it is important to understand the two stages of these interactions, their individual challenges, and the importance of not neglecting one during the execution of the other. The first stage is the strategic transportation procurement process. The second is the operational process of load tendering by the shipper and the carrier's load acceptance decision. These processes are described further in the following sections.

1.4.1 Strategic Procurement

The first stage is comprised of the strategic procurement process (see Figure 1-2). To procure contracted capacity, the shipper conducts a reverse auction. In a reverse auction, rather than the buyer responding to the seller with his willingness to pay for the good or service in a more commonly known forward auction, the seller of the good or service responds to the buyer with the price she is willing to accept. It forecasts expected annual transportation demand for each lane (origin-destination pair) within its freight distribution networks, decides if and how to package lanes together in order to make them more attractive to carriers, and invite carriers to participate by issuing a Request for Proposals (RFPs).

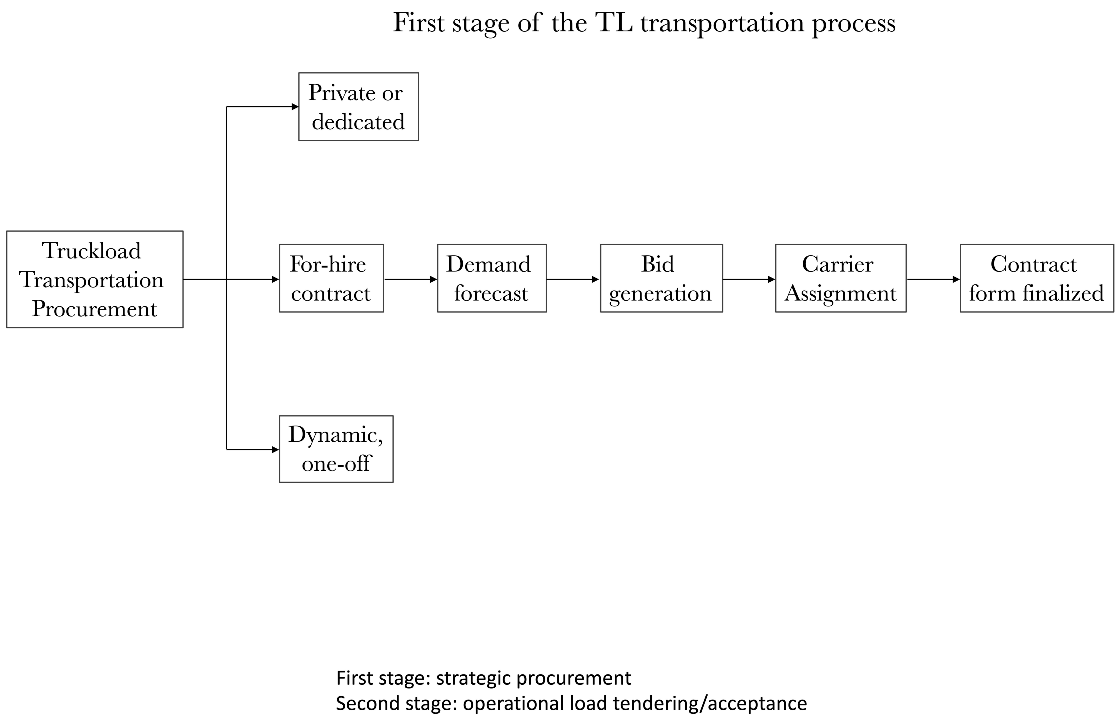


Figure 1-2: First Stage of For-Hire Truckload Transportation

The carriers respond with their sealed bid prices – the prices they are willing to accept to serve each lane they are interested in. The price is based on origin and destination characteristics (e.g. expected dwell time - time spent waiting at facilities - or likelihood of obtaining a follow-on load after drop-off), expected volume forecasted by the shipper (number of loads offered per week), the carriers' internal operating costs, fleet size, and existing network characteristics. This is referred to as the bid generation problem in the literature (e.g., [43, 140, 89]). The carrier's objective is the maximize profit across a network of shipper customers subject to network fit and capacity constraints. Thus, when determining its bid price, the carrier also considers its existing network balance of customer demand, total capacity, and how well the expected demand on each lane from this shipper fits in.

Once the carriers' bid prices are collected, the shipper chooses the carriers with which to enter contracts on each lane. This is known as the carrier assignment problem (e.g. [43, 65]). The shipper wants to encourage carriers to bid as close to their reservation prices as possible – that is, the lowest price they are willing to accept to serve the demand while maintaining the shipper's desired service levels. This last point becomes important during the second stage of for-hire TL transportation. Thus, the shipper chooses carriers with the lowest bids subject to constraints such as historical performance levels, service type, incumbency, technological sophistication, etc.

This process results in contracts between the shipper and carriers. Under these contracts, prices are set and held exclusively and privately between shippers and carriers, typically lasting one or two years. While the price is set, a unique aspect of TL contracts is that they are non-binding on both sides: in (1) demand of loads tendered from the shipper and (2) supply of capacity from the carrier. This is because unavoidable demand forecasting errors mean the expected or "awarded" volume the shipper forecasted and communicated to the carrier may not ever materialize (referred to as "Ghost Freight", (see working paper by [4]), or more than the expected volume may need to be moved. Thus, the shipper is not required to offer, or tender the awarded volume over the contract time period. At the same time, due to difficulties of ensuring a truck and driver are available at the time and location every load is ready to be moved, the carrier is not required to accept every load it is tendered. This is due to the difficulty for carriers to know with certainty that a truck will be available at the time and location each load does materialize. Since neither side wants to be penalized for uncertainties that are largely out of their control, TL contracts are flexible in this way.

Due to the possibility that contracted (also referred to as primary) carriers may reject loads, shippers construct a routing guide for each lane. The routing guide is a sequential list beginning with the primary carriers followed by a set of backup carriers. These backup carriers typically have unsuccessfully bid on the lane, thus expressing ability and willingness to serve at least some of the lane's demand. Most often, these carriers had bid higher prices than the winning carrier. The backup carriers are entered into the routing guide at their bid price, however, there is no contract in place between the shipper and backup carriers.

Not only does this mean the backup carriers' routing guide price is not binding, but there is a much lower assumption that backup carriers will produce capacity when called upon.

1.4.2 Operational Load Tendering and Acceptance

The non-binding nature of the contracts comes into play during their operationalization, or the second stage of TL transportation: the shipper load tendering and carrier acceptance decisions (see Figure 1-3).

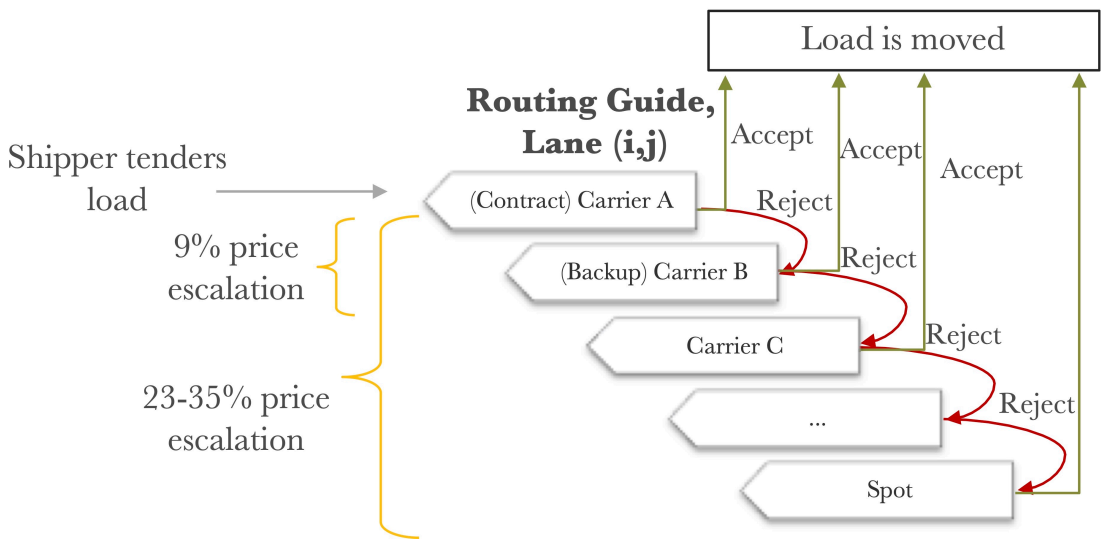


Figure 1-3: Second Stage of For-Hire Truckload Transportation

When a shipper has a load it needs moved, it tenders the load to its contracted carrier for that lane at the contracted price. This carrier can accept or reject the load for any number of reasons. If it is rejected, the shipper tenders the load to the next backup carrier in its routing guide. The prices associated with each backup carrier correspond to their bid prices, which can be 9-15% more than the contract price [2]. For this reason, shippers want to ensure their contracted carriers maintain high freight acceptance rates. This acceptance rate is one of the key performance measures shippers rate their contracted carriers on.

Like primary carriers, backup carriers may accept or reject load offers. The shipper proceeds down the routing guide until a backup carrier accepts the load or some price or time threshold is hit. At that point, the shipper may turn to the spot market where it consults a load board to find a carrier for its load. Relying on the spot market can cost the shipper 23-35% more than its contracted price.

This describes the supply side non-binding aspect. The demand side involves the amount of volume the shipper actually tenders to the carrier. During the first stage procurement process, the shipper communicates each lane's awarded volume to the carriers – that is, the number of loads per week the carrier should expect on the lane over the contract. However, due to changes in the shipper's end customers' actual demand, the shipper may tender some amount above or below this expected volume. Either of these scenarios are difficult for carriers to manage.

1.5 Freight Market Cycles

As discussed earlier, the dual-sided supply and demand flexibility in these contractual relationships are sources of unanticipated costs and uncertainty for shippers and carriers. These issues worsen when we consider macroeconomic factors of the market.

The freight transportation industry undergoes market fluctuations due to changes in the overall economy, hedging or otherwise opportunistic pricing during the transportation procurement process, truck driver shortages or churn, regulations, and extreme weather events [113]. In what is referred to as a soft market, the demand for transportation services is less than available trucking capacity, driving prices down and carriers to accept only marginally profitable or even unprofitable freight just to fill their trucks and cover costs – this is when shippers hold the more advantageous position.

During soft market conditions prices can becomes so low that it becomes difficult for carriers to cover their costs. As they typically operate with low cash flows and slim margins to begin with, many carriers - especially small ones - go out of business during soft market periods. In fact, the Wall Street Journal reports that there are 3-4 times the number of carrier bankruptcies during these times than any other [152].

On the other hand, in a tight market, demand for transportation outstrips supply, carriers have their pick of freight they want to move, prices are high, and carriers have the advantage. The industry experienced a sever cycle from a soft to a tight market around Q3

2017 [113], and markets began to soften again around Q4 2018. The markets remained soft until the onset of Covid-19 pandemic, hitting the US economy by about March of 2020.

1.6 Truckload Carrier Competitive Landscape

The characteristics of TL carriers vary dramatically, in particular their size and service type. Carriers' strategies, operating expenses, and access to shippers' business are influenced by these characteristics. Moreover, they directly impact the relationships they form with shippers.

1.6.1 Carrier Fleet Size

The carrier's fleet size - or number of trucks - dramatically impacts its operations. While the US Department of Transportation reports over 700,000 registered motor carriers in the US in 2021 (including TL and LTL), the distribution of fleet size is highly skewed. For example, 91% of carriers have 6 or fewer trucks, 97% have less than 20 trucks, and less than 200 carriers have over 500 trucks [51].

The smallest carrier size is the owner-operator. This is where a single driver or team of two drivers own from 1-3 trucks. Sometimes these carriers work directly with shippers, but more often they either lease their capacity and service to larger asset carriers. They operate by filling in the gaps in shippers' network coverage as they can me much more flexible than larger carriers. However, these owner-operators often lack resources such as dispatching services, finance departments, and strategic or tactical business knowledge to operate efficient and profitable businesses. However, while the owner-operators do not typically run highly profitable businesses, the draw is the independence that comes with self employment and the truck driving lifestyle.

Small and medium sized carriers that operate in the 10-500 truck range typically operate best focusing on regional operations. They are large enough to serve small and medium shippers' needs and can offer more specialized, niche service. These small- and mid-sized carriers can be very successful operating in this way.

While there are tens to hundreds of thousands of owner-operators, small, and medium carriers, there are just a handful of large fleet carriers. Knight-Swift, J.B. Hunt, Landstar, and Schneider are major players in the US that own and operate thousands of trucks each. Due to their scale, they can efficiently operate nation-wide and serve large and small shippers alike. Shippers tend to form long-term contractual relationships with a set of core carriers, comprised of mostly larger carriers. Important to note however, due to the sheer size and competitiveness of the market, these large carriers still function as price takers; shippers always have another provider option willing to take their freight for a lower price.

1.6.2 Carrier Service Type

Within the carrier service type category, there are asset-based carriers and brokers (or "non-asset" carriers). Asset carriers own the trucks and all equipment needed to move goods. They have direct relationships with the shippers that hire them.

Non-asset carriers (also called freight brokers or freight forwarders), on the other hand, act as the middle man between shippers and asset-based carriers. As a result, their cost structures and relationships with shippers differ significantly from those of asset carriers. Brokers' margins depend on how well they can hedge the market. They enter contractual relationships with shippers but they typically buy capacity from carriers on the spot market - often operating their own internal load boards and an army of agents on answering and making phone calls to match loads to trucks.

Shippers may strategically use brokers as contracted carriers for lanes that are difficult to cover with their core asset carrier base. Usually these are lanes with low volume and/or irregular demand patterns. They may also rely on brokers as backup options in their routing guides when their contracted carriers reject loads.

Since shippers tend to prefer to manage small core carrier bases, they can maintain a relationship with the broker that then accesses a larger pool of smaller carriers. In exchange, particularly smaller carriers use brokers to access business they would not have access to otherwise. Carriers also use brokers to reposition capacity rather than having a deadhead move and incur the associated costs of empty miles.

The role and reputation of brokers has evolved over the decades. Traditionally they were viewed as a last resort option for shippers. They were - and still are - criticized for taking too high of a cut of the transaction, introducing economic inefficiencies into the equation. However, they have also become a key strategic resource for shippers and carriers. Due to their access to vast amounts of data in the form of large volumes of transactions from multiple shipper customers and carrier providers, they can offer added value to a shipper's business. For example, they provide transportation analytics insights such as benchmarking rates to help the shipper identify how close to average market prices it is paying for certain lanes or regions. Similarly, brokers argue that they offer deep industry knowledge and attractive freight options for carriers. In the US, major players in the brokerage space include C.H. Robinson, XPO Logistics, and Coyote Logistics.

1.7 Trends in Contractual Relationship Forms

With the traditional process and players established, we discuss trends the industry is experiencing in new contractual relations. Contractual forms are in place that help spread the inherent financial risks and reduce uncertainty due to the non-binding nature of the traditional contract. The most common contract form is the long-term, fixed-price contract, typically in effect for 1-2 years [101, 40]. However, alternative contract forms have been explored in practice – and to a lesser extent in the literature.

Under a tiered volume pricing contract, for example, the shipper and carrier agree to different fixed prices for different volume levels. For carriers, more volume is not always better. A carrier serves many customers; more capacity allocated to one customer, particularly more than what has been strategically planned, throws a carrier's network out of balance and may cause it to underserve another customer. As economies of scale do not apply (see [39], a shipper that offers more volume should not expect to get volume discounts. In fact, shippers may have to pay more for higher volume.

Two additional alternative contract forms that are sometimes used in practice are guaranteed volumes and guaranteed service levels. Under the former contract form, a shipper commits to a set volume tendered to the carrier and if that demand does not materialize, it must still pay the carrier for the unused capacity - either at the load price or some fraction of it. This is colloquailly called “take or pay”. In the latter contractual form, the carrier commits to hauling a set volume. If it cannot take those loads, then it must pay.

Finally, a contractual form that has received a lot of attention from practitioners is the market- or index-based contract (see, for example, [149, 31]). With this contract, the price is no longer fixed. An initial price is decided during the procurement stage of the process, however the price is allowed to fluctuate in time up or down as the freight market does. The argument for these contract forms is that since the market introduces uncertainty and reduced contract compliance, these dynamics could actually be used to incentivize the desired outcomes, specifically by (1) reducing shippers’ reliance on higher priced and lower performance backup carriers, and (2) ensuring carriers are paid a market-competitive price for their services[2].

1.8 Outline of Thesis

The remainder of this thesis is structured as follows. In Chapter 2, we explore the question of carrier reciprocity in tight markets when they have a more advantageous position and what shippers can do to mitigate the effects of the market to maintain high contracted carrier acceptance rates. In Chapter 3, we quantify contracted carriers’ willingness to stick to the contract as its best-known alternative becomes more attractive than the contract price. We take a segmentation approach and determine this measure of contract price stickiness for different pieces of shippers’ network of lanes, carrier base, and market conditions. In doing so, we determine the potential opportunities for an alternative shipper-carrier relationship form that has received much attention from practitioners: market-based contract. In Chapter 4, we explore how shippers and carriers should design the market-based contract explored in the previous chapter, so that it results in a Pareto improvement over the status quo fixed-price contract. Finally, in Chapter 5, we summarize the key contributions of this thesis and areas to extend the research.

[2]Index-based contracts are well established in other industries in which market prices are volatile, market patterns are cyclical, both buyers and sellers want an exogenous way to determine a “fair” price. They are used in in ocean shipping, energy markets, and for many agricultural products, for example.

Chapter 2

An Empirical Analysis of Carrier Reciprocity in Dynamic Freight Markets

Abstract

Dynamic macroeconomic conditions and non-binding truckload freight contracts enable both shippers and carriers to behave opportunistically. We present an empirical analysis of carrier reciprocity in the US truckload transportation sector to demonstrate whether consistent performance and fair pricing by shippers when markets are in their favor result in maintained primary carrier tender acceptance when markets turn. The results suggest carriers have short memories: they do not remember shippers' previous period pricing, tendering behavior, or performance when making freight acceptance decisions. However, carriers appear to be myopic and respond to shippers' current market period behaviors, ostensibly without regard to shippers' previous behaviors.

2.1 Introduction

A key performance indicator used by shippers is a contracted (also called primary) carrier's acceptance ratio (PAR) - that is, the fraction of loads the carrier accepts relative to the number of loads that it is offered by the shipper on a lane a carrier has won in the strategic procurement process. When a primary carrier's PAR declines, as is seen in tight market periods, shippers must go deeper in their routing guides to find non-contracted carriers to move their loads. Elevated freight rejection and high transportation costs in tight market periods hinders overall business performance [138]. As a result, during tight periods shippers search for ways to develop relationships with their carriers and ensure consistent performance through the cycle, seeking 'shipper of choice' status, [46].

The tensions between shippers and carriers that we highlight in Chapter 1 motivate this chapter, in which we test the often-heard opinion by shippers and carriers that they each "have long memories" between market cycles.

2.2 Literature Review

A vast set of literature related to the PAR problem models the load matching and carrier assignment decisions to help shippers and carriers improve performance. Due to advances in near-real-time data availability and computational capabilities, much of the recent literature has explored dynamic versions of these problems, where the shippers' load information becomes available over time rather than being known a priori. A related, but much smaller body of literature studies shipper-carrier relationships.

2.2.1 Strategic carrier assignment decisions

Two problems in the transportation literature, which are related but distinct, are the shipper's strategic decision during the procurement process of which primary carrier to assign to each lane, and the operational decision regarding which carrier to offer each load.

[39] and [42] argue that economies of scope come about when shippers and carriers make procurement and bid decisions by considering their network as a whole rather than on a lane-by-lane basis. A combinatorial optimization approach to the procurement auction is developed to improve efficiencies; in particular, to improve how carriers communicate their cost function information to shippers and how shippers assign carriers to lanes. These studies focus on the bids that carriers submit in response to the RFP and develop formal optimization models and solutions for how shippers should analyze these submissions and assign carriers to lanes.

Building on this research, [98] considers both bid cost and carrier reputation in the shipper's combinatorial carrier assignment problem. The author finds that when taking reputation into account during the primary carrier selection process, shippers pay more in total direct costs to carriers, but there is a decrease in total hidden costs, which results in a net savings in total transportation costs. Including carrier reputation in the assignment decision is also considered by [119]. Here, the reputation-based winner determination problem where allocation of long-term contracts to carriers is decided based on both bid ask prices and carriers' reputation. The carriers' reputation is translated into unexpected hidden costs the shipper may incur when working with the winning carriers.

The above literature assumes the shipper has decided to outsource the transportation service and models the shipper's selection of for-hire carriers. However, a number of studies have approached the carrier assignment problem in which the shipper chooses between contracting for-hire carriers and serving the lanes itself with a private fleet. This set of literature includes [64], [41], [90], and [104]. For a review of the TL transportation service procurement modeling approaches including contract or spot agreements between shippers and carriers, the reader is referred to [20].

2.2.2 Operational load matching decisions

The load assignment problem finds the minimum-cost assignment of loads to carriers, where each load must be picked up at a given location at a specific time and delivered to a known destination (see [115]). [114] presents a methodology to evaluate the dynamic load assignment model with demand uncertainty and real time information using rolling horizon simulations.

Similarly, simulation approaches have been applied to study the dynamic nature of TL load matching operations by [148] and [118]. Further, [133] demonstrate an approximate dynamic programming method to accurately simulate the operations of the largest truckload motor carrier in the US, with over 6,000 drivers. The authors make use of both pre- and post-decision states in order to deal with the complexities of carriers' operations. The model includes driver attributes such as home domicile, desire to be offered loads that get them back home, and hours already worked. Load attributes including origin and destination, appointment and delivery types, and revenue, are also included to allow the company to test the value of changes in the mix of drivers, freight, and other operating policies. The authors highlight that their model can be extended to help the carrier make tender acceptance decisions at the load level based on the current state of the system and ultimately evaluate how to commit to contractual agreements with shippers.

2.2.3 Shipper-carrier relationships

The S-C relationship as it relates to market conditions and the nature of the resulting contractual agreements is the focus of [79], which studies relationship-specific investments in the trucking industry. The author argues that in trucking, the customer (i.e., shipper) rarely experiences asset specificity - that is, investments specific to partners (carriers) over long horizons as compared to other industries. He argues that with increasing long-haul trucking market thickness - that is, as more buyers and sellers come to the market - shippers and carriers utilize simple spot transactions rather than more complicated contractual arrangements.

Much of the extant transportation literature on shipper-carrier (S-C) relationships relies on surveys, interviews, and isolated experiments rather than empirical industry data, in part due to the difficulty of obtaining private company data. Carriers' freight acceptance has been the focus of a few studies based on transaction data, however. Rather than measures of the S-C relationship, most authors consider attributes of the lanes and freight in determining whether a load will be accepted. High lane volume [70], low lane volume volatility [83], high pricing [8], and high lane consistency, or cadence, [5] have been found to be positively correlated with higher primary carrier AR (PAR).

The S-C relationship, however, has been found to contribute to carriers' freight acceptance in a few studies. [170] find that contracted carriers, which shippers believe represent good relationships, outperform carriers that demonstrate poor or purely transactional relationships in terms of freight acceptance, on-time delivery, and pre-positioned capacity.

[131] analyze contract and spot market transactions and the impact the S-C relationship

on freight acceptance. They find that less frequent load offers increase the likelihood of a carrier rejecting a load, while higher offered volume, lower load offer volatility, and higher revenue transacted between the shipper and carrier increase the likelihood of carrier's load acceptance. The authors do use a measure of market condition - Spot Premium - which is the ratio of 7-day average spot prices to 7-day average contract prices in a geographic region. As the unit of analysis in the model is at the load level, this rolling ratio still indicates immediate market conditions, testing whether carriers are more or less likely to accept loads in one market or another. We extend this market condition consideration by including *previous* market behaviors into the acceptance decision model.

Each of these studies utilizes data from a single shipper, which limits their ability to generalize across types of shippers or to segment the data by S-C relationship types. Moreover, while the above literature explores the S-C relationship and its effects on carriers' tender acceptance, it does not consider previous behaviors explicitly defined by market conditions. Our research contributes to the TL transportation literature in that we analyze transactions and behaviors between many US shippers and contracted carriers and determine how the relationship between them in one market condition corresponds to performance in the next, when power dynamics have shifted.

2.3 Research Question and Hypotheses

To address the research gaps described above, we formulate the following research question, which considers the degree to which shippers and primary carriers stand by the terms of their previously defined non-binding contracts as market conditions (and power dynamics) change:

RQ: Should shippers pay competitive prices or provide high, consistent freight volumes when markets are soft to ensure carriers will adhere to contract expectations when the market tightens with high acceptance ratios of contracted loads?

In particular, do carriers have long memories (thus, act like elephants) and reciprocate high soft market period pricing, tender consistency, or high volume from shippers with consistent, high PAR in tight periods? Conversely, do carriers reciprocate low soft period pricing, tender consistency, or volume from shippers with lower tight period PAR? Or, do carriers act like goldfish with short memories, forgetting shippers' previous behaviors?

In short, our research question considers whether it is worth it for a shipper to "pay it forward" by strictly adhering to contract commitments even when lower cost options may be available. This we measure with primary carrier acceptance ratio and formulate the following hypotheses:

2.3.1 Carriers' response to previous shipper behavior

First, we focus our attention on whether carriers respond to shippers' previous behaviors. That is, do carriers have long memories (thus, act like elephants) and reciprocate good performance soft market from shippers with high PAR in tight periods? We measure "good" performance as competitive pricing, consistent tendering patterns, advanced load notice, minimal delays at pick-up and delivery, and frequent business. We consider asset-based primary carriers independently from non-asset primary carriers (which, in our context includes brokers serving as contracted carriers), as they may have different strategies, access to capacity, and perhaps tolerances for certain customer behaviors. This leads us to the following two hypotheses:

H1a: Asset-based primary carriers offer high tight market period PAR for their shippers that had demonstrated good performance in the previous soft market period.

H1b: Non-asset primary carriers offer high tight market period PAR for their shippers that had demonstrated good performance in the previous soft market period.

2.3.2 Carriers' response to current shipper behavior

Second, we consider whether carriers respond to shippers' current behaviors. That is, do carriers act more like goldfish and offer high tight market PAR for shippers demonstrating current good behaviors? Thus, we formulate the following two hypotheses:

H2a: Asset-based primary carriers offer high tight market period PAR for their shippers that demonstrate good performance in the current tight market period.

H2b: Non-asset primary carriers offer high tight market period PAR for their shippers that demonstrate good performance in the current tight market period.

2.4 Methodology

In the following section, we describe the hybrid break point detection method implemented for this study, define the market time periods, characterize the S-C pairs and describe the model we implement to answer our research question.

2.4.1 Empirical Data Summary

The data for this research is obtained from a partner company that provides logistics and transportation functions for its shipper clients. Our partner company is not a brokerage, but rather is a managed services provider assisting its clients (shippers). The relationship and all payments are directly between the shipper and the carrier with no intermediaries. This is important to note, as we intend to study the direct relationship between a shipper and its primary carriers, not how carriers interact with a third party. Many companies offer such managed transportation services. For example, Transplace, BluJay Solutions, 4Flow, C.H. Robinson, and XPO Logistics. While the latter firms are two of the largest traditional 3PL providers in the US, they both offer independent managed transportation services as described above without promoting or relying on the brokerage side of their businesses.

For our purposes, we obtain both tactical pricing information and operational load transaction data. The carriers' pricing information is obtained during each shippers' procurement event; prices carriers bid for each lane are collected, one or more carriers are selected to be contracted, and a subset of the remaining carriers that bid are retained as backup carriers in cases where the primary carriers reject loads. Typically this process is either done solely by the client shipper, or with the help of our partner company, depending on the the shipper's needs. This pricing information for each carrier - primary (i.e., contracted) and all backup (i.e., non-contracted) - is uploaded into the routing guide, which we obtain from our partner company.

The provided data further includes transactions for all of the client shippers' TL loads that originate and terminate within the continental United States, each loads' corresponding carrier tender sequence beginning with the primary carrier through any backup carriers if needed and each carriers' acceptance or rejection decisions until a carrier finally accepts the load.

After data cleaning and pre-processing, we retain just under four years of transactions; from September 2015 to May 2019. In particular, the data spans one complete market cycle into the next - that is, a soft period, a tight period, and the beginning of a second soft period, which are defined and tested in §4.2. For consistency, we consider only TL, dry van, long haul (greater than 250 miles) loads, since pricing schemes differ for other freight types and shorter lengths of haul. As a result, we maintain 1,933,299 unique loads offered by 71 shippers, tendered to 1,650 primary carriers. They comprise 7,573 shipper-primary carrier pairs operating from 4,915 origin 5-digit zip codes to 9,668 destination 5-digit zip codes. The dataset also contains geographic regions defined by our industry partner as key economic zones of shippers' business. The continental United States is divided into a set of 135 mutually exclusive and collectively exhaustive regions. Each of these regions represents an origin and a destination region.

Table 1 below summarizes the data set. In addition to these variables, the metrics described in the following subsections are computed to measure attributes of the S-C relationship.

2.4.2 Break Point Detection and Market Period Definition

First, we define the market periods by detecting breaks in market data. The time series of aggregated weekly PAR across all shippers is used as a proxy for market conditions. Previous literature ([131], [127], and [85], for example), use monthly spot market prices to characterize market conditions. In our dataset only about 2% of the total transactions go to the spot market and this value changes with market conditions as more loads go to the

Table 2.1: Data Summary

Date range	Sept 1, 2015 - May 30, 2019
Shipment Type	TL, dry van, long haul ($\geq$250 mi)
Num. Loads	1,933,299
Num. Shippers	71
Industries	Automotive, Food & Beverage or Consumer Package Goods, Manufacturing, Paper & Packaging, Other
Num. Primary Carriers	1,640 (Asset-based: 1,446, Non-asset: 194)
Num. Shipper-(Primary)Carrier pairs	7,573
Num. Origin 5-digit Zips (Regions)	4,915 (135)
Num. Destination 5-digit Zips (Regions)	9,668 (135)

spot market during tight periods, As such, our data is limited in using spot prices to define market periods as other studies have.

However, [131] show a 77.3% correlation between spot premium time series data and the Morgan Stanley supply and demand sentiment index, which surveys a broad array of shippers, carriers, and brokers. Our weekly national PAR series and this same Morgan Stanley weekly freight index from 2015 to 2019 have a correlation of -85.2%. That is, we show that high demand for capacity correlates with lower PAR - the definition of a tight market. Not only are supply and demand over time strongly correlated with spot prices as demonstrated in [131], we show that they are strongly correlated to our measure of the market, PAR. Thus we can be confident in choosing PAR rather than other metrics (e.g., spot prices, freight industry indices such as national load-to-truck ratio, truck manufacturing order rates, or other macroeconomic indicators) to define market conditions; acceptance ratios capture real-time carrier decisions.

In practice, many shipper monitor the changing market conditions by tracking the PAR along with outer routing guide failures – not just spot market price changes. In fact, one could argue that PAR is a more sensitive metric than spot premium as it captures contract abandonment by carriers. Thus, PAR is an appropriately representative and sensitive measure of how carriers respond to and communicate with shippers and, thus, of overall market conditions. Further, by considering the aggregate measure, the break points detected do not presuppose this study's results.

To define and validate the identified breaks, we combine two distinct break point detection methods. As neither method perfectly captures the changes we expect, we run each method independently and take points that are identified by both methods as validated breaks. This results in a more robust break point detection method.

Quandt Algorithm

The Quandt method, developed by [48] and expanded by [117] is the standard econometric test for a break point that is not known a priori. The method calculates a sequence of Pearson's chi-square statistics to determine when the series has shifted [68]. The algorithm considers each point in the time series as a candidate break point, τ. The data is split into sub periods at τ: A for $t \leq \tau$, and B for $t > \tau$. We estimate the vector of coefficients of a linear regression fit to each sub period and to the full series without a break (denoted by T):

$$\begin{aligned}
\hat{y}_t^A &= \hat{\beta}_0^A + \hat{\beta}_1^A X_t^A + \varepsilon^A, \qquad t \leq \tau \\
\hat{y}_t^B &= \hat{\beta}_0^B + \hat{\beta}_1^B X_t^B + \varepsilon^B, \qquad t > \tau \\
\hat{y}_t^T &= \hat{\beta}_0^T + \hat{\beta}_1^T X_t^T + \varepsilon^T, \qquad \forall t
\end{aligned} \tag{2.1}$$

The chi-square distribution is used to assess the statistical significance of the difference between the two regression models and the equality of the coefficients is tested using an F-statistic (the ratio of two chi-square distributions) [48, 68].

The break point, $\hat{\tau}$, is determined to be the point at which the maximum F_τ is achieved: $\hat{\tau} = argmax_\tau F_\tau$

Typical F-statistic critical values for such tests are not appropriate for $F_{\hat{\tau}}$ when the break point is not known a priori and the Quandt method is applied [68]. However, critical value tables for the Quandt statistic and p-value calculation methods have been developed by [9], [10], and [67].

The Quandt algorithm described above considers all points in the time series for which

$\tau \in [\tau_0, \tau_1]$, where we exclude the tails of the data from consideration as candidate points to maintain enough observations on either side to estimate meaningful regression models [117, 68]. The sample is trimmed to the interior λ to $(1-\lambda)$ range of the data. [68] suggests that typical values of the trimming parameter, λ, fall between 0.05 and 0.15. In this study we use $\lambda = 0.05$ trim and address its sensitivity by defining a transition period on either side of a detected break point when defining our market periods.

Considerations of the Quandt Algorithm

While well documented in econometric and statistics literature, the Chow and Quandt algorithms have a number of drawbacks when applied to the data presented in this paper. First, the methods are well suited for time series data containing a single break, and despite extensions developed to account for multiple breaks, the results of these extensions are sensitive to both the date range and the trim parameter selected.

The second drawback of the Quandt algorithm as applied to this research is that it tests for locations in which one can fit the best linear model to each segment of the data. However, that is not the intent of this step in our research. We merely require a method to determine when average values of the aggregated PAR has changed. Unfortunately, break point detection based on average values is not found extensively in the literature. As such, we combine the method described above with a sequential difference of means test described in the following subsection.

Jumping Mean Approach

We are interested in determining the times at which the average value of the aggregated PAR changes. To identify these points, for each date in the time series, τ, we compute the average of the PAR values for the preceding W weeks and the following W weeks. These backward- and forward-looking average values are denoted by $\overline{X}_\tau^{b,W}$ and $\overline{X}_\tau^{f,W}$, respectively and we take the absolute value of the difference between the two values, D_τ^W:

$$D_\tau^W = |\overline{X}_\tau^{b,W} - \overline{X}_\tau^{f,W}| \tag{2.2}$$

Where

$$\begin{aligned}\overline{X}_{\tau}^{b,W} &= \frac{1}{W}\sum_{i=0}^{W-1} X_{\tau-i} \\ \overline{X}_{\tau}^{f,W} &= \frac{1}{W}\sum_{i=1}^{W} X_{\tau+i}\end{aligned} \tag{2.3}$$

In doing so, we draw from a data mining context, which describes such a 'jumping mean' break point measure, where a break is detected if a difference score (which we define here as the absolute difference in means between the forward- and backward-looking window of the data) is large enough [144]. We consider the points at which the local maxima of the absolute difference occur as potential break locations, as suggested by [68] and [15] for the Quandt method.

Next, applying a standard t-test, we check the statistical significance of the difference of means at the potential breaks, τ. If $t_{stat} > t_{crit}$ at the desired significance level we consider it as a potential break point.

We test the sensitivity of this combined method to the window size, W. Windows of 6, 10, 12, 26, 34, and 52 weeks are tested and we determine that a satisfactory window size is 12 weeks, as smaller windows are too sensitive to variability in the data, detecting many immaterial break points, and larger windows over-smooth, missing potentially important breaks (see Appendix A for window sensitivity test results).

Next, we combine the results of the Quandt algorithm and this jumping mean approach. We retain only break dates that are identified by both the Quandt method and the jumping mean approach with a 12-week window. These dates are summarized in Table 2. The resulting validated break points are the first week of February 2016, the first week of July 2017, and the second week of January 2019 (see Figure 1).

Market Period Definitions

While determining a break at a single point is simple and convenient, claiming that an economic market change takes place at a single specified date is not realistic. We use the break point detection method described above as a first order approximation of a structural change point [68]. To account for seasonal effects, we take a full year of data for each

Table 2.2: Break points identified by the Quandt method and jumping mean approach

Quandt results	Jumping mean
02-05-16	02-07-16
	05-29-16
	01-29-17
06-28-17	06-25-17
	09-02-18
01-16-19	01-20-19

defined time period. By selecting a full year during each market period we average out any seasonality individual shippers or industries observe. In this way, seasonality in demand for distinct shippers or industries in one market period is compared to the same seasonality in the second market period and allows us to compare across shippers of different size and industry. Further, in order to allow for market transition buffer time, we choose the year of data that is centered between our validated break dates.

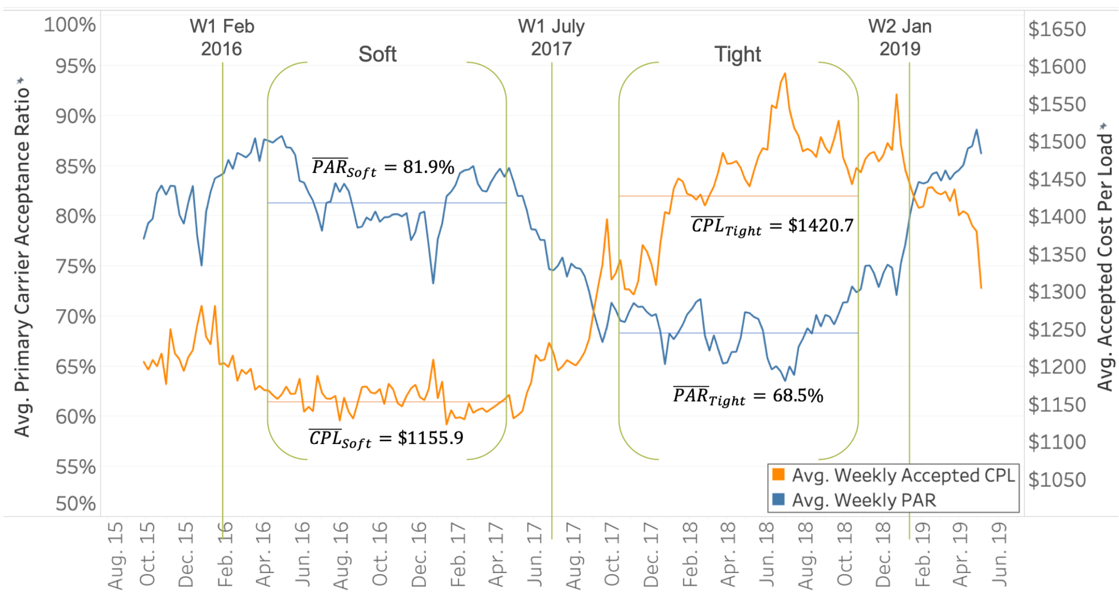
While it is reasonable to consider economic markets as continuous rather than discrete as we do here, we define our research question such that we measure market-specific behaviors, and thus must define market periods. To procure transportation services, shippers typically run annual national bid events, which are essentially seen as step changes resulting in rates and primary carrier agreements that are expected to be in effect for a year. These step change, which represent how the shipper is responding to the current market conditions, occur are different times throughout the year for different shippers. The result of cumulative shipper annual bids eventually leads to a distinct market type. Moreover, the decision to discretize the market is based on conversations with industry practitioners, who tend to refer to distinct "soft" and "tight" market periods. Our resulting market periods are defined in Table 3 below.

Table 2.3: Market Periods

	Soft	Tight
Date Range	Apr. 15, 2016 – Apr. 14, 2017	Oct. 1, 2017 – Sept. 30, 2018
Avg. Weekly PAR	81.9%	68.5%
Avg. Weekly CPL	$1155.9	$1420.7

Figure 1 shows both the average weekly aggregated PAR and the average weekly accepted cost per load (CPL) over time overlaid with our identified break points and market periods. This CPL is the linehaul price that is accepted, which includes primary carriers, backup carriers within the routing guide, and spot loads. The two measures mirror each other as expected: as the market tightens, acceptance ratios decrease and prices paid increase. As the market softens again, acceptance ratios increase and accepted prices decline.

Figure 2-1: Validated Break Dates, Market Periods, Aggregated Primary Carrier Acceptance Ratio (PAR) (left axis), and Accepted Cost Per Load (right axis)



2.4.3 S-C Relationship Measures

We select and define the S-C relationship metrics based on a combination of relationship attributes studied in the inter-firm relationship literature and in TL transportation literature, TL transportation industry reports, and conversations with our industry partner and with professionals at both shipper and carrier organizations. Thus, we have developed a set of relationship characteristics based in both theory and practice.

We define the relationship between shippers and carriers along six shipper behaviors measured in each of the market periods, The first, Primary carrier Acceptance Ratio (PAR), in the tight market is our dependent variable in our model. The remaining set of shipper behaviors is composed of pricing relative to market rates, volatility of volume offered measured, tender lead time, and origin and destination dwell times. In addition, we include one characteristic of S-C interactions across time periods (cadence of tenders), two shipper characteristics (size and industry), and two carrier characteristics (fleet size and asset- versus non-asset). We detail these measures in the following subsections and report summary statistics in Appendix C.

Primary carrier Acceptance Ratio

Primary carrier Acceptance Ratio (PAR)[1] is measured as the weekly fraction of loads that are accepted by the primary contract carrier relative to the total number of loads offered to that carrier from a specific shipper:

$$PAR_{i,j,w}^{a,b} = \frac{L_{i,j,w,accepted}^{a,b}}{L_{i,j,w,offered}^{a,b}} \tag{2.4}$$

where L is the number of loads either accepted or offered; a and b denote the shipper and carrier, respectively; i and j denote the origin and destination regions, respectively; and w is the week in which loads are offered.

High PAR in a soft period does not necessarily indicate a good S-C relationship because demand is low in soft markets and carriers search for business regardless of their relationship with shippers. On the other hand, low PAR in a soft period is a signal of a poor S-C relationship, as it is in the portion of the market cycle when carriers need loads. On aggregate, PAR decreases as the market turns from soft to tight, by definition of the market types. However, we expect that S-C pairs with an existing strong relationship will not experience a decrease in PAR as severe as that of the aggregate market PAR decline.

We calculate PAR for a S-C pair on each of their lanes, and in each week in which loads are offered between them. This value is averaged over all weeks in each of the market periods and we obtain an average weekly acceptance ratio, which is averaged again across all lanes between a shipper and carrier in each market period to obtain the S-C PAR for

[1]Primary carrier AR (PAR) is a preferred measure to AR across all carriers, where the same load may be rejected multiple time.

both periods: $PAR^{a,b}_{soft}$ and $PAR^{a,b}_{tight}$[2].

Market Rate Differential

Market Rate Differential (MRD), or the percent above or below market benchmark prices at which a load moves, measures how much of a premium (or discount) a shipper pays its primary carrier. As noted earlier, we expect that shippers paying high soft period prices, when the market is in their favor, will receive better performance in the next tight period - that is, carriers reciprocate high soft period pricing with high tight period PAR. On the other hand, we expect shippers that pay low soft market prices are taking advantage of their position and carriers respond with decreased PAR in the following tight period.

We determine MRD by calculating a benchmark price for each lane against which the S-C pairs' contract prices are to be compared. Our load data is split into the soft and tight periods to obtain lane benchmark rates for each time period. Freight transportation costs can be reduced to a linear combination of fixed costs and variable costs associated with a distance measure [53]. The benchmark rate captures the regional nature of TL pricing with lane-specific fixed effects by including origin and destination indicator variables. In doing so, we incorporate effects of outbound loads from the origin and inbound loads to the destination. These factors are important as carriers makes load acceptance decisions, in large part based on where they have available capacity, where they will ultimately end up, and the likelihood of obtaining follow-on loads.

We implement a multiple linear regression model with heteroskedastic robust standard errors in which the load linehaul price is regressed on an origin region indicator variable, a destination region indicator variable, and distance variable [18, 127].

Specifically, an indicator variable is created for all but one of the origin regions and all but one of the destination regions. The lane which corresponds to this origin-destination pair is the base case lane. The base case origin and destination regions are chosen based on their outbound and inbound volume, respectively. We measure the total outbound (inbound) volume in both of the time periods, $t \in \{Soft, Tight\}$, for each region. The base case origin

[2] We do not weight the PAR by volume offered, number of weeks in the year loads are offered, or number of lanes between the S-C pair. As a result, the relative importance of each load is higher for S-C pairs with less total volume (or fewer weeks or lanes) than that of S-C pairs with more volume (or weeks or lanes).

(destination) region is chosen if it is ranked in the top five regions by volume in both time periods. If more than one region qualifies, we take the region with the highest cumulative volume. We choose a high-demand lane as our base case to ensure it is representative of aggregate market conditions.

Our resulting base case origin region is the greater Greenville, SC area and the base case destination region is the Dallas, TX metro region. We test this choice of base case origin and destination regions among top-ranked regions by volume and find that the differences between lane benchmark values relative to one another for varying base case regions do not impact statistical results.

The vector of coefficients that results from our linear regression includes an intercept term, which is the fixed transportation cost, $\hat{\beta}^t_{base}$, a distance (i.e. variable cost) coefficient, $\hat{\beta}^t_{dist}$, I-1 origin coefficients (where I is the set of origin regions in the dataset), and J-1 destination coefficients (where J is the set of destination regions in the dataset). These 135 key economic regions defined by our industry partner are geographic clusters of transportation demand patterns. The origin and destination coefficients of our linear regression model, $\hat{\beta}^t_i$ and $\hat{\beta}^t_j$, can be considered 'price premiums', or the costs associated with an origin or destination different from the base case lane.

The origin and destination price premiums, $\hat{\beta}^t_i$ and $\hat{\beta}^t_j$, for each time period, are illustrated in Appendix B, as are $\hat{\beta}^t_{base}$ and $\hat{\beta}^t_{dist}$. In general, relative pricing of the regions hold through market periods. For example, in both the soft and tight markets, the Midwest and Southwest regions of the United States are typically higher priced regions to ship out of (i.e. positive and high origin price premiums). This is likely due to the large manufacturing hub in the Midwest the port authorities of Los Angeles and Long Beach causing high demand for outbound capacity. As a result, carriers can command higher prices. At the same time, the Northeastern region does not generate much freight volume relative to the rest of the country; carriers are willing to take lower prices in order to leave the region (i.e. negative origin price premiums in the Northeast), whereas higher prices are required to incentivize inbound carriers in the first place (demonstrated by high destination price premiums in the Northeast). Thus, these regional values define and quantify fore and backhaul lanes.

Each lane benchmark rate is calculated as follows. The base case lane benchmark price is the regression model intercept, $\hat{\beta}^t_{base}$, plus the distance coefficient, $\hat{\beta}^t_{dist}$, multiplied by the mean value of the distance variable of loads on that specific origin-destination pair, $\overline{X}_{dist}$. All other lane benchmark rates consist of the fixed cost intercept term, plus the distance coefficient and variable interaction, plus the origin and destination region coefficient price premiums corresponding to that lane, $\hat{\beta}^t_i$ and $\hat{\beta}^t_j$:

$$\hat{b}^t_{i,j} = \hat{\beta}^t_{base} + \hat{\beta}^t_{dist}\overline{X}_{dist} + \sum_{i \in I, i \neq i_{base}} \hat{\beta}^t_i X_i + \sum_{j \in J, j \neq j_{base}} \hat{\beta}^t_j X_j \tag{2.5}$$

where X_i and X_j are binary variables indicating which origin and destination is active for the lane of interest. Once the lane benchmark rates are established, the market rate differential is calculated per load, k, as the percent above or below the lane benchmark price. That is, the difference between the load linehaul price and the corresponding lane benchmark rate, relative to the benchmark rate:

$$MRD^t_{k,a,b} = \frac{LH^t_{k,a,b} - \hat{b}^t_{i,j}}{\hat{b}^t_{i,j}} \tag{2.6}$$

We obtain the market rate differential for a S-C pair, $MRD^t_{a,b}$, by averaging over all loads on each lane between the shipper and carrier, and averaging across all of their lanes for each market period separately.

In the next subsections we discuss shippers' operational performance that may impact the S-C relationship and thus carrier reciprocity in tight market conditions. A natural measure may be the the total business offered to the carrier. However, the total amount of volume a shipper tenders to a carrier does not necessarily lead to better carrier performance, as the business offered must fit the carrier's network. For example, previous literature on the benefits of information sharing finds two seemingly opposing results. In a dyadic supply chain, [75] demonstrate that supplier capacity limits the value of sharing information. However, [16] show that the benefit of information sharing is lower with suppliers with less capacity. As such, we use consistency (or CV, volatility), tender lead time, offer cadence as operational measures of good shipper tendering behavior rather than pure volume tendered.

Tendered volume volatility

TL carriers commonly call for consistency of demand at the load and lane levels. For the shipper, this means minimizing variation in tendering behaviors. Industry reports from J.B. Hunt, one of the largest TL carriers in the US, and C.H. Robinson, a leading third party logistics (3PL) provider in the US, cite increased consistency as an action shippers can take to improve their relationships with carriers, improve freight acceptance, reduce cost per load, and allow carriers and their drivers to better optimize profitability ([81] and [46]). Further, previous studies find consistency measures to be significant indicators of spot market load acceptance [131], reduced cost per load [70], and reduced routing guide failure [5].

We measure variability of volume tendered by a shipper as the coefficient of variation (CV) of the weekly volume offered. This normalized measure indicates what percent of the mean the standard variation is. It is used rather than raw standard deviation in order to compare the variability measure across dissimilar lanes. A shipper with high tendered volume variability is expected to have low PAR because inconsistency causes difficulty for carriers' network planning. As a result of high variability, either trucks will not be available when and where shippers need them, or carriers will not make an effort to ensure these inconsistent shippers are served. Thus, we expect that higher volatility from shippers leads to lower PAR in tight markets.

Tender lead time

In order for carriers to balance capacity utilization with availability, they need enough time to respond to the tender requests, ensure a truck is available, and reposition it to the pickup location on time. According to J.B. Hunt, in addition to consistency and predictability, carriers seek reasonable lead times, which allow them to create a schedule and optimize driver's hours [81]. For example, [37] found that with greater tender lead time, a shipper saw lower variability in prices paid for loads. We measure this tender lead time (TLT) as the number of days between when the load is first offered to the primary carrier to when it needs to be picked up.

It is important to note however, that with too many days of advanced notice, much can change in that time. For example, the appointment time may need to change, the carrier may not have capacity available as expected because of previous service delays, or the shipper may have been tendering the load to gather information of general carrier availability or willingness to serve for the tendered price. This point is illustrated by the numerical results of [169], which find that with a two-day TLT, the carrier's profit increases 22% over the base case of one day, but three days of advanced load notice only increases the carrier's profit by 6% over the base case. [93] and [127] study the impact of lead time on spot prices. The former consider a dummy variable for loads with greater than 8 days of lead time, while the latter finds that the impact of lead time drops off quickly for TLT longer than two days. [148] model the benefit of advanced load information sharing (i.e., TLT) to minimize carrier's total cost and find that carriers seek longer TLT. However they do not address the potential non-linearity in benefit from increased TLT addressed by some of the other studies.

Origin and destination dwell time

Carriers are often concerned with delays that occur during pickup and delivery. In a study of shipper behaviors that impact carrier performance [46] finds that carriers cited dwell times in their top shipper characteristics important to price and service decisions. Dwell time becomes even more important when one considers the regulations drivers face in terms of hours of service (HOS) laws[3]. Particularly in times when capacity is tight and carriers want to retain drivers, drivers' time utilization goes hand in hand with asset utilization and thus, carriers' attitudes toward shippers [81].

Dwell time is also considered in the literature; [169] and [148] include dwell time in their modelling approaches as a cost incurred by carriers.

Delays during delivery can often be more problematic than pickup because drivers may be

[3]In essence, a driver has 14 hours of "on-duty" time with a required 30-minute break. 11 of those 14 hours can be spent driving and 2.5 hours can be spent on all other activities, including pickup, delivery, safety inspections, and shutdown. However, it is not uncommon for live loading and unloading to take longer than 2.5 hours. When these activities do, it eats into the drivers' precious 11 hours of "on-duty, driving" time they could otherwise be using to make a profit.

able to make up some of that added time during the move. However, delays at the destination may make the driver late to the next job's appointment, which may then add further delays to that pickup, thus, proliferating the problem. We expect then, that destination dwell times may have higher impact on PAR than that at origins.

Offer cadence

In line with carriers' desire for consistency and the ability to plan for demand, the frequency, or cadence, at which shippers tender loads is often attributed to carrier's freight acceptance. In a general supplier-customer setting, [120] study inter-firm behaviors and argue that interaction frequency is a key success factor to supplier-customer relationships. [131] includes the number of days since the previous load was offered to a spot carrier by a shipper as a measure of load offer frequency. Interviews with practitioners and industry reports by J.B. Hunt and C.H. Robinson demonstrate that tender cadence, as measured by weeks in which loads are offered to the carrier, is a measure carriers use to assess shipper relationship ([81] and [46].

We measure frequency at which a shipper offers loads to a carrier as the number of weeks in a year in which the shipper offers at least one load to its primary carrier. We compare a metric, offer cadence, or the percentage of weeks during the year in which the shipper tenders loads to its primary carrier. For individual S-C pairs in our dataset, this percentage for the soft period and the tight period are highly correlated (with correlation coefficient of 0.90) so we use an average measure across the two market periods. Shippers that offer loads more frequently are expected to receive higher PAR, as the carriers rely on frequent loads to move trucks around their network and ensure trucks are available in the right location and at the right time for the portfolio of shippers served.

Table 4 summarizes the S-C attributes we consider and the literature and the academic and industry reports that have previously considered such attributes. This is not meant to be a comprehensive list, but a subset of the literature most closely related to our research question at hand.

Finally, as the main focus of our study is to examine the impact of shipper behaviors and market conditions on tight market PAR, we control the impact of other factors that

Table 2.4: S-C Relationship Metrics in the Literature

Study	PAR	Pricing	Volatility	TLT	Dwell	Cadence
[5]	✓	✓				✓
[37]		✓		✓		
[46]	✓	✓		✓	✓	
[70]	✓	✓	✓			
[81]			✓	✓	✓	✓
[83]	✓	✓	✓			
[93]				✓		
[120]						✓
[127]		✓		✓		
[131]	✓	✓	✓			✓
[148]				✓	✓	
[169]				✓	✓	
[170]	✓					

characterize the shippers and the carriers, described below.

Shipper and carrier characteristics

In this section, we describe the attributes of the shippers and carriers themselves that we include in our model to control for fixed effects. This allows us to determine if certain types of shippers or carriers behave or experience impacts from market changes differently from others. These fixed effects include shipper size, shipper industry, carrier size, and carrier service type.

Shipper size: Previous S-C relationship literature has been limited in the scope of available shipper data, however our dataset consists of hundreds of shippers. Thus, we contribute to the literature by considering the impacts of shipper characteristics. First, we consider the shipper's size, which we measure as the log (base 10) of average total annual volume tendered to all primary carriers. Larger shippers may be better insulated from the impacts of the market and maintain better carrier relationships. On the other hand, smaller shippers may invest in relationships with a smaller core set of carriers, and thus have better carrier relationships. We test if a shipper's sizes impacts how it experiences carrier performance as the markets shift.

Shipper industry: We similarly utilize the granularity of our available data by including the shipper's industry vertical in our model. Each shipper falls into only one of six industry categories: automotive, food & beverage or consumer package goods (F&B/CPG), manufacturing, paper and packaging, or a final catch-all category, other.

Carrier fleet size: Carriers with different attributes operate - and thus make decisions - differently. We first consider asset-based carriers' fleet size, measured as the log (base 10) of the carrier's tractor count, as an indicator of tight market PAR. For smaller carriers such as the single truck owner-operator, or those with only a few trucks, when demand is very high as is seen in tight markets, accepting tendered volume becomes more difficult than for larger carriers, as these small carriers have fewer available trucks to cope with the additional demand.

Carrier service type The second carrier attribute we consider is the carrier's service type. We split our data into S-C pairs for which the primary carrier is an asset-based carrier, which can be further segmented by fleet size as describe above, and a second set for which the primary carrier is a non-asset provider such as a brokerage. These providers do not own trucks, yet they may be contracted with shippers to provide services in a similar way to the asset carriers. Instead, they match their contracted shippers' needs with a vast pool of available capacity composed of many carriers. Brokerages utilize their extensive network of disagragate capacity to tailor the shipper-carrier match to the needs of both parties.

We are aware of one study that considers the differences in behaviors between asset and non-asset carriers, specifically, on spot load pricing strategies [128]. The author uses auction theory to address the differences in how asset and non-asset carriers consider the two main decisions made in auctions - whether and how much to bid. He finds that non-asset carriers bid more frequently and higher than their asset counterparts. This study, one of the few empirical studies of actual firms' decisions repeated auctions over time as the author claims, underscores our claim that asset-based carriers and non-asset carriers behave differently and justifies our consideration of them separately.

In our dataset, the non-asset carrier for these S-C pairs is defined as the brokerage itself, not the carrier that ultimately matched and moved the load. That is, we do not have visibility on which trucking company actually moved the load. Thus, we can characterize the shipper-third party relationship by including non-asset carriers separately from shipper and asset-based carrier pairs. Our dataset of S-C pairs contains 159 of these shipper-non-asset primary carrier pairs. We include the non-asset carriers in a separate model (i.e., models 1b and 2b).

2.4.4 Model Specification

As we are interested in studying the S-C relationship between different markets, our unit of analysis is at the S-C pair level and our dependent variable is tight market PAR.

As this dependent variable is a continuous fraction that can take values between and including 0 and 1 – that is, if the primary carrier accepts none or all of the tendered volume from that shipper in the market period, respectively - and we directly observe this outcome, we implement a generalized linear model. Specifically, we use a beta regression [107]. We choose a logistic link function using a maximum likelihood estimation method and robust standard errors to allow for misspecification of the prior distribution, as discussed in [109], [62] and [111].

Beta regression uses the beta distribution as the likelihood for the dependent variable:

$$f(y_i \,|\, a, b) = \frac{y_i^{a-1}(1-y_i)^{b-1}}{\mathrm{B}(a,b)} \tag{2.7}$$

Where $\mathrm{B}(\cdot)$ is the beta function defined by

$$\mathrm{B}(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{2.8}$$

a and b are the shape parameters, and Γ is the Gamma function [61].

We choose $a = 4$, $b = 1$, which fits our tight market PAR distribution with a coefficient of determination of the Q-Q plot of our tight market PAR and beta distribution of 0.931. The maximum likelihood estimators are calculated over a and b numerically through an iterative fitting process (see [105] and [95]).

The beta regression with logistic link function allows us to perform a logistic transformation of the bounded dependent variable, transforming it to the real number line, and still retain extreme values (i.e., 0 and 1) [61]. The logistic link model is defined as follows:

$$\begin{gathered} logit(y_i) = log\Big(\frac{y_i}{1-y_i}\Big) = x_i^T\boldsymbol{\beta} \\ y_i \sim \text{Beta}(a,b) \end{gathered} \tag{2.9}$$

The regression coefficients, $\boldsymbol{\beta}$, are interpreted as the log odds ratio of the dependent variable, tight market PAR, for each independent variable, x_i. We can transform the coefficients back to actual tight market PAR by exponentiating eq. (9) and solving for y_i.

As we are testing four hypotheses, we develop four models. All four models include a categorical variable for shipper industry vertical, and omit Manufacturing as our baseline. Models 1a and 2a test whether asset-based primary carriers consider previous soft period shipper behaviors in tight market period freight acceptance decisions (H1a) and whether they consider current market tight market period shipper behavior - that is, if they are myopic (H2a), respectively. These models include the continuous variable, carrier fleet size, as a predictor of tight market period PAR.

Models 1b and 2b correspond to hypotheses H1b and H2b, respectively. As such, rather than carrier fleet size, these models include a binary variable indicating whether the carrier is asset-based - coded as a 1 - or not - coded as a 0. This means that the reported coefficient for the asset binary variable in these models is associated with the additional (log odds ratio of) tight market PAR that would result if the carrier is asset-based (i.e., $x_{asset} = 1$ rather than 0).

2.5 Results

In this section, we first discuss whether carriers have long memories; that is, the results of the beta regression model of tight period PAR on soft market shipper behaviors (i.e., models 1a and 1b). We then discuss whether carriers are responding to current tight market shipper behaviors with the results of models 2a and 2b.

2.5.1 Carriers' responses to previous shipper behavior

Table 5 reports the results of the beta regression analysis of soft market shipper behaviors on tight market PAR for asset-based primary carriers and for all primary service provider types (models 1a and 1b for H1: Do carriers remember shippers' previous soft market behavior?).

Hypothesis 1a is not supported, and except for one soft period shipper behavior, hypothesis 1b is also not supported. This result means that in general, shippers' soft market period behaviors do not significantly impact their primary carriers' tight market period PAR. However, soft market PAR (which is not a shipper behavior, but rather an indicator of previous *carrier* behavior) is a predictor of tight period PAR. This suggests that on average, primary carriers tend to maintain their acceptance performance across markets.

The results of model 1b suggest that for non-asset primary carriers, destination dwell time in the soft period as a significant predictor of tight period PAR. This finding, that non-asset service providers may be more sensitive to soft market period destination dwell time than asset-based carriers, is explained by our industry partner. For asset-based carriers, one driver may handle the long-haul move to get the load into the destination region, drop the trailer at a regional facility or distribution center, and move on to the next load. A regional driver will then take the loaded trailer to the final destination. In soft markets, asset utilization is low and asset carriers may have the flexibility needed to have multiple trucks handle a single load than their non-asset counterparts, which have less control over shuffling drivers and trucks around. This may explain the greater sensitivity to destination dwell time in the soft market observed when non-asset carriers are included in the analysis of models 1b as compared to model 1a.

A number of control factors that describe the S-C interaction across market periods, shipper-specific characteristics, and carrier-specific characteristics are statistically significant in models 1a and 1b. These factors are also significant in models 2a and 2b, and we discuss them further in section 5.2, using coefficients found from model 2a.

The main result of H1 is that primary carriers do no remember how shippers treated them in the past, when markets favored the shippers and they may have acted opportunistically. Put another way, shippers do not benefit from higher PAR in tight markets for having good performance or competitive pricing in previous soft markets.

Table 2.5: Model 1 results, soft market shipper behaviors

Variable	Model 1a Asset-based primary carriers	Model 1b Non-asset primary carriers
Constant	-0.4818	2.681
	(0.6803)	(1.209)
PAR_{soft}	2.630***	1.275**
	(0.3574)	(0.6828)
MRD_{soft}	0.0067	0.0111
	(0.0040)	(0.0091)
CV_{soft}	0.0334	-0.1092
	(0.2708)	(0.4889)
TLT_{soft}	-0.0255	-0.0251
	(0.0222)	(0.0589)
Origin dwell$_{soft}$	0.0288	-0.0606
	(0.0350)	(0.1515)
Destination dwell$_{soft}$	-0.0664	-0.3592**
	(0.0563)	(0.1546)
Offer cadence	0.8143***	0.4902
	(0.2421)	(0.4515)
Log shipper volume	-0.1462**	-0.1692*
	(0.0630)	(0.1029)
Automotive	0.9791***	-1.057**
	(0.2370)	(0.5292)
F&B/CPG	0.2416	-0.8392*
	(0.1839)	(0.4945)
Paper & Packaging	0.0261	-1.330***
	(0.2084)	(0.5355)
Other	0.3821	-1.096
	(0.3851)	(0.7793)
Log carrier fleet size	-0.0449	NA
	(0.0327)	NA

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

2.5.2 Carriers' responses to current shipper behavior

Next, we discuss the results of models 2a and 2b, which test whether tight market shipper behaviors impact tight market PAR (H2).

The results, summarized in Table 6 demonstrate that H2a and H2b are supported: carriers respond to current tight market period shipper behaviors. In particular, tight period pricing (i.e., market rate differential), consistency and cadence of tendering patterns, and destination dwell times in tight market periods have significant impacts on PAR in the same tight market period.

Table 2.6: Model 2 results, tight market shipper behaviors

Variable	Model 2a Asset-based primary carriers	Model 2b Non-asset primary carriers
Constant	2.1214***	3.583***
	(0.5835)	(0.1.113)
MRD_{tight}	0.0250***	0.0461***
	(0.0043)	(0.0088)
CV_{tight}	-0.7876***	-0.8715*
	(0.2655)	(0.4885)
TLT_{tight}	-0.0115	0.0989
	(0.0216)	(0.0969)
Origin dwell_{tight}	0.0388	-0.1997
	(0.0327)	(0.1308)
Destination dwell_{tight}	-0.1572***	-0.2955**
	(0.0492)	(0.1345)
Offer cadence	1.0986***	1.144**
	(0.2474)	(0.4922)
Log shipper volume	-0.1317**	-0.1433
	(0.0633)	(0.1244)
Automotive	1.0406***	-1.002**
	(0.2234)	(0.5167)
F&B/CPG	0.2750	-0.9841*
	(0.1821)	(0.5622)
Paper & Packaging	-0.0350	-1.288**
	(0.2115)	(0.5275)
Other	0.0550	0.7379***
	(0.3719)	(0.7856)
Log carrier fleet size	-0.0523	NA
	(0.0310)	NA

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

While the intent of this study is to indicate which shipper behaviors, S-C relationship measures, and shipper and carrier characteristics in different market periods impact tight market PAR, not to predict tight PAR, in this section we illustrate the relative impact of the significant variables with the models' resulting coefficients and offer some industry context. However, it is important to note that these models and their quantitative results have low predictive power.

As an illustration of model coefficients interpretation, the intercept term of model 2a, 2.1214, indicates that the model predicts tight period PAR to be 0.8932, or 89.32% of tendered volume is accepted by the primary carrier. We obtain this result by plugging into eq.(9) and solving for PAR:

$$log\Big(\frac{PAR_{tight}}{1 - PAR_{tight}}\Big) = 2.1214$$

Shippers that pay their primary carriers above tight market rates on average observe higher tight market PAR. Again, we plug in the MRD_{tight} coefficient, 0.0250 with the constant coefficient, and solve for PAR:

$$log\Big(\frac{PAR_{tight}}{1 - PAR_{tight}}\Big) = 2.1214 + 0.0250$$

which gives us $PAR_{tight} = 95.95\%$. This means that all else held equal, an increase of 1 percentage point in MRD_{tight} results in an increase of 6.65 percentage points in PAR_{tight}. Discussions with shippers and carriers indicate that reasonable pricing changes may be in the range of 1-3%.

Next, we observe that shippers' volatility in tight market period tendering behavior (i.e., CV of weekly offered volume) negatively impacts tight PAR. Recall that CV measures the ratio of standard deviation to mean of weekly tendered volume, as a percentage. Thus, one percentage point increase in tight market CV corresponds to an 11.3% decrease in PAR (from 89.32% to 79.19%). This result is consistent with previous literature ([131], and [5]), which study load- and lane-level freight acceptance.

The results of model 2a further suggest that carriers are sensitive to tight period destination dwell time: an additional hour of tight market dwell time that a primary carrier experiences results in a 1.75% decrease in tight market PAR (from 89.32% to 87.73%). Moreover, not only are shippers with high destination dwell times susceptible to reduced PAR, but they likely incur detention fees at facilities with excessive wait times. Our analysis, however, focuses on linehaul prices rather than these accessorial charges.

Some shipper and carrier characteristic are found to be statistically significant in all four models. Of these interaction factors across market periods, the cadence of tenders offered, is a significant contributor to tight market PAR. For example, using results from model 2a, a 10% increase in offer cadence – which corresponds to 5.2 additional weeks per 52-week market period – corresponds to a 7.69% increase in tight PAR (from 89.32% to 96.17%).

The results of all four models indicate that shipper's size is an indicator for its primary carriers' tight market PAR. The coefficient for the log of the shipper's annual volume is negative, which implies that larger shippers tend to experience lower tight market PAR. Continuing with model 2a coefficient results, a shipper that is 10 times larger than a fellow shipper experiences tight period PAR that is 1.46% lower.

In addition, results of models 1a, 2a, and 2b suggest that industry vertical impacts tight market period PAR. Automotive shippers tend to see higher tight market PAR than shippers in the manufacturing industry (recall that the base case industry vertical – i.e., the variable which is omitted from the categorical industry vertical variable in the regression – is Manufacturing). In fact, model 2b indicates that except for those in the paper and packaging industry all industry verticals see higher tight period PAR from non-asst carriers than those in the manufacturing sector.

The numerical illustrations discussed above demonstrate that, while certain tight period shipper behaviors and characteristics impact tight period PAR, in isolation, most of these behaviors may have small impacts on aggregate tight period PAR. In combination, however, they may have greater impact.

In summary, the combination of results that do not support H1a or H1b but do support H2a and H2b, suggest that carriers are myopic: they do not reciprocate good behaviors from shippers in previous soft markets with higher PAR in the following tight market. However they do offer higher PAR in the tight market to shippers that pay higher prices and offer loads more consistently *in the same market period*.

2.5.3 Backup carrier price premium

The preceding results indicate that shippers that pay below market prices in the tight market period observe lower PAR in the same tight periods. These shippers may consider paying more in the tight period to improve PAR. How much the shipper should be willing to pay above its existing contract prices is its Backup Premium: the percent of contract prices the shipper ultimately pays its backup carriers that accept the load (regardless of routing guide depth) or spot carriers as a result of primary carrier rejections:

$$Backup\ Premium = \frac{E[Accepted\ Linehaul\ Price|Primary\ Carrier\ Rejects]}{E[Accepted\ Linehaul\ Price|Primary\ Carrier\ Accepts]} \tag{2.10}$$

A Backup Premium value less than 1 indicates contract prices of primary carriers that accept loads are higher than the prices backup carriers charge, or that the shipper often ends up on the spot market and finds capacity at lower prices than contract prices. A Backup Premium of less than 1 may also indicate that loads tend to be rejected by primary carriers on the shippers' lower priced lanes more often than on its higher priced lanes, as the prices used for the Backup Premium are linehaul prices rather than the normalized MRD prices.

The average Backup Premium for shippers in the soft market is 1.009. That is, on average, shippers pay 0.9% more than their anticipated contract prices as a result of primary carrier rejections in the soft market. However, in the tight market, shippers' average Backup Premium is 1.182. Thus, primary carrier rejections cost shippers 18.2% more than their contract prices in the tight market. This further demonstrates that not only is primary carrier freight acceptance lower in tight markets, but shippers pay the price for it. This may help shippers negotiate with primary carriers leading into tight markets to increase prices (i.e., MRD), but not by more than their expected Backup Premium.

2.5.4 Summary of findings

Our results suggest that carriers have short memories - they act like goldfish - when market conditions change. Shippers that "pay it forward" in soft markets with high pricing relative to the market or consistent volume do not necessarily reap better carrier behavior (i.e. higher PAR) when markets tighten. Put another way, shippers that pay above market prices in the soft market are not more likely to see high PAR in the tight period. In addition, the converse holds: shippers that pay below market prices in soft market periods are not more likely to see PAR decrease as the market tightens. Instead, carriers appear to be myopic. They respond to higher current market period pricing, consistent and frequent load tendering, and low dwell times with higher PAR.

We find that larger shippers are more likely to see lower tight market PAR. This may be explained by the fact that larger shippers interact with more primary carriers, some of which may have a good relationship with the shipper and continue to provide service during tight markets with higher PAR, while other carriers - perhaps those with lower prices or those that serve more volatile lanes or facilities with slower drop off times - may not prioritize the shipper's freight in the same way. Thus, this spread in carrier response to their shippers may be wider for larger shippers and contribute to the negative relationship between shipper size and tight market PAR. While of course this is not a characteristic shippers can change, it is an important confounding factor to include in the model and allow us to understand how different segments of the industry experience market cycles.

2.6 Implications

In this chapter, we implement a hybrid approach to detect market cycles in TL transportation time series data, which combines structural stability tests and a forward- and backward-looking scoring process. We partition the TL transportation market into two distinct time periods in which the power differential between players shifts from shippers to carriers. Despite prevalent discussions in the literature of the impacts of market conditions on transportation partners, none have considered the impacts of behaviors in one market on behaviors in future market periods.

Our results extend previous literature on S-C relationships, in particular [131] and [170] in two ways. First, we verify previous authors' results (each of which consider a single shipper) with our empirical study, which generalizes these findings over a large number of US shippers and carriers across different industry verticals and shipper sizes. Second, we follow individual S-C pairs' interactions over time and include shippers behaviors during markets that favor them in the model of carriers' acceptance decisions when the market is in their favor. We find that there is a propensity for carriers to forget shippers' previous market period behaviors; carriers tend to respond to shippers' behaviors and performance in the current market period rather than based on previously demonstrated behavior.

These results imply that for shippers, it is not necessary to "pay it forward" with exceptional performance or pricing in soft markets, but it is important to prioritize particular behaviors in the tight market to ensure high PAR. First, shippers should price competitively relative to market rates. While there are likely diminishing returns to establishing contract prices that are well above market prices, as is suggested both by our empirical analysis and by our industry partner's experience, paying slightly above market prices does result in higher PAR. Shippers may refer to their historical Backup Premiums, particularly at the lane level, to guide the extent of appropriate price increases.

Second, shippers should seek ways to enable consistent tendering behavior. This may include improving forecasting models, contracting a larger set of carriers on particularly volatile lanes (e.g., split the demand on lanes to multiple primary carriers) to smooth out volatility, or bundling lanes for primary carriers such that the volatility of the package of lanes served by each carrier is low. Similarly, shippers can package low volume lanes (i.e., those with low offer cadence) with higher volume lanes to a single carrier so that the cadence of tenders of a set of lanes is more attractive to that carrier.

Finally, shippers' destination facilities must improve dwell times in order to ensure high tight period PAR. The importance of this behavior in particular is underscored by its significance in both the models that test soft market behaviors and tight market behaviors. To reduce dwell times, shippers may improve appointment scheduling, ensure adequate staffing during peak delivery times, expand drop trailer opportunities (as opposed to live unloads, which naturally take longer), or clarify instructions or maps at larger facilities to help drivers find the right loading dock quickly. Driver experiences can make or break shipper reputations across the carrier community. Addressing this problem may be particularly difficult for instances in which loads are being shipped to a downstream customer; the shipper does not own the destination facility and is unable to adjust operations. In these cases, shippers may need to work with their consignees, or recipients, to reduce destination dwell times.

2.7 Limitations and Future Research

As dynamic market conditions impact overall business operations for shippers and carriers, and both sides report interest in developing relationships with one another to ensure better performance, this and future related research have both academic and practical implications. As noted earlier, our model allows us to determine which market-specific behaviors, S-C relationship factors, and shipper and carrier characteristics are important in determining tight market PAR. However, the predictive power is low for our model and as such, we are limited in our ability to quantify direct impact of changing behaviors. Future research may build on these findings and propose a more game theoretic model of shipper-carrier reciprocity to analyze and quantify the impact of specific decisions and behaviors.

A second limitation of this work is that we have limited knowledge of carriers' behaviors outside of the load accept or reject decision. While we know whether the carrier is primary or not with a given shipper, the contracted price, and actual volume tendered and accepted, we do not know what other lanes and shippers the carrier serves exogenous to the given dataset, how well the business realized in our dataset fits into the carriers' overall network, or even expected demand from a shipper on each lane. Shippers' bid cycles occur at effectively random times (from the carriers' perspective) throughout the year and our data includes neither bid timing nor contracted volume. We inherently assume that volume offered to a carrier is approximately that which has been agreed upon during the procurement process. As such, we do not account for surge or unexpected volume from a shipper for which carriers have not planned outside of the effects that are captured by our volatility of volume offered metric. In such cases, our measure of carrier PAR reflects poorly on the carrier as opposed to on the shipper. Future studies may consider how acceptance ratios and routing guide compliance change relative to contracted and unplanned freight volumes.

Finally, in order to retain S-C pairs that interacted in both market periods we removed S-C pairs that interacted in the soft market period but not in the tight market period. There are a multitude of reasons why a S-C pair may interact in one year but not in the future. The shipper may have run a procurement event between the two time periods and for any number of relationship reasons or business fit reasons, the carrier may not have submitted a bid to serve that shipper anymore. This is of course a carrier decision. Or in the same scenario, perhaps the carrier did submit a bid but the shipper chose another carrier to serve the lanes that the first carrier had been serving in the past. This is the shipper's decision. Finally, if the carrier does not show up in the data again with that shipper, it could be because a load was not tendered to the carrier by the shipper. In this case again, it is the shipper's decision not to interact with the carrier. However, we don't have enough information to discern which scenario is or make appropriate assumptions regarding which partner chose to terminate the relationship. As such, we must eliminate these S-C pairs.

The authors of this study will consider a full second soft market period and address the converse research question: do shippers act like elephants or goldfish? Should carriers "pay it forward" by adhering to contract commitments in tight periods to ensure shippers reciprocate by upholding their end of the bargain when market soften again?

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 3

Opportunities for market-based freight contracts: a Segmentation Approach

Abstract

In the for-hire truckload market, firms often experience unexpected transportation cost increases due to contracted transportation service provider (carrier) load rejections. The dominant procurement strategy results in long-term, fixed-price contracts that become obsolete as transportation providers' networks change and freight markets fluctuate between times of over and under supply. We build behavioral models of the contracted carrier's load acceptance decision under two distinct freight market conditions based on empirical load transaction data. With the results, we quantify carriers' likelihood of sticking to the contract as their best known alternative priced load options increase and become more attractive; in other words, carriers' contract price stickiness. Finally, we explore carriers' contract price stickiness for different lane, freight, and carrier segments and offer insights for shippers to identify where they can expect to see substantial improvement in contracted carrier load acceptance as they consider alternative, market-based pricing strategies.

3.1 Introduction

As discussed in Chapter 1, contracts between shippers and their for-hire TL carriers are non-binding on both sides. When the shipper's tendering patterns are difficult to manage or the contract prices of the loads are not competitive enough with the market - as we demonstrate in Chapter 2 - contracted carriers tend to reject loads more frequently. In addition, when general freight markets become constrained during tight market periods, carriers are more likely to reject shipper's unattractive, low-priced freight in favor of higher-priced alternatives.

When primary, contracted carriers reject loads, shippers then tender the load to backup carriers in their routing guide for that lane. If the list of backup carriers is exhausted, the shipper may have to rely on the spot market. Shippers can find carriers that have an available truck going from the region in which their load needs to be picked up and going to the region their load needs to be dropped off.

Shippers find a carrier on the spot market for independent transactions, but there is no single spot market price at any given time, even for spot loads on the same lane. Instead, when we refer to the "spot market price" it is the average of a range of prices. Figure 3-1 illustrates this point.

It depicts actual loads fulfilled on the spot market from 6 shippers on a single lane in October 2017. Each day, there may be one, many, or no loads that require spot capacity. Each load is fulfilled at a different price set by the individual carrier and depends on the carrier's network structure at that moment. In the remainder of this study, we discuss the dynamic average lane-specific spot price, depicted in the figure, which represents an underlying distribution of individual realized spot load prices.

3.1.1 Contractual relationships and the best known alternative

Spot market prices are highly volatile and represent the dynamic nature of immediate balance between trucking supply and demand. For shippers and carriers alike, the spot market represents alternative price options to a contracted load. When trending spot prices are high relative to contract prices, carriers may be tempted away from adhering to their contractual

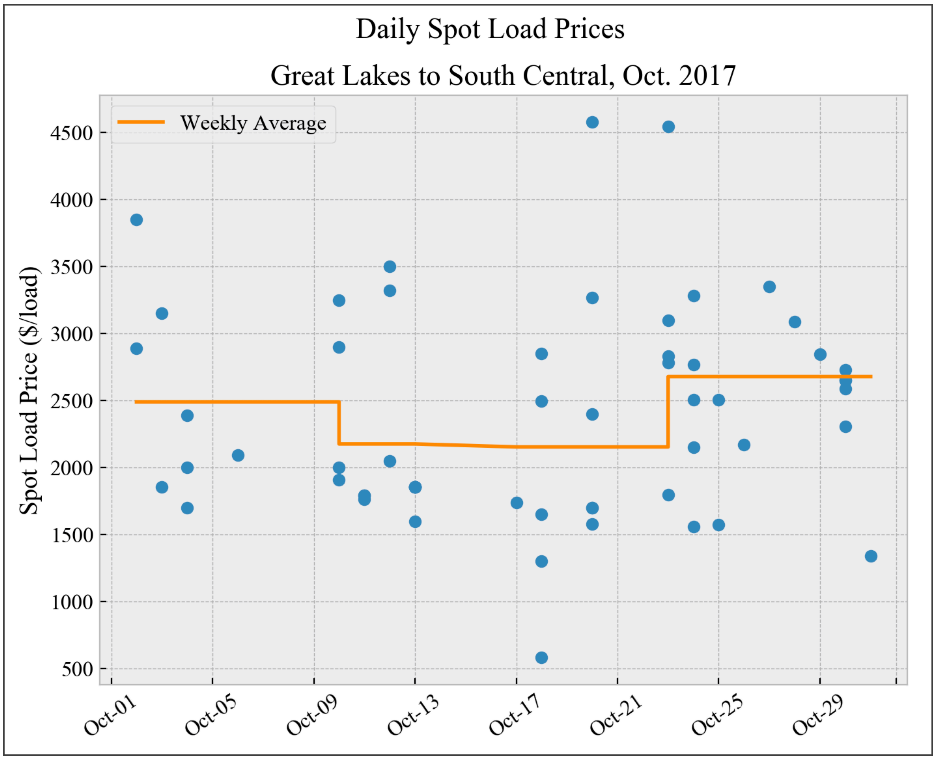


Figure 3-1: Daily Spot Prices for Single Lane

agreements. A primary carrier can either uphold its contractual commitments with the shipper by accepting loads and maintaining the relationship, or it can act opportunistically by rejecting loads and offering its capacity on the spot market with higher expected profit. However, by reneging on its contract and rejecting loads, the carrier risks potential future business from that shipper.

About 72% of for-hire TL freight is accepted by the primary carrier, while about 5% is fulfilled by spot capacity [40]. The remaining 23% is moved by backup carriers within the shipper's routing guide. When spot market prices are higher than average contract prices - a condition referred to as a tight market - shippers typically pay a price premium of 35% above the contract price if they end up using spot carriers [5].

Even when backup carrier prices or spot market prices are below a shipper's contracted carrier's price, there are substantial benefits to working with primary carriers rather than backup or spot carriers. The shipper's RFP is a vetting process. Over the course of weeks, sometimes months, the shipper and its bidding carriers not only communicate expected demand, service expectations, and pricing, but the carriers have also have demonstrated fast and simple communication processes, and technological sophistication to integrate with the shipper's electronic data interchange (EDI) and payment systems.

An account manager is assigned by each of the contracting parties. This person continuously evaluates performance over the course of the contract and is the point of contact when performance issues arise. As a result, the value for the shipper of working with a primary carrier includes the ease of business processes and relationship developed. Due to its transactional nature, none of these process or communication channels are established for spot loads. Furthermore, the spot market price is not known to the shipper in advance. Thus, both service level and actual price to be paid to spot market carriers are largely uncertain. Therefore, it is typically in shippers' best interest to encourage primary carriers to accept loads, even if the spot market price is expected to be better (i.e., lower) than contract prices.

As freight markets fluctuate between periods of over and under supply - soft and tight markets, respectively - shippers' fixed contract prices can become stale. As a result, the amount of freight that is accepted and moved at the original contracted price declines over the course of the contract. To mitigate this issue and ensure contract prices remain market competitive, there has been growing interest from both shippers and carriers to explore market-based pricing into their portfolio of TL freight contracts. Understanding carriers' willingness to stick to (or defect from) their contracted load price helps practitioners identify the most promising network, lane, freight, and carrier segments for this market-driven approach.

The remainder of this chapter is structured as follows. Section 3.2 discusses practitioners' considerations regarding index-based contracts, Section 3.3 offers a review of the relevant literature, Section 3.4 summarizes our research question and the hypotheses tested, and Section 3.5 describes our empirical dataset and model specifications. We present the results of the models in Section 3.6 and their implications, particularly for market-based pricing consideration, in Section 3.7. We conclude in Section 3.8 with a discussion of limitations of this research and areas for further exploration.

3.2 Market-based Freight Contracts

For every load a contracted carrier is tendered on a contracted lane, it can either accept the load or reject it and fill available capacity with a load on the spot market. Carriers may also have other contracted shippers on the same or similar lanes and thus other contracted business available as an alternative option. However, because the shipper has no knowledge of its contracted carriers' other customers and because spot prices and contract prices move up and down together, albeit with some lag (see [113]), we use the average lane-specific spot market price to represent carriers' alternative options.

As the average price of spot market loads available to the carrier increases relative to the contracted price, an opportunistic carrier with low contract price stickiness will be incentivized away from accepting the contracted shippers' loads. For each load offered to a contracted carrier, we calculate its Spot Rate Differential (SRD), or how much the current lane-specific spot price is above or below the load's contract price, as a percentage:

$$SRD_{k,i,j,t} = \frac{(SpotPrice_{i,j,t} - ContractPrice_{k,i,j,t})}{ContractPrice_{k,i,j,t}} \times 100 \tag{3.1}$$

where k is the load tendered to the carrier on lane with origin i and destination j at time t.

This Spot Rate Differential is the percent difference between the contract price and spot market price at a given time and the key metric for carrier contract price stickiness. Building off of the literature described in Subsection 3.3.3 we define contract price stickiness as the rate at which the contract price must change (relative to the current lane-specific spot price) as factors exogenous to the contract change.

A market-based pricing strategy aims to appeal to carriers with low contract price stickiness. Market- or index-based pricing is commonly used in industrial, agricultural, and energy commodities markets. Similar to the truckload market, these products see cyclical supply and demand fluctuations. Market-based pricing helps simplify price negotiations and increase transparency for long-term contracts between sellers and buyers, particularly when price volatility over time is a concern [136]. Interest in indexed pricing has grown in the truckload freight market as both shippers and carriers seek ways to mitigate the risks incurred by freight market cycles.

Figure 3-2 demonstrates the dynamic nature of the truckload freight market. We present the Truckload Linehaul Index reported by Cass Information Systems, Inc. - a leading industry provider of information and payment solutions - and national average contract and spot market prices from our empirical dataset from September 2015 to January 2020.[1].

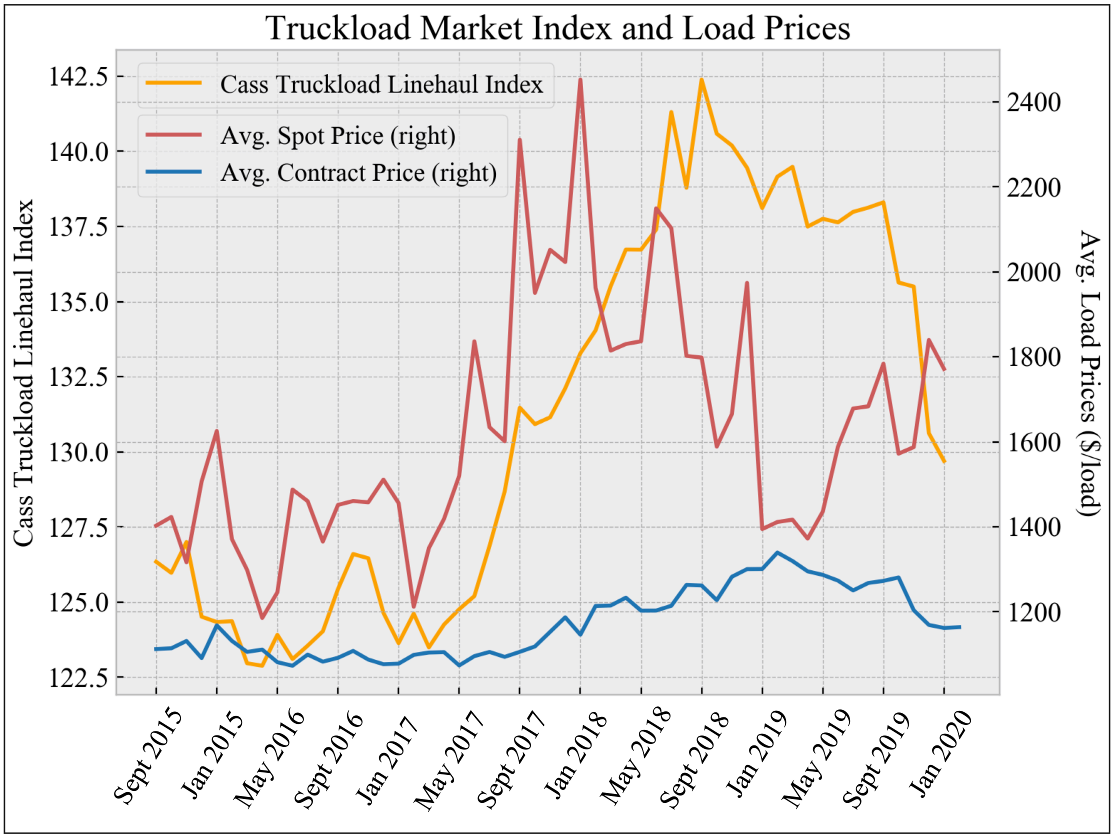


Figure 3-2: Industry Index, Spot, and Contract prices

The variations in market index and spot and contract prices over this time period reflect both soft and tight market periods. Until July of 2017, the industry experienced a soft market, where we see low prices and index values. Following this time, the industry experienced a very tight market, with high prices and severely deteriorated primary carrier acceptance rates. In fact, in the first soft market period, average primary carrier acceptance rates were 81.9%. This number dropped to 68.5% in the subsequent tight market period. In Chapter 2

[1]The Cass Truckload Linehaul Index is comprised of 95% contract load transaction data and 5% spot load data. Discussion of how our empirical dataset appropriately represents dynamics in the truckload freight market is found in Section 3.5

we provide quantitative justification of these market periods and, along with [5], [139], and [2], analyze market-driven primary carrier acceptance rates.

As the freight market cycles between soft and tight periods, index-based pricing strategies aim to allow contract prices to adjust automatically with changes in the broad market. Otherwise, shippers manually adjust prices through full RFPs or "mini-bids" (targeted carrier or lane price adjustments specifically focused on under-performing segments, typically resulting in short-term fixed-price agreements). With index-based strategies, the contracted price between a shipper and primary carrier increases or decreases with a market-based indicator such as those reported by Cass Information Systems Inc. or DAT Freight & Analytics.

A handful of large shippers that operate in the consumer packaged goods, food and beverage, and manufacturing industries, and their carriers (both asset and non-asset) have explored index-based pricing strategies in practice. Those that have to date, however, have only done so on a small set of trial or pilot lanes [137]. In addition, some carriers, particularly 3PLs and brokerages (i.e., non-asset providers), have dedicated resources to offer dynamic pricing options for their shipper customers (see [50], [125], [150], and [26]).

Despite the potential benefits of market-based pricing strategies, the advantages of traditional fixed-price contracts remain. Shippers and carriers alike seek consistency and predictability - both in terms of demand or capacity availability and prices. Shippers run annual RFPs to establish fixed contract prices around which they can budget. Carriers intend to provide high-quality service to their customers but need to know they will be able to cover their costs in the process. Keeping this in mind, we aim to help shippers and carriers understand where this traditional fixed-price approach is effective and where an alternative market-driven approach may be beneficial.

3.3 Literature Review

In this research, we quantify carrier's willingness to stick to contracts by constructing empirical behavioral models that include industry dynamics not captured in previous literature.

3.3.1 Supply chain contracts

Much of the extant literature on supply chain contracting explores the ways contracts help coordinate or share risks between buyer and supplier [87]. Agents seek risk-sharing contracts to encourage both sides to remain committed to the contract terms through buy-back [110, 56], revenue sharing [35], and options [19] contracts. (See [34] and [134] for surveys of these contracting mechanisms.)

The non-binding TL freight contract is similar to the self-enforcing agreements studied in the context of repeated games literature, which explores agreements that result in both parties maintaining the terms of the agreement over time without the interference of an external party [145]. Instead, informal mechanisms are used for contract enforcement. As the transportation provider considers each load, if it sticks to the contract by accepting the load, it maintains that relationship with its shipper. On the other hand, the transportation provider can defect from the contract, reject the load, and damage the relationship with each rejection (see [129]). This self-interested, opportunistic behavior that violates the existing contract may be influenced to higher priced alternative options on an external spot market [159, 160, 157]. However, the value of future potential business - i.e., the shadow of the future - encourages carrier contract compliance [74, 108].

For shippers, the incentive to stick to their contracts stems from the initial investment and vetting involved with securing the contracted carriers in the first place. The procurement process discussed in Chapter 1 takes months of preparation, communication, and analysis to establish contracted carriers on each lane. Transaction Cost Theory considers the costs incurred as a result of the information search, negotiation, and monitoring involved in the exchange of a good or service between agents – effectively, the frictions of a transaction [160, 162]. The theory justifies the use of contracts as a means of defining the terms of inter-firm agreements [163], made particularly important with uncertainty and high frequency of interactions [162]. [101] underscores the benefit of contract use, noting that the contract collects many repeated heterogeneous transactions under one common pricing policy, reducing the transaction costs needed to repeatedly negotiate prices for each transaction.

3.3.2 Shipper-carrier relationships

We further motivate the research by expanding the extant body of literature on shipper-carrier relationships, specifically that on primary carrier load acceptance decisions. Rather than direct measures of the shipper-carrier interactions, however, much of the literature considers attributes of the lanes and freight to determine primary carrier acceptance rate (PAR). PAR is measured as the percentage of loads a primary carrier accepts relative to the number of loads it is tendered on its contracted lanes. High lane volume [70], low lane volume volatility [83], high prices [8], and high lane consistency, or cadence, [5] have been found to be positively correlated with higher PAR, all else equal.

The shipper-carrier relationship itself, however, has been found to contribute to carrier PAR in a few studies. [130] analyze contract and spot market transactions and the impact of the history of the shipper-carrier relationship on freight acceptance. The authors find that less frequent load offers increase the likelihood of a primary carrier rejecting a load, while higher offered volume, lower load offer volatility, and higher revenue transacted between the shipper and carrier increase the likelihood of carrier's load acceptance. [170] define a "good" shipper-carrier relationship as one in which a contract is in place and find that contracted carriers outperform non-contracted carriers in terms of freight acceptance, on-time delivery, and pre-positioned capacity. These two studies each obtain data from a single shipper, which limits their ability to generalize across types of shippers or segment their datasets to offer insights specific for types of shippers' business.

The impact of market dynamics on carriers' freight acceptance decision is studied in Chapter 2 [2]). The authors consider two distinct market conditions: soft, where shippers have low demand relative to available capacity and thus corresponding low contract and spot prices, and tight, where demand for TL capacity outstrips supply and prices are high (spot prices typically exceed contract prices). The authors find one of the main contributor to primary carriers' tight market acceptance decision to be how competitive the shippers' contracted load offer price is with prevailing market prices. [130] also consider a measure of market condition - Spot Premium - in determining carrier load acceptance decisions and find that carriers consider overall market conditions and alternative priced loads when making contracted load acceptance decisions.

The nature of existing TL contracts, external market conditions, and their impact on carriers' load acceptance decisions are studied in [129]. The authors find that explicit contracts (those that have more formally defined service level expectations) as compared to implicit contracts illicit higher primary carrier acceptance rates from carriers. In addition, and similar to the findings by in chapter 2, as the market tightens and becomes more attractive to carriers, the benefit seen by these explicit contracts diminishes.

Finally, [94] consider TL carriers' reservation prices and willingness to accept loads. Using spot market prices and a discrete choice survey experiment, the authors find that carriers' contracted load acceptance decisions are impacted by higher-priced hypothetical alternative load options. We expand this research by using empirical data to study actual carrier decisions and explore how these decisions differ across shippers' freight networks, carriers' service types, and overall freight market conditions.

3.3.3 Supplier price stickiness

While our research builds on the transportation literature, we also draw from that which models price stickiness - in other words, the rate at which prices change in response to internal firm or external market dynamics. Much of this literature focuses on how suppliers' prices change as a collective at the industry level, rather than the micro, firm level.

The most relevant stream of literature to our research consists of a set of studies focused on producer pricing stickiness to customer demand patterns or to external market prices. Manufacturers adjust prices due to underlying costs and customer demand changes [97] and beliefs that competitors are also changing prices (known as coordination failure) ([24] and [25]).

This coordination failure comes into play in the TL context during the strategic RFP stage. Carriers do not coordinate with one another on pricing of lanes. However, while carriers submit lane bid prices according to their own network fit, if they want to be competitive, they must also factor in their beliefs about how other carriers may be pricing the lanes they want. Submitting bids that are too high may mean another carrier is awarded the business; too low and they may be chosen for the contract, but at a price that cuts into critical profit margins. This coordination failure may result in a race to the bottom and cause the shipper to award poor-fitting carriers to lanes.

Studies by [59] and [60] use a survey of over 11,000 firms across Europe and find that most firms consider both historical market information and expectations of the market when making pricing decisions. The authors conclude that not only is price stickiness influenced by market conditions, it is also related to customer relationships as defined by explicit and implicit contracts. These findings underscore truckload practitioner sentiments and findings by [129] that a combination of market conditions and standing shipper-carrier relationships are key determinants for carrier pricing and load acceptance decisions.

The above studies also find that individual firms' rate of price change tends to be slower than that of the general market. This time- and market state-dependent supplier price stickiness is further explored by [96], [66], and [11]. Similar results are observed in the TL freight context as carriers' contract price changes tend to lag those of spot prices (see [113]). In addition, contract freight rates rise more quickly than they fall.

The extant literature shines light on how producers make (predominantly strategic) pricing decisions. However, there are few studies of if or how suppliers choose to offer their product or service at an operational level to contracted customers or to non-contracted customers on a spot-like market. We aim to add to the literature in this way.

3.4 Research Question and Hypotheses

As discussed above, the long-term, fixed-price contracts between shippers and their TL carriers are non-binding in volume and capacity commitments. They often result in degraded performance and unexpected price escalations as exogenous market conditions change. Alternative contract forms, in particular those based on indexed pricing, have drawn recent attention. With this research, we aim to help shippers and carriers determine the most promising areas within their networks for index-based pricing in their portfolio of freight contracts. Thus, we address the following research question:

For which segments of shippers' networks - lanes, volume, demand patterns,

and carriers - should they consider index-based contracts?

To answer this question, we model primary carrier load acceptance decisions based on the load's offered price relative to the carrier's best-known alternative option: the current lane-specific spot market price. In this way, we can demonstrate carriers' contract price stickiness. We formulate our first hypothesis as:

H1: A load is more likely to be accepted by a primary carrier the higher the contracted price is relative to the current, lane-specific spot market price.

This claim is the basis for our behavioral assumptions that carriers prefer higher priced loads [159, 157].

Next, we consider carriers' willingness to stick with contract prices for different lane and freight types. We measure how primary carriers' likelihood of accepting a load is impacted as the offered price of the load changes relative to the spot market price for these lane and freight segments of interest. First, we consider a lane's distance. While carriers each have individual preferences and strategies relating lane distances, we aim to isolate segments to determine general patterns. We define the following hypothesis:

H2: The likelihood a primary carrier accepts a load increases with lane distance and further increases as the load's contracted price increases relative to the current, lane-specific spot price.

Next, we consider how price adjustments impact primary carrier load acceptance for different lane demand patterns. Shippers' infrequent and inconsistent tendering behaviors lead to lower primary carrier acceptance (see Section 3.3.2). For lanes on which a shipper tenders loads inconsistently, it may be more difficult or require longer distances driven empty for a carrier to re-position capacity when a load is tendered. A higher price incentive may improve the probability the carrier accepts the load.

We expect that index-based pricing for lanes with inconsistent demand will result in higher primary carrier acceptance and less reliance on backup or spot alternatives. Moreover, carriers would benefit by being able to better cover additional internal costs for serving unanticipated demand. This leads us to the following two hypotheses:

H3: The likelihood a primary carrier accepts a load increases with higher shipper tendering frequency on the awarded lane and further increases as the load's contracted price *increases relative to the current, lane-specific spot price.*

H4: The likelihood a primary carrier accepts a load increases with lower shipper tendered volume volatility on the awarded lane and further increases as the load's contracted price increases relative to the current, lane-specific spot price.

In the above hypotheses, we capture lane-level demand patterns that make carrier acceptance difficult. Next, we measure a load characteristic: surge volume. Recall, the contract agreement between a shipper and a primary carrier includes a (non-binding) awarded or expected volume However, due to poor forecasts, network changes, or other unanticipated end customer demand, shippers may tender more loads to a contracted carrier than the awarded volume in a given week.[2] This additional surge volume above the carrier's allocated capacity to that lane may be more difficult or costly to cover and may require an additional price incentive for the carrier to accept the surge loads. Thus, we identify each load that is tendered above the awarded weekly volume as surge loads and test the following hypothesis:

H5: The likelihood a primary carrier accepts a surge load increases as the load's contracted price increases relative to the current lane-specific spot price.

Finally, we aim to determine which carrier types are best suited for indexed pricing. Carriers fall into one of two categories based on the services provided: asset-based carriers that own the trucks and trailers used to move freight, and non-asset providers, often referred to as brokers or third party logistics (3PL) providers. This latter group of providers remove the shippers' burden of securing capacity and act as the middle man between shippers and asset-based carriers. Brokers access a vast pool of typically smaller, asset-based carriers, aggregate their capacity, and match it to shippers' demand. These brokers, or non-asset providers, typically buy and sell transportation based on their expectations of the general market; their profit margins are tied to how well they are able to manage the cycles. We test the following hypothesis:

H6: Non-asset primary carriers are less likely than asset-based primary carriers to stick to fixed-price contracts as spot market prices rise.

[2] A similar issue for primary carriers may arise if the shipper tenders much *less* than the awarded volume. While this is an important issue to address, it is out of the scope of the present research.

We test this hypothesis by calculating the rate of primary carrier load acceptance (PAR) as Spot Rate Differential changes. A higher rate of change (i.e., decrease in PAR as spot price increases relative to contract price) suggests lower contract price stickiness. A similar approach is taken by [94] to measure carriers' load price elasticity.

We further characterize asset carriers based on their fleet size. We expect that larger carriers may be better insulated from fluctuations in the market and may be more willing to forego opportunistic spot market loads than smaller carriers. We formulate our final hypothesis as follows:

H7: Larger asset carriers are more likely than smaller carriers to stick to fixed-price contracts as spot market prices rise.

By exploring each of these hypotheses individually, we can measure carriers' contract price stickiness for different freight segments of interest and identify which areas of a shipper's network each carrier type may be most amenable to index-based pricing. In this way, we address our central research question.

3.5 Carrier load acceptance model specification

In this section, we summarize our empirical data, describe our carrier acceptance model, and define the variables we use to build the model. Our partner company - a major US freight transportation management firm - provides us with transaction data over four years (2015-2019) of all of the TL loads for 68 shippers of various sizes and industry verticals and their 412 (primary and backup) carriers, both asset and non-asset. The data represent each load's tender sequence. This includes the date, time, and price at which each load is tendered to the primary carrier with its accept or reject decision, and, if needed, backup carriers' accept or reject decisions. The reported tender sequence continues down the routing guide waterfall until a carrier - primary or otherwise - accepts the load and the price the load is accepted at, including an indication if it ultimately is moved on the spot market.

The set contains 1.7 million long-haul loads (i.e., loads that move a distance greater than 250 miles)[3] all of which originate and terminate in the continental US. From the subset of

[3]We use this long-haul distinction because pricing structures for the alternative, short-haul moves, differ

tenders to primary carriers, we model the probability a load is either accepted or rejected (a binary outcome) by the primary carrier based on associated load, lane, shipper, and carrier characteristics described in the following subsections.

Logistic regression models are widely used in econometric literature to isolate the relationships between a binary dependent variable and independent input variables. Moreover, these models are the predominant modeling choices for authors studying producer and consumer price stickiness [97, 45, 38]. We adopt a logistic model choice as well and, to allow for non-linearity between independent and dependent variables [103], we discretize continuous input variables and include them as categorical variables.

Our dataset is comprised of multiple load accept/reject decisions by each carrier. As such, we must account for repeated measures of the same individual carrier. With repeated measures data, typical (logistic) regression modeling neglects to account for the correlations between the set of decisions made by the same individual. This within-subject correlation results in inefficient estimators. In other words, the calculated estimators have a greater spread around the true population values. Instead, General Estimating Equations (GEEs) model the average response of an individual in the population [91, 17]. The regression coefficient estimates in GEE models consider the covariance matrix between the outcomes in the sample associated with the same individual. GEE estimators reduce to those obtained through OLS if the dependent variable is normally distributed and no within-individual response correlations exists [69, 63].

We use a GEE model with a logistic link function where the coefficient estimates represent the marginal increase or decrease in the (log) odds a carrier accepts a load:

$$logit(y_{c,k}) = log\Big(\frac{y_{c,k}}{1 - y_{c,k}}\Big) = x_{c,k}^T\boldsymbol{\beta} \tag{3.2}$$

where each individual primary carrier, c, makes a binary accept/reject decision for each load it is tendered, k (the positive outcome, $y_{c,k} = 1$, is a load acceptance), and the matrix of explanatory variables, $x_{c,k}$, are the lane, freight, shipper, and carrier variables described in the following subsections.

from those we consider and discuss in this research.

3.5.1 Load Spot Rate Differential

We define the Spot Rate Differential (SRD) in Section 3.2 and Eq. 3.1 as the percent difference between the current lane-specific spot market price and the contracted price of a load. The SRD calculation requires knowledge of each lane's spot price at the time a load is offered. While we do not have such data consistently across all lanes and time, given the breadth of our dataset, we can calculate benchmark spot market prices.

Our dataset approximately represents the general freight market trends as it is comprised of many shippers' load tenders across the continental US. We corroborate this claim by comparing two statistics from our dataset to external industry data. First, we measure the correlation between the time series of average primary carrier acceptance rate (PAR) in our dataset, a real-time primary carrier behavior, and the Morgan Stanley Freight Index, which represents overall practitioner sentiment of the market's supply and demand. The correlation between the two time series is 85.2% (see [130] for a similar justification process). Second, the correlation between our national average contract linehaul price and that of the Cass Truckload Linehaul Index is 91.8%. We conclude that our dataset is sufficiently representative of the overall freight market.

Next, we reconstruct spot prices for each origin region to destination region combination. Regions of the US differ in attractiveness to carriers based on the business opportunities that are expected in those areas. For example, regions with high outbound demand are typically more attractive for carriers to accept loads going into because they are more likely to easily find a follow-on load. To account for the dynamic nature of spot prices, we include the year and month indicator in the spot price model.

[53] shows that point-to-point transportation costs (e.g., linehual costs) result from a combination of fixed and variable costs. Based on this, we model spot prices using multiple linear regression with heteroskedastic robust standard errors. We regress the point-to-point linehaul price of loads that are accepted on the spot market in our dataset on origin and destination region binary variables and a month and a year binary variable. These represent fixed costs. We include a continuous distance variable, which represents the carriers' variable costs [18, 127]. We further detail the price benchmarking methodology in Chapter 2.

Our base case lane (i.e., the binary variables omitted to avoid multicollinearity) originates in the Lower Atlantic region of the US and terminates in the South Central region in January of 2016. These ommitted variable choices correspond to the most volume (i.e., greatest number of observations) within each categorical variable. The lane-specific spot price for a given time, $Spot_{i,j,t}$ is defined as:

$$\begin{aligned} Spot_{i,j,t} = \hat{\beta}_{base} + \hat{\beta}_{dist}\overline{X}_{dist}^{(i,j)} + \sum_{i \in I, i \neq i_{base}} \hat{\beta}_i X_i + \sum_{j \in J, j \neq j_{base}} \hat{\beta}_j X_j \\ + \sum_{m \in M, m \neq m_{base}} \hat{\beta}_m X_m + \sum_{y \in Y, y \neq y_{base}} \hat{\beta}_y X_y \end{aligned} \tag{3.3}$$

The intercept term, $\hat{\beta}_{base}$, is the fixed cost of the base case, and $\hat{\beta}_{dist}$, is a distance (i.e. variable cost) coefficient associated with the average distance of loads between i and j, $\overline{X}_{dist}^{(i,j)}$. The fixed cost of the origin and destination regions that are different from the base case are represented by the I-1 origin coefficients, $\hat{\beta}_i$ (where I is the set of origin regions and X_i the corresponding I binary variables indicating in which region the load originates), and J-1 destination coefficients, $\hat{\beta}_j$ (where J is the set of destination regions and X_j the corresponding destination binary variables). The 15 mutually exclusive and collectively exhaustive regions are key market areas defined by our industry partner representing geographic clusters of transportation demand patterns. The origin and destination coefficients of our linear regression model, $\hat{\beta}_i$ and $\hat{\beta}_j$, can be interpreted as spot price premiums associated with an origin or destination different from the base case lane.

Finally, $\hat{\beta}_m$ and $\hat{\beta}_y$ measure the dynamic, time-based changes in spot prices for each month and year, respectively, and capture both seasonal and underlying market structural trends.[4] As a notational simplification, for the remainder of this paper, we combine m and y for a time-dependent variable by denoting it with a subscript t.

With the results of this model, we calculate a single average spot price for every origin-destination-month-year combination in the dataset. As discussed in Section 3.1, this spot

[4] We choose to include month and year variables (as opposed to a single variable that treats each month in the time frame separately - that is, $m \in \{1, 60\}$ and no X_y variable) to ensure enough observations are in each time-based outcome variable for robust coefficient estimates.

price represents the average of a distribution of underlying spot load prices. We use this average spot market price to calculate the Spot Rate Differential (SRD) for each load, k, using Equation 3.1 as the key metric to test carriers' contract price stickiness.

3.5.2 Lane distance

Carriers' preferences regarding lane distance may vary depending on individual operations. Each carrier tries to maximize asset utilization - in other words, miles driven carrying a paid load. We measure a lane's distance by the number of days it takes to drive that distance, assuming a driver can drive 400 miles per day based on federally mandated hours of service regulations and actual driving patterns [81]. For a more detailed discussion of carrier economics, see [21], [101], and [33].

Our lane distance measure is broken into five binary "travel-days" variables based on whether the lane is expected to take up to 1 day to drive, 1-2 days, 2-3 days, 3-4 days, or more than 4 days. Rather than using a continuous variable for distance (or travel-days), we decompose the variable into these categories to avoid a linear assumption on the relationship between carrier acceptance and distance.

3.5.3 Lane demand cadence

In the supplier-buyer relationship, frequency of interactions points to more positive relationship; infrequent demand patterns are problematic for suppliers [120]. This aspect of the relationship is particularly true in the TL industry. Inconsistent or infrequent tendered volume from one shipper (the customer) makes it difficult for the transportation suppliers (carriers) to plan where and when they need to position capacity to balance their networks and serve each of their other customers.

One metric carriers use to measure shipper performance is tender cadence, or the frequency at which loads are tendered [81, 46]. [130] incorporate this frequency of interactions as a contributor to load acceptance by measuring the number of days since the previous load was offered to a carrier by a shipper. The authors find that carriers are less willing to accept loads that are tendered at unpredictable frequencies. Moreover, in Chapter 2, we demonstrate that primary carrier acceptance increases with the percentage of weeks in which loads are tendered to that carrier on that lane [2].

We use this latter measure to characterize a lane's frequency. For each load, we calculate the percent of the preceding four weeks in which that shipper tendered at least one load to that carrier on that lane. The resulting cadence metric is a discrete measure, taking on values of 0%, 25%, 50%, 75%, or 100%. Our input variables to the carrier acceptance decision models are binary variables indicating which of these discrete values the cadence measure takes.

3.5.4 Lane demand volatility

In addition to the frequency of interactions, consistency of demand is an important factor for carriers to anticipate capacity needs [81, 46]. Both practitioners and the literature note the importance of reducing tendered volume variability to improve carrier freight acceptance, reduce the cost of loads, and help carriers better utilize their capacity [70, 130, 5, 2].

We incorporate a measure of lane-level tendering volatility between a shipper and primary carrier acceptance. Each load is assigned the corresponding lane-level tendering volatility, which we calculate as the shifted 4-week rolling average week-over-week change (measured as the square difference) in tendered volume from the shipper to the carrier on that lane:

$$Vol_{i,j,t} = \sqrt{\frac{\sum_{\tau=t-1}^{t-4}(d_{i,j,\tau} - d_{i,j,\tau-1})^2}{4}} \tag{3.4}$$

where load k is tendered on lane (i, j) in time period t and $d_{i,j,\tau}$ is the number of loads tendered on the lane in week τ. We consider only weeks in which loads materialize and are tendered to the carrier - in other words, weeks in which $d_{i,j,\tau} > 0$. This is because we want a measure of the volatility of the *materialized* volume. We already capture how frequently there are no-volume weeks with the Cadence measure.

3.5.5 Surge volume

Next, we consider surge volume, a load characteristic defined as tendered loads that are above the awarded, or expected, weekly volume. Carriers report that they can typically manage to make capacity available when the number of loads tendered from a shipper in a week is within about 10% of the lane award volume. However, as volume reaches and surpasses 20% of the award volume, carriers often are unable to serve the excess demand. Moreover, shippers commonly call for contracted carriers to flex up with increased demand, often up to 20% above the awarded volume, as a stipulation of the service level expectations in the contract [135]. Such service level expectations highlight the non-binding nature of the contract: while agreed upon and defined in the contract, they are not legally or contractually enforceable by the shipper. The main incentive for the carrier to uphold them is the promise of continued business from the shipper (i.e., the shadow of the future).

While the expected volume is part of the information communicated to carriers during the RFP, many shippers do not keep careful record of the awarded volume to each carrier on each lane after the bid is complete. As such, our dataset does not include the primary carriers' awarded volume for each lane. Instead, we use a proxy for this awarded volume: the preceding shifted 4-week rolling average of the tendered volume to the primary carrier on the lane. [130] use a similar proxy to rank loads. The authors measure the average daily volume on a given lane over the 30 days preceding the load of interest and denote how the load's rank within the day measures relative to that 4-week rolling average daily volume. We similarly rank each load within the week and categorize it based on how its rank compares to the awarded weekly volume proxy.

Each load is assigned to its corresponding surge category. If the load's rank within the week is less than or equal to the awarded volume proxy, it is given a Surge category of *Within Mean*. If the load's rank is more than the average but less than or equal to 10% above the mean, it is in the *Up to 10% Surge* category, if it is more than 10% but less than or equal to 20% of the average volume, it is in the *Up to 20% Surge* category, and if the load's rank is more than 20% above the awarded volume proxy, it is in the *Over 20% Surge* category. We expect the higher the surge volume category a load is in, the less likely it is to be accepted by a primary carrier.

3.5.6 Carrier service type

Asset and non-asset carriers offer different services which drive their relationships with shippers and how they interact with the overall market. On one hand, asset carriers own the trucks (tractors) they operate. They have fixed available capacity to manage across their networks and serve customers.

On the other hand, non-asset carriers - otherwise known as brokers or freight forwarders - do not own the tractors or trailers that move their customers' loads. Instead, non-asset providers match shippers with asset-based carriers. The advantage to shippers of working with a non-asset carrier is that they are able to aggregate smaller (asset) carriers' capacity to serve the shipper's needs. Often, shippers aims to maintain a manageable carrier base size (i.e., number of contracted carriers) across their networks. Rather than contracting with many, small carriers, they prefer to allow brokerages to manage these carriers. The benefit for smaller asset carriers of working with a broker or 3PL is that often they might not otherwise have access to some shippers' business. For larger asset carriers, the broker may provide opportunities to fill backhauls or moves to re-position a truck that does not have contracted volume and would otherwise be an empty, unpaid trip.

A shipper may set up a contract with a non-asset provider on a lane to lower the risk that loads go to the less predictable, volatile spot market. Brokers set prices with shippers for a contract term length and then typically pay (close to) spot market prices for their asset carriers over the course of the contract. They aim to hedge the market and, over time, still make a reasonably sustainable margin. When spot market prices are high, they may be paying carriers more than they are receiving from their contracted shippers. However when spot market prices are lower in a soft market, they receive higher payments from their fixed-price contracts than the price at which they are buying capacity on the market.

Because of the difference between how asset and non-asset providers utilize markets, we predict non-asset carriers are more likely than asset carriers to be pulled from their contracted loads when spot prices are high and more attractive than contract prices. Thus they would be more responsive to index-based pricing.

Using an auction theory lens, this distinction in behaviors between how asset and non-asset carriers approach bid pricing is addressed in [128]. We aim to expand on this by modeling the behaviors of asset and non-asset carriers separately.[5] In this way, we extract each type of carriers' contract price stickiness for different freight and lane segments independently.

3.5.7 Asset carrier fleet size

Within the asset provider segment, carriers' cost structures vary by fleet size. For example, large carriers may be more able to absorb variations in market prices than would a smaller, owner-operator that owns a single truck and trailer. The distribution of asset carrier fleet size - both across the industry in the US and in our dataset - is highly skewed; about 60% of total for-hire carriers in the US are independent owner-operators, and 96% of fleets have fewer than 20 trucks. To account for this, we include the log of the carrier's fleet size (tractor count) as our measure of asset carrier size.

Figure 3-3 depicts the skewed distribution of carriers' number of tractors in the fleet and the normalized distribution of the log of carrier fleet size from our dataset.

3.5.8 Other fixed effects

While we do not make formal hypotheses regarding the following variables, we do control for their fixed effects in our model and report their effects on carrier acceptance decisions. These include the shipper's size, measured by the log of its total monthly tendered volume across all lanes, and the shipper's industry vertical.

Indicators for the lane's origin and destination regions are also included to account for their relative attractiveness. As discussed in Section 3.5.1 and Chapter 2, carrier accep-

[5] Some carriers also offer both services. For example, large asset carriers J.B. Hunt, Schneider, and Knight-Swift all run a brokerage division of their business. In our dataset, we can distinguish between whether a company's asset or non-asset arm was awarded or tendered loads because they fall under different SCAC codes. SCACs, or Standard Carrier Alpha Codes, are unique 2- or 4-letter codes assigned by The National Motor Freight Traffic Association, Inc., (NMFTA) to identify transportation companies and are used throughout the freight industry for consistent carrier identification.

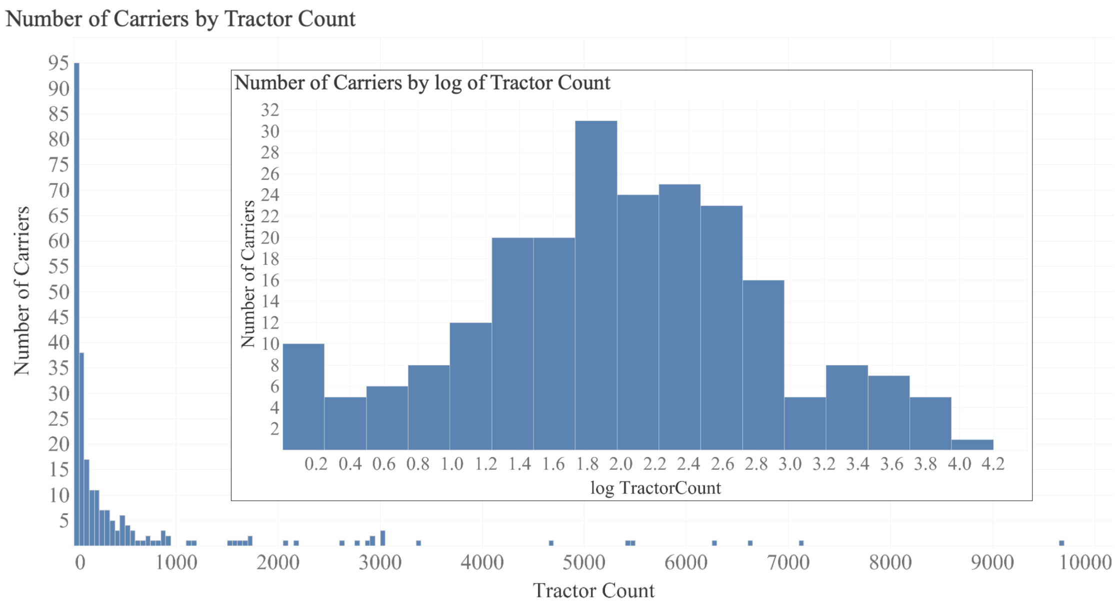


Figure 3-3: Distribution of Asset Carrier Fleet Size

tance is expected to vary across different inbound or outbound regions due to the business opportunities present at the origin and destination [2]. We include in our carrier acceptance model a binary variable for each of the 15 regions of the US defined by general market demand patterns of our transportation management partner company, at both an origin and destination.

The results of our model indicate which load, lane, and carrier characteristics are statistically significant indicators for predicting primary carrier acceptance rate (PAR) at a load transaction level. Further, we address our carrier price stickiness research question by quantifying the changes in PAR as the contracted load price relative to the spot market price (i.e. SRD) changes for the resulting statistically significant freight and carrier segments.

3.5.9 Market Condition

Building off the research outlined in Section 3.3, we expect carriers' contract price stickiness to differ depending on the general market condition. Our dataset spans two distinct market conditions: a soft market observed from the first week of February 2016 to the first week of July 2017 and again after the second week of January 2019 until the end of the time frame covered by the dataset; and a tight market observed before the first week of February 2016 and from the first week of July 2017 to the second week of January 2019. We justify these market periods in Chapter 2 by identifying the weeks in which a statistically significant change is observed in the underlying structure of the TL freight market [2]. We develop four distinct acceptance models: (1) asset carriers in soft markets, (2) non-asset carriers in soft markets, (3) asset carriers in tight markets, and (4) non-asset carriers in tight markets.

3.6 Results

In this section, we summarize our results by discussing the full GEE logistic regression models and use them to predict the (log) likelihood a load is accepted for each hypothesis. The regression results are tabulated in the electronic companion. The separate reported tables comprise a single model with the input variables described in Sections 3.5.1-3.5.8.

As validation of each of our four models, we report the Brier Score, which is the appropriate scoring metric when probability outcomes are the desired result [153, 106, 167, 29]. The Brier Score, BS, is the mean square error between the predicted probability an observation is in the the positive class, p_n (here, an accepted load, $y_{c,k} = 1$), and the actual outcome, $o_n \in \{0, 1\}$:

$$BS = \frac{1}{N} \sum_{n=1}^{N} (p_n - o_n)^2 \tag{3.5}$$

where N is the total number of observations in the dataset. The Brier Score takes on values between 0 and 1. Better models have lower Brier Scores.

To build our models, we segment each of the datasets for the four models into a training set on which we fit the model and a test set to measure model performance with a 70%:30% split. In each of the four pre-split datasets, the relative frequency of accepted loads is much higher than that of rejected loads. To develop unbiased models that do not naively favor accepted load predictions, we use a stratified sampling technique to define each of the training and test sets such that the ratio of the number of accepted loads to number of rejected loads in each split set is the same as that of the original pre-split dataset. In this way, the validated models can better predict carriers' load acceptance or rejection probabilities on new data.

Model	Brier Score
Asset carriers Soft market	0.049
Asset carriers Tight market	0.032
Non-Asset carriers Soft market	0.075
Non-Asset carriers Tight market	0.064

Table 3.1: Test set Brier Scores

The Brier Scores for each of our four models on their respective test datasets are reported in Table 3.1. Recall that the closer a score is to 0, the better. A "good" Brier Score largely depends on the dataset itself. However, we draw from the literature for reasonable comparison. [167] report test set Brier Scores for the best calibrated probability prediction models of five datasets ranging from 0.012 to 0.204. Similarly, [153] report calibrated model Brier Scores for 34 datasets between 0.042 to 0.319. All four of our models' Brier Scores sit below the average Brier Score value of the best fitting models in these previous studies. Thus, we conclude that our models perform well in predicting loads acceptance probability.

3.6.1 Carrier price stickiness by service type and market condition

The results of the GEE logistic regression models are reported in Table B.1. Hypothesis H1, which addresses carriers' overall contract price stickiness is supported: in general, primary carriers are more likely to accept a load with a contract price that is higher relative to the current, lane-specific spot market price. The statistically significant coefficients in Table B.1 show that as SRD increases and the spot price rises above contract prices, the probability a carrier accepts the load decreases for both carrier types in both market conditions. This suggests that in general, carriers can be incentivized away from their contracted loads as the external spot market increases relative to those contract prices.

Figure 3-4 illustrates this behavior for each carrier type and market condition. To construct the figure, we plot the probability a load is accepted at each of the statistically significant SRD values for a base case lane and fit a linear model to these plotted results./footnoteWe show only the segments of these linear models over the SRD ranges that apply to the associated market condition. In other words, soft market models apply when SRD negative (spot market prices are below contract prices) and tight market models apply when SRD is positive. The slope term of each linear model represents the carrier's contract price stickiness, or willingness to stick to contracted loads as spot market prices change relative to contract prices in the corresponding market condition.

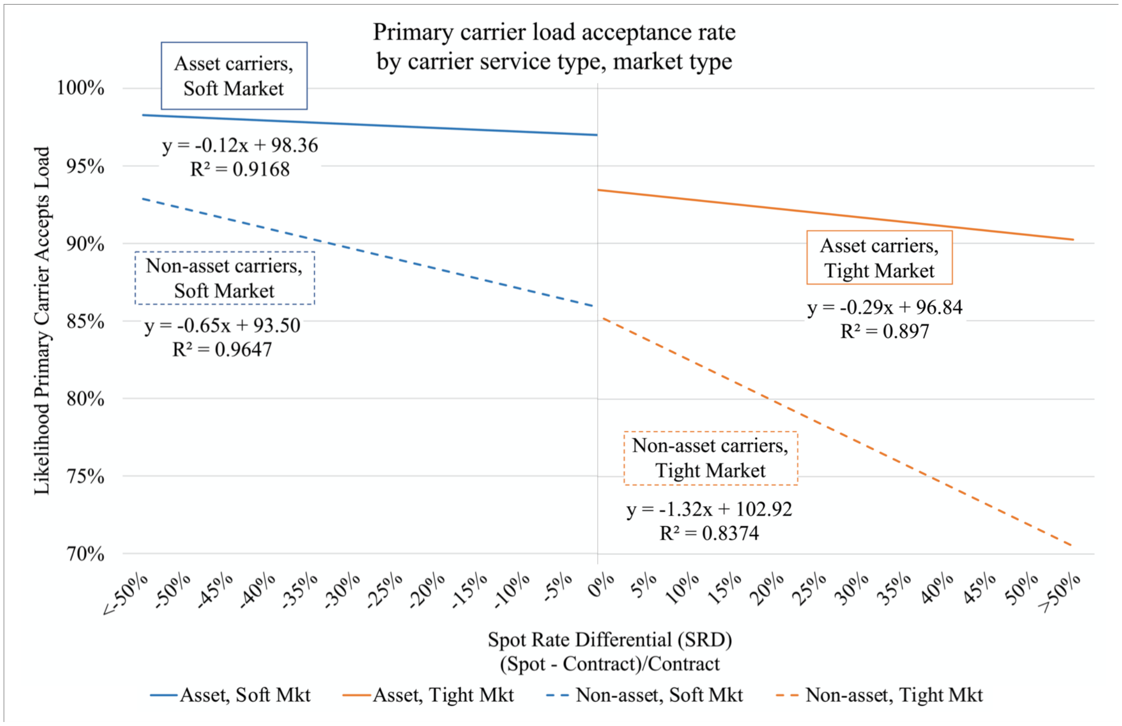
Both carrier types are about twice as likely to be pulled from contract load prices in a tight market as they are in a soft market: asset carriers have a slope of -0.29 in tight markets as compared to -0.12 in a soft market and non-asset carriers have a slope of -0.65 in soft markets and -1.32 in tight markets. Moreover, H7 is supported: non-asset carriers are about five times more likely to be pulled from contracted loads than their asset-based counterparts in each market condition.

For shippers considering indexed contract pricing strategies, these results suggest that all else equal, non-asset carriers are more likely to respond to market-based pricing than asset carriers and that both carrier types may be more responsive to such pricing strategies in tight markets than soft.

Next, we discuss the results for specific freight and lane segments highlighted in our hypotheses. We demonstrate carriers' contract price stickiness by holding spot price constant at $1000 for both soft and tight markets and plotting the contract price needed to maintain a 90% likelihood the primary carrier accepts the contracted load over the range of values each freight or lane segment of interest takes. We choose 90% PAR because most shippers expect at least this level of service from their contracted carriers. Many may expect higher acceptance rates, however this threshold represents basic service level expectations in the shipper-carrier relationship. Our method of presenting contract price stickiness follows those discussed in Section 3.3 on general supplier price stickiness and responses to factors exogenous to the supply contract.[6]

[6]The figures identify segments where primary carriers require higher price incentive to stick to their contracts as the spot price approaches or surpasses the contract price. Contract prices of categorical variable values for which the GEE regression model coefficients are not statistically significant (see B.1) are presented as the contract price associated with the baseline value.

Figure 3-4: Carrier Contract Price Stickiness Models



3.6.2 Contract price stickiness: lane demand consistency

Results for both lane tendering frequency and volatility are summarized in B.4. They show that hypotheses H3 and H4 are supported particularly for asset carriers in both market conditions. Carriers prefer lanes with more frequently and consistently tendered volume; that is, carriers are more willing to stick to their contract prices on lanes with high frequency load tendering and lower week-to-week volume volatility.

Importantly, the shipper often cannot control its load tendering consistency; it is subject to external factors such as its own customers' demand patterns, inbound suppliers' schedules, and congestion on roadways or at ports. While the shipper does not control when its demand for trucks materializes, it does control what carriers are offered the loads when they do appear. Moreover, the shipper has historical knowledge of its lane demand patterns. Thus, the shipper's decision here is what bid prices to accept and how to tender loads that do appear for lanes with historically irregular cadence or inconsistent volumes.

Tendering cadence

The results weakly support Hypothesis H3: lanes on which loads are tendered less frequently see lower primary carrier acceptance. Moreover, both asset and non-asset carriers in both market conditions are less willing to stick to their contract prices for these low cadence lanes. (See B.4.)

Figure 3-5 shows that as lane tendering cadence decreases, asset carriers' contract price needed to maintain 90% PAR increases about 7% in soft markets (to \$682 per load from \$617) and 2.5% in tight markets (to \$857 from \$830).

Similarly, non-asset carriers require a 6% contract price increase in soft markets to maintain 90% PAR on low cadence lanes. While they do not appear to make load acceptance decisions based on lane tendering cadence in tight markets, these non-asset carriers set the highest contract prices for all cadence levels in tight markets. This may be because shippers often use non-asset brokerage services specifically for infrequently tendered lanes. These providers may be familiar with the undesirable, infrequently tendered lanes and knowingly set higher contract prices during the RFP.

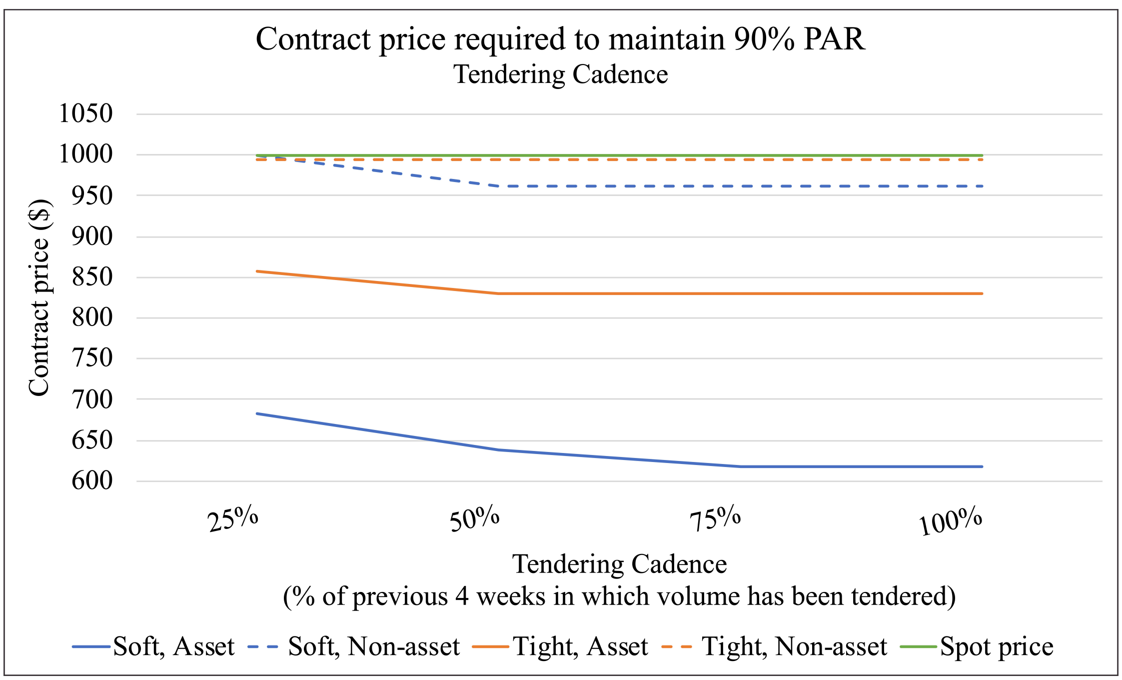


Figure 3-5: Contract Price Stickiness, Tender Cadence

The results suggest that both carrier types are willing to stick with contract prices on moderate and high cadence lanes under both market conditions. However, lanes with very infrequent load tenders require higher contract prices relative to going spot market prices. Shippers may see better primary carrier acceptance on these low frequency lanes with market-based pricing strategies. Important to note, transportation practitioners and the literature both highlight that there must be enough business between the buyer and supplier (on a lane) to offset the additional effort involved in introducing either non-traditional, or more explicit contracts [129, 137].

Tendered volume volatility

Next, we consider the tendered volume volatility on a lane from a shipper to its primary carrier. Consistent with Hypothesis H4, higher tendered volatility leads to lower load acceptance probabilities, particularly for asset carriers. The GEE logistic regression model results are presented in Table B.4 and Figure 3-6 below demonstrates carrier contract price stickiness. In soft markets, asset carriers stick to their contract prices for lanes with week-to-week tendered volume volatility up to 50%.

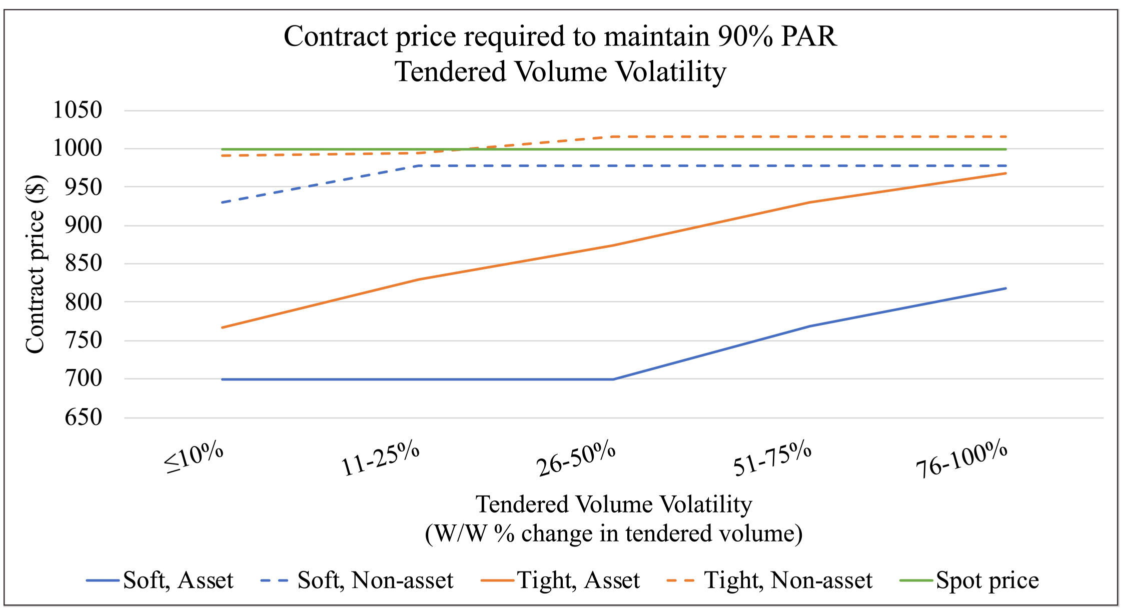


Figure 3-6: Contract Price Stickiness, Lane Volatility

However, as tendering volatility increase, they are more likely to be pulled from the contract priced loads unless contract prices are much closer to the current spot market price. Specifically, contract prices needed to maintain 90% acceptance are 17% higher on these high volatility lanes than low volatility lanes in soft markets.

In tight markets, as compared to soft markets, even higher contract prices are needed for asset carriers to maintain high acceptance rates. Moreover, these price escalations begin at lower volatility lanes (i.e., 11-25%). Contract prices needed on high volatility lanes are 26% higher than those on low volatility lanes in tight markets. This suggests that in tight markets, asset carriers are more easily pulled from their contracts on lanes with lower tendering volatility than in soft markets.

Non-asset carriers are more willing to stick to their contract prices at high lane volatility in both market conditions as compared to their asset-based counterparts. Contract prices needed for non-asset providers to maintain 90% load acceptance rates remain steady for lanes with more than 10% week-over-week change in tendered volume, with an increase of only 5% in soft markets for moderate and high volatility lanes.

Shippers should pay close attention to moderate and high volatility lanes in their networks. These lanes are particularly difficult for asset carriers to accommodate due to fixed capacities and a network of many customers' demand they continuously balance while non-asset providers do not. For the lanes on which a shipper's demand volatility is difficult to control or smooth out (by splitting the volume and tendering to multiple carriers, for example), it may be beneficial to introduce a market-based pricing strategy that adjusts contract load prices as spot market prices change to better incentivize primary asset carriers to accept loads.

3.6.3 Contract price stickiness: surge volume

Table B.5 summarizes the GEE logistic regression models' results demonstrating carriers' likelihood of accepting loads considered surge volume. Hypothesis H5 is strongly supported for asset carriers: asset primary carriers are less likely to accept loads that are above the lane awarded volume - particularly those over 20% above awarded volume - than loads within their expected weekly tendered volume. Put another way, asset carriers are less willing to stick to their contract prices for excessive surge volume, especially in tight markets.

Figure 3-7 shows that the contract price required to maintain 90% acceptance for asset carriers in soft markets increases by 9% for loads that are over 20% above the awarded volume. These asset carrier contract prices further increase in tight markets. For surge volume over 20% above the award in tight markets, shippers pay $940 per load - almost at the $1000 spot market price level - which equates to a 13% increase in contract price from the awarded volume loads.

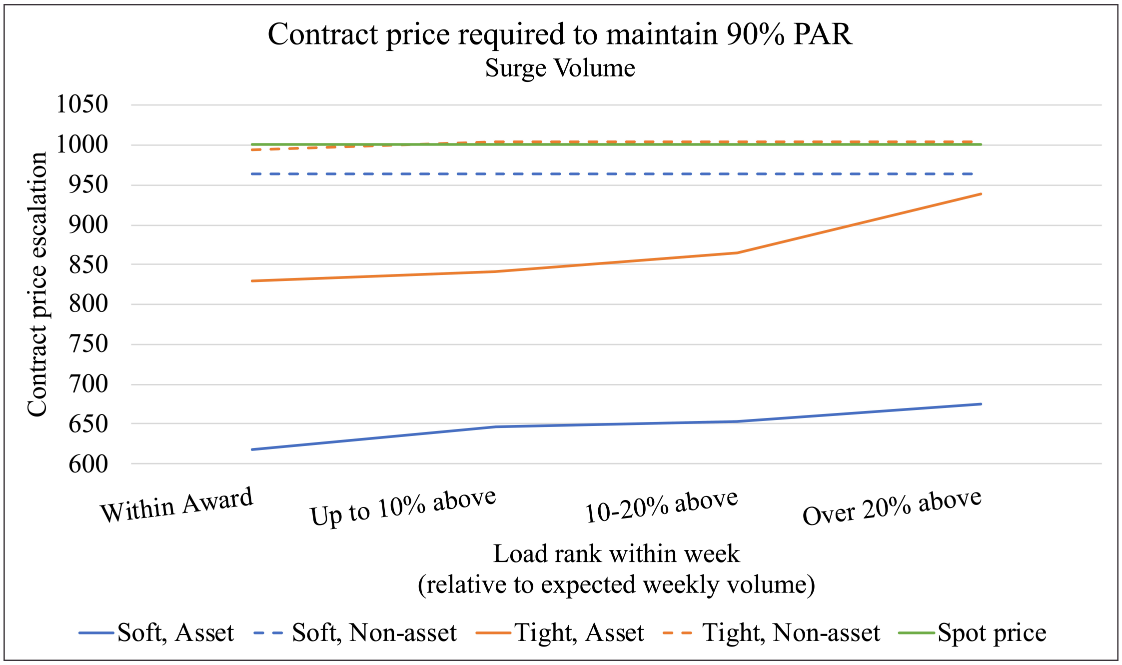


Figure 3-7: Contract Price Stickiness, Surge Volume

The contract price required to maintain high acceptance rates for loads up to 10% and 10-20% above awarded volumes with asset carriers stays relatively steady. This moderate surge volume requires contract price increases of 4-6% in both soft and tight markets. Asset carriers are willing to stick to their contract prices for these mid-level surge loads. This may be because shippers often communicate an expectation of primary carriers to accept anywhere from 10-20% above the awarded volume without reducing acceptance or other carrier performance metrics in their (non-binding) contract service level agreements. Although asset carriers may have more difficulty providing this additional capacity than non-asset carriers, they may have factored some additional capacity into their strategic capacity allocation decisions during the original RFP. Surge loads over this 20% threshold require higher contract prices, especially in tight markets, and asset carriers are more willing to be pulled from contracts for better priced alternatives on the spot market.

Non-asset carriers, on the other hand, do not appear to be incentivized away from contracts with higher priced alternatives on surge volume, even for loads that are more than 20% above the awarded volume. This difference between asset and non-asset carriers is expected. Non-asset carriers are not limited in capacity in the way asset carriers are. While non-asset carriers must also balance supply and demand of capacity, they are better equipped to serve excess demand because they can access a large pool of asset carriers and aggregate capacity accordingly. Moreover, non-asset carriers already require higher contract prices to maintain 90% acceptance for all volume - in fact, very close to the spot market price - regardless of surge classification. Thus, they may already be adequately incentivized to offer that capacity.

Shippers may see the most benefit from implementing indexed pricing applied to specific surge volume loads with asset carriers. This type of volume-based pricing strategy has been seen in practice. It is referred to as tier pricing, where levels of surge volume are set at a higher fixed price - determined during the strategic RFP - than that of the base awarded volume.

3.6.4 Contract price stickiness: lane distance

Table B.3 summarizes the lane distance portion of the GEE logistic regression model results. It shows weak support of Hypothesis H2: non-asset carriers prefer longer lanes. In soft markets, loads on shorter lanes - specifically those with only a one-day travel time - are less likely to be accepted than all other lane distances. In tight markets, loads on longer lanes (three- and four-day lanes) are more likely to be accepted than loads on shorter lanes. However, there is no statistically significant difference in probability of load acceptance for asset carriers on different lane distances, suggesting that asset carriers do not consider lane distance in their load acceptance decision.

3.6.5 Asset carrier contract price stickiness by fleet size

The preceding sections consider carriers based on their service type: asset and non-asset. However, asset carriers range widely in their fleet sizes (Figure 3-3), which directly impact their internal cost structures, tolerance to price fluctuations, access to network model optimization software to plan strategically, and as a result, willingness and ability to accept loads. Table B.6 summarizes the GEE logistic regression models' results for an asset carriers' fleet size. Hypothesis H8 is not supported: all else equal, we do no find statistically significant evidence that the size of the asset carrier (measured here as log of Tractor Count) indicates its primary acceptance rate in soft or tight markets.

3.7 Discussion and Implications for Index-based Pricing

In this study we explore supplier price stickiness in response to customer demand patterns and exogenous market conditions. We take the case of firms that outsource their TL transportation service needs and aim to help them determine when, where, and with which transportation suppliers they should consider alternative pricing strategies from the standard, fixed-price contracts. In particular, we consider dynamic market-based pricing by measuring primary carriers' contract price stickiness, or the change in contract price relative to the going spot price needed so carriers maintain a reasonable load acceptance service level.

1. Shippers should not delay or skip RFPs or mini-bids as markets tighten. Both carrier types (asset and non-asset) are twice as likely to stick to their contract priced loads in a soft market than a tight market. Shippers can use this to inform the timing of their strategic procurement event as market conditions may be changing. For example, say a shipper runs its annual strategic RFP in January. In the preparation months leading up to the event, the shipper may observe that spot market prices or other leading indicators of general market conditions are rising. Some shippers in this situation may consider delaying the RFP and keeping the current contracted prices that had been set during a softer, lower-priced market rather than opening itself up for carriers to lock in higher prices. Alternatively, the shipper may choose to run the RFP anyway, perhaps allowing for slightly increased contract prices, and assume it can expect good primary carrier acceptance rates for the next year even as markets further tighten. Finally, the shipper may take a middle-ground approach and proactively increase rates for certain core carriers on specific important lanes to keep them out of an RFP.

However, the results of our study suggest that no matter the decision, if spot prices continue to rise as the market tightens further, the contract prices will become less competitive (i.e., their spot rate differential will increase) and primary carriers with low contract price stickiness will begin opportunistically rejecting loads at higher rates, defecting to either fresher contract prices with other shippers or to the spot market. Therefore, when markets are tightening, shippers need to continuously ensure their contract prices stay competitive with the going spot market prices, regardless of their RFP timing. One approach shippers can take as market conditions become more constrained is to execute smaller, focused, more frequent "mini-bids" with core primary carriers on the most important, susceptible lanes. These mini-bids result in shorter term (e.g., 30-, 60-, or 90-day) contracts. Another approach is to implement a market-based pricing strategy, as we propose in this study, which effectively ensures contract prices are up-to-date.

2. Non-asset carriers are best suited for market-based contracts. All else equal, asset primary carriers are five times more likely to stick to their contract priced loads than non-asset primary carriers, or brokers, in both soft and tight markets. In fact, across most network and freight segments, shippers must pay these non-asset primary carriers just about spot market prices to maintain high load acceptance rates. This suggests that brokers may respond to index-based contracts better than asset-based providers.

3. Market-based pricing shows promise for volatile demand, low frequency lanes, and surge volume. Asset primary carriers are less likely to stick to their contract priced loads on lanes with infrequently tendered loads and high week-to-week volatility of tendered volume. Shippers often have little control over when loads actually materialize and carriers' capacity is needed. However, they do know historical demand patterns on their lanes. They can control which carriers they tender the loads that do materialize and the pricing strategies they are willing to implement. Our results suggest a market-based pricing strategy for these low cadence or high volatility lanes would better incentivize asset carriers to stick to their contracts.

In addition to these lanes types, shippers can expect asset carriers to respond to market-based pricing for surge volume - in other words, loads that are in excess of the expected weekly volume communicated during the RFP. This finding relates to the two stages of TL transportation. During the first stage, the strategic RFP, the shipper communicates the expected weekly volume on each lane that the carriers bid on. Once the carrier wins the lane, it plans to allocate capacity for that expected volume, perhaps with an additional 10-20% buffer. It also uses this knowledge of expected demand to balance the demand for trucks across its existing network and to inform its bid strategy in other shippers' RFPs. However, during the second stage, the operational stage when loads materialize and are tendered to that carrier, the actual demand may well exceed the capacity for which the carrier has planned. Asset carriers in particular are constrained in their total capacity available at any one time. To maintain high acceptance rates of loads that exceed 20% over the awarded volume on a lane, carriers require a price incentive. This is an important segment of shippers' freight: consistently across our time horizon, 10-15% of total tendered volume is in this surge category.

In fact, a tier pricing strategy where surge volume is priced higher than the baseline awarded volume is used by some shippers and their larger or core carriers, precisely for the reasons described. However, these tier-based rates are still fixed at the time the contract is established. A dynamic market-based price, on the other hand, would ensure the tiered price incentive remains competitive with the current market.

4. Indexed pricing may result in lower load acceptance if applied in soft markets. Important to note that we discuss market-based pricing with the implication that contracted load prices increase as market prices increase in a tight market, but also decrease as prices decrease in soft markets - in other words, "symmetric" indexing. Shippers may want to consider using dynamic market pricing as an incentive for primary carriers whereby indexed contract prices only increase as the market tightens, but settles to the competitive, fixed contract price when the market prices are decreasing. Otherwise, for segments that require a contract price *premium* (i.e., higher contract price than spot market price) for high primary carrier acceptance, the shipper would expect to see decreased primary carrier acceptance when the indexed price symmetrically decreases.

Our study suggests that there is still a place for the widely used fixed-price contracts. On lanes where tendered volume is consistent and frequent, both shippers and carriers prefer a set price to plan and budget around. While these prices may also become outdated as spot market prices increase, carriers are more willing to stick to their contracts for attractive freight.[7] Moreover, competitive fixed price contracts may still be best suited for soft market conditions.

3.8 Limitations and Future Research

Our empirical modeling results offer both academic and practical contributions. First, we add to the econometric literature on supplier price stickiness in response to demand and exogenous market dynamics. We do so by quantifying TL transportation suppliers' price stickiness as market prices change and for different customer demand patterns, freight segments, characteristics of shippers' networks, and supplier service types. Our work explores real-time decision that previous literature overlooks; transportation suppliers have the spot market as a dynamic alternative option to their contracted business for every transaction.

In addition, we add to the literature specific to shipper and carrier relationships, particularly with a focus on how changing market conditions impact behaviors. We do so by utilizing a uniquely extensive and detailed dataset. Previous empirical studies in the space have been limited in that their data represent only a single shipper's business, and often contain limited or no tendering or carrier acceptance service level information.

This study is not without its limitations. Precisely due to the observational nature of our dataset, we cannot control for outside influences on carriers' acceptance decisions. We attempt to account for these factors by segmenting the contract price stickiness analysis by lane, load, and carrier types and other fixed effects that have been previously identified as contributing factors. In doing so, we build models with good performance (i.e., very low

[7] While low tight market period primary carrier acceptance rates are undesirable, the majority of freight still moves under fixed contract rates even in tight markets. However, we find evidence that a market-based approach would improve primary carrier acceptance rates.

Brier Scores). However, there may still be alternative explanations for carriers' acceptance decisions.

Notwithstanding these limitations, our study further adds to the existing shipper-carrier relationship literature by isolating the impact of market prices on the way carriers make freight acceptance decision for their contracted shippers. As discussed in the motivation of this research, there has been growing interest in index-based pricing in the TL industry. This research serves as a starting point. We demonstrate where shippers can expect to see carrier acceptance behaviors most influenced by dynamic, market-based pricing methods. An interesting future stream of literature could develop strategies for shippers and carriers to design and implement these freight contracts.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 4

Expanding the freight contract portfolio: Index-based freight contract design under uncertainty

Abstract

Long-term, fixed-price contracts in the truckload transportation sector have been the prevailing relationship form between firms and their for-hire transportation service providers. While advantageous for planning and budgeting purposes, these contracts do not lend themselves well to contexts in which supply, demand, and market uncertainties are prevalent. We propose an alternative, market-based contract for the truckload sector to reduce the unanticipated costs and performance degradation that result from the standard fixed-price contract. With a uniquely detailed and expansive dataset that captures compounding uncertainties, we build empirical transportation provider decision models. We apply them to index-priced demand across a set of design, implementation, and segmentation strategies and present those that offer Pareto improvements over the status quo fixed-price contract. Finally, we validate our models with a pilot study of indexed contracts implemented by a large agricultural firm in the US. We employ propensity score matching to quantify the causal effect of indexed pricing on transportation suppliers' contract compliance and costs the firm incurs during a constrained market period.

4.1 Introduction

In addition to the traditional fixed-price contract, shippers and contracted carriers may engage in a portfolio of contractual designs to help either encourage contract stickiness from both parties, or mitigate noncompliance and resulting unanticipated costs described above. In this chapter, we explore how shippers and carriers can include an alternative contractual form - index-based contracts - into their contract portfolios to help share the risk brought about by shippers' demand uncertainties, carriers' capacity uncertainties, and market volatility. A market-based contract has received attention from practitioners. For example, some large players such as Schneider - one of the largest transportation providers in the US - along with others have been pushing to scale these dynamic market-based contracts [124]. However, its potential benefit has not yet been demonstrated and remaining concerns over how to design such contracts has obstructed widespread implementation.

There has been extensive research on risk-sharing contracts for supply chain partners (e.g., [35, 36, 110]). The combination of demand- and supply-side uncertainties with market uncertainties observed in the TL transportation context, however, has received little attention, despite the sector's importance to business performance and the overall economy. We aim to address the practical challenges as well as contribute to the literature by demonstrating the contract designs that result in benefits to both buyer and supplier.

The remainder of the chapter is organized as follows. In Section 4.2 we summarize the relevant literature and in Section 4.3 we describe our model formulation and the contract policies and designs we test. In Section 4.4 we detail our carrier acceptance model selection methodology and our steps for choosing the best contract design and policy options. We discuss our modeling results in Section 4.5. In Section 4.6, we describe a case study of a large US shipper implementing an indexed contract pilot. We validate our models and conduct a causal inference study on the impacts of these contracts on carrier performance. Finally, we discuss the research implications in Section 4.7.

In this research, we explore yet another contractual form that aims to mitigate risks and incentivize contract compliance: index-based pricing. Shippers and carriers alike have demonstrated interest in such alternative pricing mechanisms, but few have implemented them. Much of the holdup comes about when discussing the details of how to design the contract.

Thus, we formulate the following research question:

How should shippers and carriers design index-based pricing into their portfolio of transportation contracts that result in a combination of reduced expected costs and higher primary carrier acceptance for the shipper while increasing expected revenue for the carrier?

4.2 Literature Review

In this section, we discuss the relevant literature on contract and transaction cost economics, supply chain contracting under uncertainty, and freight transportation contracting. The extant literature on contract economics, games, and transaction cost economics is vast, and we do not conduct an exhaustive review. Instead, we highlight the relevant themes and how our freight transportation contracting context fits into the larger bodies of literature.

4.2.1 Contract theory

In a very general sense, a contract can be described as a game: a contingent plan specifying the set of actions available to each agent and a reward schedule specifying transfers following each sequence of actions in that game. A contract defines the outcome or payment structure of a set of potential future realizations of the world and resulting decisions. If the agents each have private information, or if information arises over time, there may be stages of communication in which information is revealed.

The foundational economic literature on contract theory, which studies how agents construct contractual agreements, typically in the presence of asymmetric information and an array of other uncertainties, is often attributed to Kenneth Arrow. Agents are commonly modeled as maximizers of von Neumann-Morgenstern utility functions – that is, they seek to maximize their expected utility given a set of probabilities of possible future states of the world [151]. We draw from this literature by employing an expected cost and revenue framework, describe in Section 4.3.

4.2.2 Game theory

The contract literature often seeks the optimal contract in the abstract space of all possible contracts. This is quite a difficult problem. Instead, we take a game theoretic approach, often - though not necessarily explicitly - taken by authors studying supply chain contracting and procurement strategies, in which the performance (e.g., profit or cost outcome) of a set of available contracts and corresponding parameters are compared.

The non-binding TL freight contract is similar to the self-enforcing agreement studied in the context of repeated games literature, which explores agreements that result in both parties maintaining the terms of the agreement over time without the interference of an external party [145]. One of the most prominent results of self-enforcing agreements is the Nash equilibrium. A natural extension of this literature considers the conditions under which agreements are vulnerable to renegotiations after repeated rounds of play [1].

4.2.3 Supply chain contracting and procurement under uncertainty

The theories discussed in the above section have been specifically applied to supply chain procurement. The reader is referred to [36] for a survey of the game theoretic models applied to supply chain management and a review of models that seek to achieve equilibria, particularly in non-cooperative games.

A common topic in supply chain contract literature is that of channel coordination, which identifies the contracts that result in the players' Nash equilibrium coinciding with the optimal contract of the supply chain as a whole. Agents seek such risk-sharing contracts to encourage both sides to remain committed to the contract terms. The risk-sharing, or coordinating supply chain contracts that are commonly found in industry and in literature are the buy-back contract and the revenue-sharing contract [35, 110]. See [134] and [34] for review of these models. In the present chapter, we determine the optimal market-driven index-priced contract design as a means for both shippers and carriers to share the risks incurred by both sides due to the TL market dynamics and uncertainties. We then determine whether their use (Pareto) improve, or coordinate, the buyer-supplier channel (as per [87]).

The benefits of long-term fixed-price versus short-term, dynamic-priced contracts has been explored in [58], [154], and [49].The authors find that long-term, fixed-price contracts may not always be optimal and discuss conditions under which short-term contracts may be justified. Specifically, only when market conditions are stationary - in the mathematical sense in which the mean or variance of the market price do not change over time - over the duration of the contract terms, then the long-term, fixed-price contract is best. The non-stationary market condition represents the freight context with which we are concerned. It underscores why the traditional, fixed-price TL contract may not be suitable given the dynamic nature of freight markets.

As there are trade-offs between selecting particular contract types, a few authors have considered a portfolio approach to supply chain contracting. [155] determines the optimal design of a portfolio of long- and short-term contracts with access to a spot market from the perspective of a gasoline supplier. [100] develops a framework to analyze the performance of a portfolio of supply contracts. The authors construct a portfolio of a long-term contract, an option contract, and spot market transactions. They design the optimal contract parameters, and determine the replenishment policy to maximize expected profits in a multi-period environment. They find that in a general manufacturer-supplier context, the use of portfolio contracts increase expected profit as compared to traditional single long-term contracts while also reducing financial risk. However, the authors do not consider the supplier's revenue as a result of these contracts. We apply the particular nuances of the TL industry to provide recommendations for a portfolio contract strategy consisting of indexed and fixed-price contracts. We further incorporate a channel-coordinating perspective and identify the portfolio of contracts and designs that Pareto improves shipper and carrier outcomes. That is, either one or both parties are made better off (e.g., lower expected cost or higher PAR for the shipper and higher expected revenue for the carrier) without the other being made worse off.

At scale, determining a portfolio contract strategy for shippers is highly complex and may prove to be analytically intractable (e.g. [49]). Thus, we take an empirical approach to numerically determine best strategies.

4.2.4 Freight contracts

The uniqueness of freight contracts has drawn attention in the literature. [39] makes the case for the use of contracts between shippers and carriers. The author demonstrates how economies of scope come about when shippers and carriers make procurement and bid decisions by considering their network as a whole rather than on an individual lane-by-lane basis. A combinatorial auction approach is designed to improve efficiencies.

The merits of contracts in trucking are further studied by [79]. The author models the impact of market thickness - that is, number of buyers and sellers utilizing the market - on shippers' and carriers' use of contracts and finds that as the thickness of long-haul trucking markets increases, simple spot transactions are used by both parties more frequently than are contracts.

Recognizing the uncertainties and price fluctuation in TL transportation and the interest in alternative, more flexible contractual forms, [149] explore the applicability of a futures or options contracts. The authors find that when demand uncertainty is high, some lanes in the US lend themselves well to futures contracts by offering value to both shippers and to carriers. The authors do not, however, explore the characteristics of lanes - or other freight characteristics - that make market-based contracts most attractive. This latter question is addressed in Chapter 3.

[31] and [32] study freight contracting strategies. In [32], given a menu of contract options, the shipper and carrier may choose a single contract to hold over a finite time horizon. [32] includes a form of fixed start-up, or information costs for each contract (i.e., the transaction costs associated with running procurement events and constructing routing guides). [31] discusses the lock-in costs that encourage ongoing interactions between a shipper and carrier when they enter contracts with one another (i.e., the shadow of the future discussed in Chapter 3).

While providing a theoretical foundation for freight contracting, [31] and [32] overlook three critical aspects of the US freight industry landscape. First, the capacity and volume uncertainty introduced by the non-binding nature of freight contracts is omitted. That is, in the author's formulation, the shipper tenders all realized demand to the contracted carrier and the carrier accepts all of this volume at the contract price. Any demand in excess of the awarded volume is served on the spot market. Of course in reality, shippers may tender above or below awarded volumes and a carrier may reject any amount of tendered volume. We explicitly consider this in our expected cost formulation of the freight contracting problem.

Second, Brusset's models assume that any deviation from the contract terms results in immediate termination of the contract. In practice, when a primary carrier does reject load tenders, the contract still remains in effect. Moreover, shippers often tender surge volume to contracted carriers. They may pay the same contract rate, or the contracted carrier may accept the excess volume but at a price premium. Alternatively, the excess volume may be rejected by the primary carrier but picked up by a backup carrier.

Finally, in order to avoid trivial solutions in which the shipper always chooses to use the spot market, the author models cost of using the spot market to be higher than that of using any contract. This is not always the case - in particular, during soft market periods. By adopting such an assumption the authors overlook the fluctuations in market conditions that are pervasive in the freight industry [2, 113].

In this research, we propose a contracting approach built off of empirical models and captures real-world truckload industry complexities not considered in previous literature. The approach aims to offer practical strategies for shippers under more realistic scenarios representing different suppliers (carriers), freight, and network characteristics.

4.3 Model Formulation

In this section, we describe the order of the decisions a shipper makes, the contracting design elements and policies we test and how we measure their performance relative to one another.

4.3.1 Decision sequence

The shipper's decision as to whether and how to implement an index-based contract on a lane with a carrier is formulated as follows. The shipper's procurement event corresponds to time t_0, when the shipper must decide which one of a set of contract policies, $\pi \in \Pi$, it should implement with a single primary carrier, C_p, on a single lane from origin i to destination j, that will be in effect for T periods.

In addition to the contract policy and design, the shipper also decides the initialization price, F_0, which defines both index-based policies. This initialization price is based off of an exogenously determined benchmark price for the lane: BM_{i,j,t_0}. In practice, shippers have access to such benchmark prices through memberships to freight analytics services. It is also information the shipper collects through its reverse auction process. This lane-specific benchmark price completely defines the Fixed price contract on this lane. We assume the exogenous index to be used is predetermined by this shipper and carrier.

The knowledge or information available to the shipper at t_0 is defined by the vector κ. It contains potential contracted carrier's service type (i.e., asset or non-asset), the current lane spot price, S_{i,j,t_0}, and benchmark price, BM_{i,j,t_0}, the current market condition, the previous periods' tendered volumes, $D_{i,j,t}$, spot and benchmark prices, index values, In_t, and historical lane characteristics, $L_{i,j,t}$ (as defined in Section 4.4.2) all for $t < t_0$. The shipper also has an understanding or expectation for each subsequent period of its future demand, spot market and benchmark prices, and the probability the contacted or primary carrier accepts loads at P_π, the contract price for that load defined by the policy chosen in t_0 - that is, $\Pr[A_k \mid C_p, P_\pi]$.

In each subsequent period, $t > t_0$, the index value, spot price, and actual demand are realized. Each load, k, is first tendered to the primary carrier, C_p, at the contracted price, P_π. The primary carrier accepts the load with probability $\Pr[A_k \mid C_p, P_\pi]$. As we discuss in Section 4.4.2, this primary carrier acceptance decision depends on a set of load, lane, and carrier characteristics.

If the primary carrier accepts the load, the shipper pays the offered price, P_π. If the primary carrier rejects the load, the load is offered to backup carriers, C_b. When shippers must rely on backup carriers after primary carrier load rejections, they pay a 'backup premium', bp_t, or a price escalation above the *fixed* contract price on the lane. The value that the backup premium takes on may reach up to 18% or more depending on how deep into the routing guide the shipper must go until a backup carrier accepts the load and the current market conditions (denoted by t) [2, 5]. We model the price a shipper pays backup carriers as $P_{b,t} = bp_t \times BM_{i,j,t_0}$. When tendered a load, k, the backup carrier accepts the load with probability $\Pr[A_k \mid C_b, P_{b,t}]$.

If the backup carrier accepts the load, the shipper pays $P_{b,t}$, otherwise the load is offered on the spot market at the current lane-specific spot price, $S_{i,j,t}$ and is accepted with probability 1. The expected cost setup encapsulates the non-binding nature of the shipper-carrier contact - that is, contracted carriers reject loads with some probability. For each load, k, and policy, π, the shipper has expected cost:

$$\begin{aligned}
\mathrm{E}[\text{cost} \mid \pi]_k = {} & \Pr[A_k \mid C_p, P_\pi] \times P_\pi \\
& + (1 - \Pr[A_k \mid C_p, P_\pi]) \times \Pr[A_k \mid C_b, P_{b,t}] \times P_{b,t} \\
& + (1 - \Pr[A_k \mid C_p, P_\pi]) \times (1 - \Pr[A_k \mid C_b, P_{b,t}]) \times S_{i,j,t}
\end{aligned} \tag{4.1}$$

At the end of the fixed time horizon, T, the performance of each contract strategy and contract design can be summarized by two metrics (1) the total expected cost per mile (CPM) of the strategy:

$$\mathrm{E}[\text{CPM} \mid \pi] = \frac{1}{K} \sum_{\forall k} \frac{\mathrm{E}[\text{cost} \mid \pi]_k}{\text{Dist}_k} \tag{4.2}$$

where K is the total number of loads realized over T and Dist_k is the distance of the lane on which load k is tendered, in miles; and (2) the average primary carrier acceptance ratio over the time horizon:

$$PAR_{C_p,\pi} = \frac{1}{K} \sum_{\forall k} \Pr[A_k \mid C_p, P_\pi] \tag{4.3}$$

Shippers want low expected costs, and high primary carrier acceptance. Thus, the shipper will choose the Indexed contract that offers PAR and expected cost at least as good as the benchmark Fixed price contract. On the other hand, the carrier will agree to the Indexed contract if it results in contract price at least as good at the benchmark Fixed price contract. Thus, we present the design parameters values that offer a Pareto improvement over the Fixed price contract.

4.3.2 Contracting policies

We define the set of contract policies to test based on input from practitioners and previous literature. The indexing policies must be simple enough to implement and agreeable for both shippers and carriers. Our aim is to demonstrate how a small set of indexed pricing policies compare to the baseline, or status quo policy, which is the long-term fixed-price contract. Further extensions of this research may explore additional policies.

Our two policies are defined below and summarized in Table 4.1. First, π_0 is the baseline Fixed-price policy in which all loads are offered at the lane benchmark price, BM_{i,j,t_0}. Second, π_1 is the Indexed policy in which all loads are offered at the indexed price at the time, In_t.

Table 4.1: Contract policy summary

Policy	Name	Offered Price	
π_0	Fixed-price	BM_{i,j,t_0}	$\forall k$
π_1	Indexed	In_t	$\forall k$

4.3.3 Indexed contract design

We expect that the performance of each index-based policy depends on certain contract design choices that the shipper makes upfront. Our experience with shippers and carriers suggests that the largest barriers to implementation of index-based contracts in the TL industry relate to the design details we discuss here.

Index choice

First, the shipper and carrier must agree on the index to be used. It must appropriately reflect the dynamics of the freight to which it is applied. Further it must be transparent in the data that it comprises and generally accepted across the industry [23]. Within the TL Dry Van freight industry, a few such indexes exist. These include several DAT National indexes, the Cass National Truckload Linehaul Index, the Stephens Freight Index, and the Morgan Stanley Freight Index.

We assume the the index chosen is the Cass National TL Linehaul Index. It is public and freely available consistently from 2005 [44]. It is comprised of \$2 billion in transactions, 95% from contract transactions and 5% from spot transactions on lanes that span the continental US. The index is a good representation of the general industry, as this aligns with published estimates of the industry-wide ratio of contract to spot transactions; [40] estimates 90-95% of the total TL market spend is established through contracts.

Further, the Cass TL Linehaul Index is calculated based solely on the linehaul prices of the transactions, which represents contractually defined fixed prices in the market. It does not include fuel surcharges or other accessorial charges added to the base price that reflect specific load or time-based attributes that are sometimes included in other industry indexes.

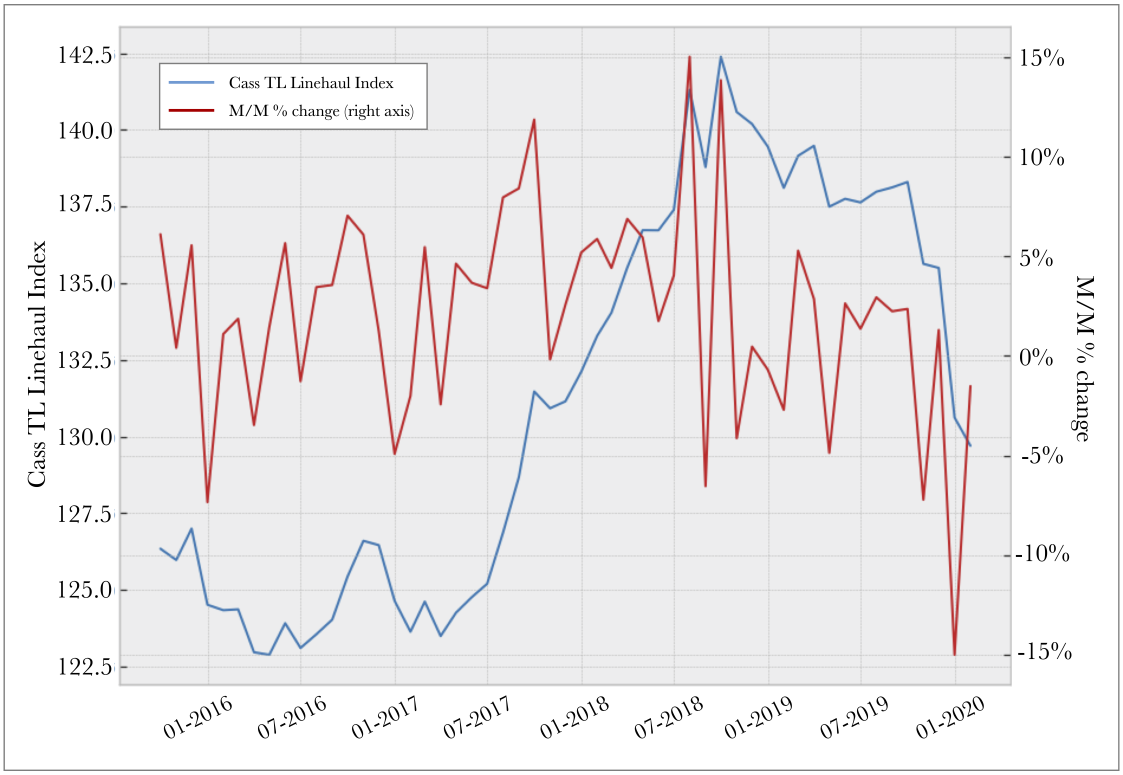
Rather than taking the index value itself, we use the percent change of the index to measure market change - typically week-over-week or month-over-month change, depending on update frequency of the index available. The percent change of the indexed value for time t is defined as $\hat{I}_t = (I_{t-1} - I_{t-2})/I_{t-2}$. The percent change in index value, $\hat{I}_t$, can take both positive and negative values, which allows the indexed price defined in Equation 4.4 to increase or decrease.

We use the index values of the preceding two periods, I_{t-1} and I_{t-2}, because index values are often reported for the previous period. That is, in time t, I_{t-1} would be the most recently available index value. Figure 4-1 shows the Morgan Stanley Index from 2015 to 2020 and the month-over-month percent change.

The offered indexed price for loads during period t, In_t, is defined by increasing or decreasing the indexed price calculated for the previous period, $t-1$, by an amount determined by the percent change of the index in that period as well. The initial period's indexed price is based off of the fixed price, F_0. Thus, we have

$$\begin{aligned} \text{In}_{t_0} &= BM_{i,j,t_0}, & t &= t_0 \\ \text{In}_t &= \text{In}_{t-1}(1+\hat{I}_t), & t &> t_0 \end{aligned} \tag{4.4}$$

Figure 4-1: Cass Truckload Linehaul Index Freight Index, M/M percent change



Collar:

Shippers may choose to implement a collar, or an upper and lower bound on the amount the indexed price can fluctuate. One disadvantage of an indexed price is that it adds price uncertainty for shippers and carriers alike. As a safeguard against large swings in the market, a collar can dampen the price volatility: $\hat{I}_t \in [\underline{c}, \overline{c}]$ and here we assume symmetry $c = |\overline{c}| = |\underline{c}|$:

$$Offered\ Price,\ collar = \begin{cases} \text{In}_t, & \underline{c} * BM_{i,j,t_0} \leq \text{In}_t \leq \overline{c} * BM_{i,j,t_0} \\ \underline{c} * BM_{i,j,t_0}, & \text{In}_t < \underline{c} * BM_{i,j,t_0} \\ \overline{c} * BM_{i,j,t_0}, & \text{In}_t > \overline{c} * BM_{i,j,t_0} \end{cases} \tag{4.5}$$

Where, the lower bound, $\underline{c} < 0$, and the upper bound, $\overline{c} > 0$, represent percentages below and above which the shipper and carrier do not want the indexed load cost change to exceed due to indexing. Note that we apply the collar symmetrically, meaning the indexed price is allowed to increase or decrease by the same amount. An alternative approach, is an escalator design; a special case of the collar design where the indexed price is used as a price increase to incentivize carriers to accept loads but the indexed price does not decrease below the initial price, F_0. That is, $\underline{c} = 0$ and $\overline{c}$ is some positive value or is infinite in the case where the indexed value is unbounded from above.

Initialization price

Finally, the shipper and carrier agree to an initialization price for the Indexed contract in t_0, F_0, as described in Equation 4.4. Shippers want to know how well each policy and design performs if the choice of initialization price, F_0 upfront is too high or too low. Determining what that price should be is difficult due to market uncertainties. Thus, one of our design choices is α, or how F_0 relates to the going lane benchmark price at t_0:

$$F_0 = \alpha * BM_{i,j,t_0} \tag{4.6}$$

Where $\alpha > 0$ and is a multiplier on BM_{i,j,t_0}, the benchmark price on lane (i, j) at time t_0.

We test the relative performance of these contract designs and policies on empirically derived demand. We use an extensive transaction dataset to model the best design choices as a shipper considers the trade offs between expected costs and carrier acceptance rate performance.

4.4 Methodology

In this section, we describe the price mechanisms and policies we test, our model selection methodology, and the model specifications we choose.

4.4.1 Indexed price mechanism design and contract policies

We utilize a highly detailed empirical dataset that contains 2.2 million long-haul (i.e., loads that moved a distance greater than 250 miles)[1], dry van loads, all of which originate and terminate in the continental United States. The observations represent the load tenders and carrier acceptance and rejection decisions for TL loads for 47 shippers and their 308 asset and 74 non-asset primary carriers plus 430 backup carriers. The data span over four years (2015-2019). This encompasses a Soft market period from the beginning of the dataset to July 2017 and from January 2019 onward and a Tight market period from July 2017 to January 2019 (see Chapter 2 and [2] for market period justifications).

In Chapter 3 we describe promising segments of a shipper's freight and network to apply market-based contracts. We split our data according to these segments (see Appendix C.1), apply each indexed design, predict carriers' acceptance behaviors for each, and determine the expected costs and primary carrier acceptance rate under each policy. Thus, we demonstrate the best index-based contract design for each segment.

In addition to the variables and data described above, we have our four empirically trained carrier acceptance prediction models of the probability each load is accepted for each primary carrier service type and market condition combination: (i) Asset primary carriers in a Tight market, (ii) Asset primary carriers in a Soft market, (iii) Non-Asset primary carrier in a Tight market, and (iv) Non-Asset primary carrier in a Soft market. Moreover, we model backup carrier acceptance probabilities as well. The method to train, validate, and test these models is described in detail in Section 4.4.2. The resulting probabilities are inputs to the expected cost formulation of Equation 4.2, which we calculate for each carrier service type and market condition combination.

4.4.2 Carrier acceptance decision model

We build our empirical carrier acceptance decision models on load transaction data described in Section 4.4.1 . First, we must develop an accurate model to predict the probability that a carrier accepts or reject a load, given the characteristics of the load, lane, shipper

[1] We use this long-haul distinction because pricing structures for the alternative, short-haul moves, differ from those we consider and discuss in this research.

tendering behavior, and price. We must also consider whether the carrier is a primary or backup carrier, as the agreements, incentives, and relationships differ between shippers and their primary carriers as compared to their backup carriers (see Section 4.4.2). As such, we develop five distinct carrier acceptance response models: four for each of the primary carrier service type-market combination, and a fifth for backup carriers.

To do so, we split the data into one set of the load tenders from a shipper to its primary carrier on the lane (regardless of if the primary carrier accepts the load), and a second set of all load tenders from a shipper to backup carriers on the lane, which only occur if the primary carrier has rejected the load. We then model the probability that a load is either accepted or rejected (i.e., a binary outcome) by a carrier based on multiple associated load, lane, and carrier characteristics. For brevity, rather than describing each in detail here, we refer the reader to Chapter3 for a detailed description of each variable we include in our model. Table 4.2 below describes each one.

Table 4.2: Carrier Acceptance Model Input Variables

Variable*	Description
Spot Rate Differential (SRD)*	Difference between spot price and offered price, as percentage of offered price
Lane demand cadence	Percent of the preceding 4 weeks in which loads are tendered to primary carrier by shipper on the lane
Lane demand volatility	Four-week rolling average week-over-week percent difference in tendered volume from shipper to primary carrier on the lane
Load surge category	Load's rank within the week relative to the expected weekly volume as a percentage
Asset carrier fleet size*	Number of tractors in the carrier's fleet, categorized
Distance*	Number of miles between origin and destination locations

*Variables that are included in both the primary carrier and the backup carrier acceptance rate models. Otherwise, only included in the primary carrier model.

Imbalanced data

Primary carriers are expected to maintain high acceptance rates. While shippers' service level requirements differ, carriers may be expected to accept at least 90, 95, or even 99% of loads tendered to them. Of course, this level of service may drop due to capacity availability or contract prices becoming out of date as the freight market changes over the course of the contract[2].

In our dataset that spans multiple market cycles, the average acceptance ratio of primary carriers is 82%. In other words, the frequency of observations in our primary carrier dataset in which the load is rejected is much lower than the frequency of the observations in which the load is accepted. Thus, our segmented primary carrier dataset is imbalanced. Predictive classification models trained on imbalanced data are often biased toward the majority class, or the class containing most observations; here, the "accepted" class. In our case, the imbalanced problem only affects the primary carrier model, as the backup carrier dataset is balanced.

The class imbalance problem has been addressed in the literature (e.g., [73, 143]). It is found in many real-world contexts such as credit card fraud detection [30, 12], corporate bankruptcy prediction [6], and rare disease diagnosis [84]. Models based on imbalanced data suffer from high error rates in classifying the minority class because there are many more instances of majority class than minority class observations/footnoteFor example, a model trained to maximize accuracy (the percent of all observations that are correctly classified) on an imbalanced dataset with a ratio of majority to minority class of 80:20 could achieve 80% accuracy by simply predicting all observations to be in the majority class without even considering the observations' features.. Thus, common model performance measures such as accuracy are ill-suited for imbalanced datasets, particularly when good predictive performance is desired for both majority and minority classes.

Algorithmic approach

To obtain models with high predictive performance on our imbalanced data we compare a set of classification algorithms and choose the one and its corresponding best-tuned hyperparameter values [166, 142, 92] for our primary (and backup) carrier load acceptance decisions. See Appendix C.2 for more detailed discussion of the different approaches to dealing with imbalanced data. We test a set of algorithmic methods commonly used in the literature

[2] It is typical for shippers to expect primary carriers to flex up capacity availability and maintain this level of tender acceptance for "surge" volume up to 10% - sometimes more - above the awarded volume.

for binary classification on imbalanced data: Logistic regression (with and without penalty terms), clustering with k-Nearest Neighbors, and ensemble methods, in particular random forests.[3] For each model described below, our input variables are those described in Section 4.4.2 and we model Asset primary carriers and Non-Asset primary carriers in each market period separately.

Logistic Regression: Logistic regression models are a common choice for binary classification. One powerful advantage of the logistic regression model is it allows the modeler to isolate the relationship between each input variable and the binary response variable. However, many real-world data may not abide by the assumptions required to apply (log) linear models. We expect other models may better represent our empirical data with better predictive performance.

Regularization: In addition to the logistic regression model without penalty terms discussed above, we consider three regularization extensions: logistic regression with an L1, or linear penalty term (Lasso); logistic regression with an L2, or quadratic penalty term (Ridge), and logistic regression with a linear combination of L1 and L2 penalty terms (Elastic Net). These three models tend to improve the predictive performance of the non-penalized logistic regression model by minimizing or eliminating the influence of unimportant input variables [147, 71].[4]

Clustering: We consider a clustering technique commonly used in classification problems, k-Nearest Neighbors (kNN). The model makes no assumptions on relationships between variables and is generally easy to interpret [71].

Ensembles: Finally, we consider an ensemble classification technique, a random forest, which is commonly used due to its good classification performance across many data types.

[3]We do not test another common classification approach, the support vector machine (SVM) algorithm, because of long run times associated with SVM and large datasets - i.e., dataset larger than tens of thousands of records, much smaller than our primary carrier tenders dataset.

[4]The L1 penalty forces the estimated coefficient of some input variables to zero in the regression error minimization calculation, resulting in a simpler model and, in many cases, better predictive performance than the non-penalized regression model. On the other hand, the L2 penalty shrinks the coefficients of unimportant input variables, but does not allow them to equal zero. This is advantageous when the model includes input variables that are highly correlated, but still offer important information about the data. The L1 penalty would choose one of the variables to keep and force the coefficients of the others to zero, whereas the L2 penalty would retain all variables but reduce the relative importance of those that contribute to higher error. The Elastic Net regularization allows the benefits of each L1 and L2 and balances the two penalty terms with a regularization parameter, λ.

In fact, random forests have been shown to handle imbalanced data well [28], particularly when the imbalance rate is below 15% [164, 54, 112, 168].

Scoring metric While some classification metrics described in Appendix C.2.3 are better suited for the imbalanced problem than others, they still do not capture the specific problem at hand. In our case, good classification performance of each class is not the most important outcome of the prediction model. Our model is to be used as an input to an expected cost formula. As such, we need a model that has good performance predicting the *probability* an observation is in each class.

There has been limited literature focusing on probability estimates on imbalanced data, however there has been work studying how to obtain good probability predictions in general [106, 167]. For probability predictions, typically a Brier Score is used [29]. The Brier score, BS is the mean square error between the predicted probability an observation is in the positive class, p_n^+, and the actual outcome, $o_n^+ \in \{0, 1\}$. A 1 indicates the observation is in the positive class and a 0 indicates it is not: $BS = (1/N)\sum_{n=1}^{N}(p_n^+ - o_n^+)^2$, where N is the total number of observations in the dataset.

The Brier Score takes values between 0 and 1. Better models have lower Brier Scores. However, this standard Brier Score does not apply well to imbalanced datasets as it does not capture whether a model is biased toward the majority class; The formulation averages the error term over all observations, N. A model that predicts the probability of observations in the majority class well (i.e., low probability predictive error of majority class observations) will still have a low Brier Score even if it poorly predicts the minority class, because there are fewer observations in this class to outweigh the low error on the majority class. To overcome this issue, [153] propose a stratified Brier Score, where we calculate class-specific Brier Scores:

$$\begin{aligned} BS^+ &= \frac{1}{N^+}\sum_{n=1}^{N^+}(p_n^+ - o_n^+)^2 \\ BS^- &= \frac{1}{N^-}\sum_{n=1}^{N^-}(p_n^- - o_n^-)^2 \end{aligned} \tag{4.7}$$

The authors show that the stratified Brier Scores, BS^+ and BS^-, are effective measures of how well a range of different supervised classifiers predict class probabilities for both majority and minority classes of known imbalanced datasets. Thus, we adopt the stratified Brier Scores to evaluate and compare carrier load acceptance prediction models: logistic regression, regularized logistic, kNN, and random forest. We describe our algorithmic approach including model hypertuning and model selection in the following section.

4.4.3 Model hypertuning and comparison

Because no single model performs best on all datasets - imbalanced or otherwise - we adopt a cross validation method to determine the best model to predict the probability a load is accepted by a primary carrier. That is, the model which results in the lowest combination of stratified Brier Scores. Details of our hypertuning process are in Appendix C.3. Results of the model selection for the Asset primary carriers in Tight market conditions are reported in Table 4.3.

Table 4.3: Hypertuned prediction model Brier Scores: Asset, Tight market

Model	Standard BS	Majority BS^+	Minority BS^-
Logistic Regression	0.178	0.178	0.175
L1 Regularization*	0.178	0.177	0.181
k-Nearest Neighbors	0.078	0.032	0.282
Random Forest	0.047	0.036	0.097

*L2 and Elastic Net results excluded because L1 regularization always outperforms L2 in the hypertuning step, and Elastic Net with regularization parameter, $\lambda = 1$ (corresponding to L1), is best.

The results show that the hypertuned random forest outperforms all other hypertuned models in predicting the probability a load is accepted by a primary carrier due to its low majority and minority class Brier Scores. We apply the same method for the dataset of Asset carriers in Tight markets reported above to all other carrier-market condition combinations. The relative performance of the random forest holds in all cases. Thus, we use the balanced random forest as the empirically derived probabilistic model for our expected cost equations.

Although the backup carrier data are balanced, we use a similar backup carrier acceptance model selection method comparing the models' to ensure good predictive performance.

The backup carrier input variables include the load's spot rate differential, origin and destination region indicator variables, and distance. Recall that other shipper tendering behavior, lane, and load characteristics included in the primary carrier model do not pertain to non-contracted, backup carriers. The resulting best model for both carrier segments is the random forest with *uniform* class weighing.

4.4.4 Design and policy choices tests

Next, we test our contract design and policy choices. For each design choice parameter (i.e., choice of α to determine initial fixed price and c to determine size of the collar) we apply a range of values for each parameter for each network segment, market condition, and carrier service type. We calculate the resulting primary carrier acceptance (PAR) (Eq. 4.3) and expected cost per mile (Eq. 4.2) for all parameter values.

Recall, α determines how much above or below the lane benchmark price at which the indexed pricing initiates and is a multiplier on that lane benchmark price. Therefore, a value of 1 corresponds to using the actual lane benchmark price to initialize the indexed contract. We range α from 0.01 to 2.5. For the collar parameter, we range c from 1% to 100%. This allows the indexed price to fluctuate up or down around the fixed benchmark by that capped amount.

The "best" value of α and c depend on a shipper's cost and performance tolerances. Therefore, we report the PAR vs. expected cost per mile curves. These curves show the rate at which PAR increases with changes in expected costs (resulting from dynamic indexed contract prices and reliance on backup or spot options) as we change the design choice parameters - in other words, the marginal increase in PAR for an increase in expected costs. In this way, we capture diminishing returns and assist the shipper in determining the design that captures the most benefit for a cost it deems reasonable.

Once we have the best design choice parameter values, we compare the three policy choices to one another. That is, we compare the PAR and expected costs if the shipper implemented the Indexed policy (π_1) or the Surge only policy (π_2) as compared to the traditional Fixed-price policy (π_0).

4.5 Results

A shipper will choose an Indexed contract that offers lower expected cost or higher primary carrier acceptance, or both, as compared to the status-quo Fixed price contract. We report the contract design parameter conditions for each segment and carrier-market condition combination in which the Indexed contracts results in a Pareto improvement over the Fixed price contract policy. In other words, the segments and their corresponding Index-based contract policy in which both parties are at least as well off and at least one is better off. This occurs when either the shipper experiences lower expected cost or increased PAR and the carrier sees increased expected revenue. In this section, we summarize the key findings.

We present the particular Indexed contract designs for two scenarios:

Case I: The carrier's expected revenue is greater than or equal to that of the benchmark Fixed-price contract **and** the shipper's expected cost is less than or equal to the benchmark Fixed-price contract **and** the shipper sees higher or the same primary carrier acceptance rate (PAR) relative to that of the benchmark Fixed-price contract;

Case II: The carrier's expected revenue is greater than or equal to that of the benchmark Fixed-price contract **and** the shipper sees higher or the same primary carrier acceptance rate (PAR) relative to that of the benchmark Fixed-price contract at some reasonable expected cost increase relative to the benchmark Fixed-price contract. A reasonable increase in expected cost increase is set to 10%, which reflects what a shipper typically pays for its backup carriers between both Soft and Tight markets (see [5]).

Finally, the choice of implementing a collar and corresponding c value impacts the primary carrier acceptance, shipper's expected cost, and carrier's expected revenue to a small extent. At values of $\alpha < 0.5$, the collar does not impact the results to a large extent. At values of $\alpha > 1.1$, the collar threshold value is hit more often. Thus, at the optimal values of α from 0.9-1.1, presented in Tables 4.4 and 4.5, the resulting acceptance rates, costs, and revenues are not impacted by different values of α. As such, we do not report values of c in these summary tables.

The summary of benefits for Case I is reported in Table 4.4 all with no collar on the price. Indexed-based contracts will be agreed to under the Case I conditions only in Tight markets on three of our segments: High Volatility and Moderate and Low Cadence Lanes. In all cases, the initialization price should be set to about 5-10% below the lane benchmark price.

Table 4.4: Indexed contract design improvements, Case I

Segment	Market Condition	Carrier Type	Exp. cost reduction	PAR increase	Exp. rev. increase	α
High Volatility	Tight	Asset	1%	3%	8%	0.95
		Non-Asset	5%	2%	1%	0.90
Mod. Cadence	Tight	Asset	1%	3%	8%	0.95
		Non-Asset	5%	4%	3%	0.90
Low Cadence	Tight	Asset	1%	4%	9%	0.95

On High Volatility lanes, the shipper sees the best reduction in expected cost (5%) with a 2% increase in PAR by implementing the Indexed contract with Asset carriers - who see a 1% increase in expected revenue - rather than Non-Asset carriers. On Moderate Cadence lanes, the Non-Asset carrier offers better expected cost reduction (5%) and PAR increase (4%) than the Asset carrier. Finally, on Low Cadence lanes, only Indexed contracting with Asset carriers results in the conditions for Case I.

Next, we summarize the outcomes for Case II in Table 4.5. Case II differs from Case I in that here, the shipper accepts an increase in expected cost up to the average backup premium of 10%, for an increase in PAR. We expect that in doing so, the shipper will have a higher increase in PAR than in Case I. Similar to Case I, here in order for the carrier to agree to the contract terms, its expected revenue should be at least as good as that of the benchmark Fixed price contract.

Table 4.5: Indexed contract design improvements, Case II

Segment	Market Condition	Carrier Type	Exp. cost **increase**	PAR increase	Exp. rev. increase	α
High Volatility	Tight	Asset	1-4%	8-11%	19-28%	1-1.05
Mod. Cadence	Tight	Asset	4-7%	6-9%	17-26%	1-1.05
Low Cadence	Tight	Asset	3-9%	6-10%	17-33%	1-1.10

The results suggest that Case I and II only occur during Tight market conditions. This is because the freight market index - and resulting index-based price - decreases during

Soft markets. In order for the carrier's expected revenue to increase relative to the Fixed benchmark priced contract in the Soft market conditions, the initialization price would have to be so high relative to the benchmark price that the shipper's expected cost increases well above our 10% threshold. In this case, our models offer carriers a starting point for discussions with shippers to consider their willingness to accept higher expected costs for higher load acceptance from contracted carriers with which a relationship is in place and a history of demonstrated performance.

4.6 Case Study: Index-based TL contracts, pilot experiment in practice

We work with a large US agricultural shipper to implement a pilot of index-based contracts in its long-haul TL network. For purposes of anonymity, we refer to the shipper as AgCorp. We use their load-level transaction and carrier acceptance/rejection data for two purposes: to (i) validate our carrier behavioral models, and (ii) quantify the causal effect of index-based pricing in the freight transportation context.

4.6.1 Pilot implementation and data description

During the height of the Covid-19 pandemic in the US - summer and fall of 2020 - AgCorp ran its annual TL transportation procurement event. Due to concerns of elevated contract price bids it received and carrier acceptance levels plummeting - similar to the rest of the shipper community during this time - AgCorp implemented index-based contracts with two of its Non-Asset carriers.

The contract policy and design choices correspond to those we model and test in the preceding sections. The contracts apply indexed pricing to all of the loads (i.e., policy π_1) on the ten chosen pilot lanes, all originating from a major inbound city in its network. They use the month-to-month actual dollar change in DAT's national TL index to update the indexed contract price monthly with no collar. They choose the initializing price, F_0 to be the lane-level benchmark price defined by DAT at the start of the contract, BM_{i,j,t_0} (i.e.,

$\alpha = 1$). Finally, the contracts are implemented during a Tight market period.

The pilot lane characteristics are characterized by the segments discussed in the preceding sections: award volume, demand cadence and volatility, and surge patterns. On average, the award volume is about 1 load per week. Average volatility is 1.47 (i.e., on average, the week-to-week tendered volume is 147% more than the awarded volume of 1 load per week. For example, 2-3 loads per week). Cadence and Surge volume category splits are reported in Table 4.6. The average length of haul is 1,224 miles.

Table 4.6: Summary of Pilot Lanes

Awd. Volume (# loads/wk)		1
Volatility		1.47
Cadence Category (% of lanes)	0%	15.7
	25%	13.9
	50%	14.8
	75%	20.4
	100%	35.2
Surge Category (% of lanes)	Up to Mean	49.1
	10% above award	4.6
	10-20% above award	0
	>20% above award	46.3
Avg Lane Distance (miles)		1,224

AgCorp provides detailed transaction-level data of every load it tenders to each of its long-haul, TL carriers from October 2018 through April 2020. The data consist of each carrier's accept or reject decision and subsequent tendering if the load is rejected, the load price, and lane details. Moreover, we are provided with the award information: expected volume and price of each awarded carrier for the three bid events AgCorp conducted (2018, 2019, and 2020). As the pilot indexed contract program ("treatment") is implemented in October 2020, we have data that represents pre- and post-implementation shipper and carrier behaviors and decisions. We also have benchmark prices for every lane from the Freight Market Intelligence Consortium (FMIC), a well-established freight analytics group.

4.6.2 Model validation

First, we use these pilot data to validate our carrier acceptance models - the engine of the contract design questions. To do so, we apply our trained non-asset, tight market carrier acceptance model (recall, a "weighted" random forest model using input factors listed in Table 4.2) to the loads that are tendered under the index priced contracts. This subset of the pilot study data is similarly imbalanced to the training dataset, so it is reasonable to expect our models to perform well.

In the imbalanced dataset context, a commonly used measure of the model's ability to predict both majority and minority classes is the precision versus recall curve [123, 80, 86], its Area Under the Curve (AUC), and their harmonic mean, the F1 score [27]. Precision measures the fraction of True Positive to Positive Predictions and Recall measures the fraction of the positive observations that are successfully classified (see Appendix C.2.3. Finally, the F1 score is the weighted average of Precision and Recall. It takes both false positives and false negatives into account:

$$F1 = 2 * (Recall * Precision)/(Recall + Precision)$$

Applying our trained model of Non-Asset carrier in Tight market acceptance to the pilot data, we get an AUC of 98.6% and an F1 score of 96.3%. From these results, we conclude that our trained model is a very good predictor of carrier acceptance for new, unseen data under an actual indexed pricing environment. This validates the results from our contract design models.

4.6.3 Causal effect of index-based freight contracts

Next, we utilize the experimental pilot data and a causal inference approach to isolate the effect of index-based contracts on primary carriers' acceptance rates and costs of loads that the shipper pays. Due to the non-random assignment of carriers and lanes to indexed contracts (i.e., "treatment") we use propensity score matching to estimate the average treatment effect and bootstrap sampling to estimate the standard errors of this two-stage optimization problem [14].

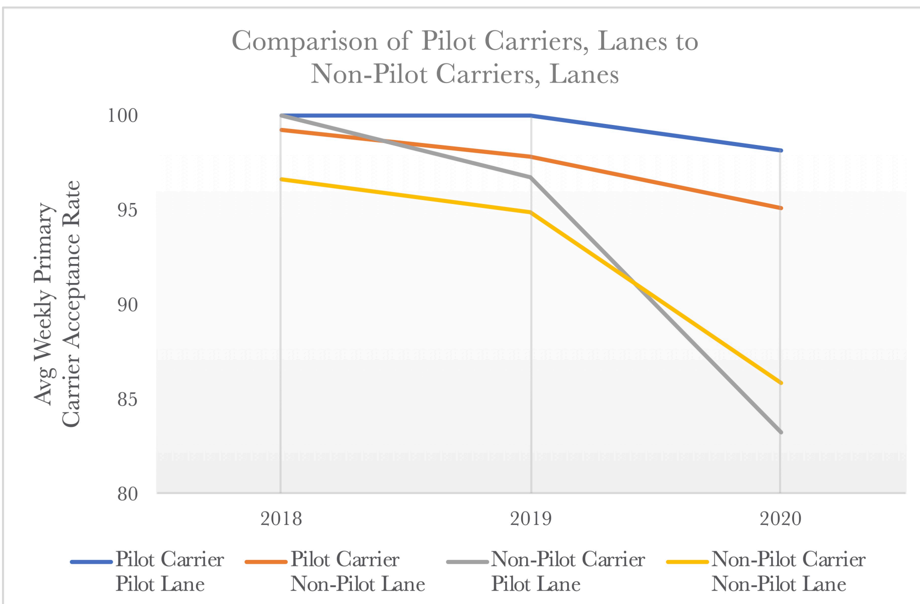
Research question

As the indexed contracts are implemented during a Tight market period, all primary carriers' acceptance rates decline during this time period. However, as demonstrated in Figure 4-2, the treated carriers (i.e., Pilot Carriers on Pilot Lanes) do not show nearly as dramatic a reduction, even compared to themselves on different lanes under Fixed-price contracts (Pilot Carriers on Non-Pilot Lanes). This leads us to the key research questions:

1. Is the higher maintained PAR of Pilot Carriers on Pilot Lanes in the Tight market *because* of the indexed contracts or due to some other confounding carrier or lane factor(s)? - e.g., lane demand patterns such as volume, volatility, cadence, or carrier service type.

2. Do Index-based contracts result in lower costs for the shipper?

Figure 4-2: Primary Carrier Acceptance Rate by pilot carrier-lane combination



4.6.4 Causal Inference Methodology

In control trials as this, it is typical to consider the impact of a policy intervention or treatment on a sample of the population where we have some individuals assigned to a test group that receives the treatment, and a control group of individuals that do not receive treatment. We could measure the average difference in the outcome of interest as the difference between these two populations. But, even if the assignment of individuals to groups were completely randomized we would potentially have uncontrollable, unmeasured, or confounding variables that introduce bias in the estimate of the treatment effect.

Instead, we can use matching techniques [141]. These methods pair treated observations with untreated observations that are similar across a set of potentially confounding factors. In this way, we construct a pseudo-counterfactual for each treated observation. The difference in outcome between the treated and untreated matched observations is the treatment effect. Averaging this value across all matched observations results in the average treatment effect. Figure 4.7 describes the matrix of carrier-lane-treatment conditions and how they relate as potential matches to the treated observations.

Table 4.7: Treated and Untreated Observations and Matching Potential

Treatment Condition	Carrier Lane	Pilot	Non-Pilot
Pre-	Pilot	Good potential match	Potential match
	Non-Pilot	Potential match	Control
Post-	Pilot	Treated	Good potential match
	Non-Pilot	Good potential match	Control

In our pilot setup, the assignment of treatment is not randomized. AgCorp chooses specific carriers with which to implement indexed contracts, and a subset of lanes originating from an important origin location. In these contexts, Propensity Score Matching (PSM) has been demonstrated to result in unbiased, accurate treatment effects [14, 141]. Moreover, PSM outperforms pure matching methods when there are dimensionality concerns.

We choose a set of potentially confounding factors (lane and carrier characteristics) that may be influencing the observed outcome (i.e., primary carrier acceptance rate and load prices). For example, in Chapter 3 we demonstrate that all else equal, Asset carriers and Non-Asset carriers have different acceptance rates and pricing concerns. The Pilot Carriers are all Non-Asset carriers. Thus, we hypothesize that carrier service type is a factor that may be influencing Pilot Carriers' acceptance behaviors. In addition, since the Pilot Lanes all originate from a major depot area within AgCorp's network, they may be particularly attractive lanes than other lanes, which may also be influencing Pilot Carriers' decisions [2]. Thus, the factors we include are those described in Table 4.2 and an indicator of each origin and each destination.

The PSM algorithm is as follows. First, we model the likelihood each observation is in the treated group based on this set of potentially confounding factors. This is typically done with logistic regression.[5] The "treatment" outcome (i.e., load is tendered under Indexed contract) of each observation is regressed on the potentially confounding attributes. The resulting likelihood is the observation's propensity score. This propensity score effectively condenses the confounding factors into one value. Second, each treated observation is matched with an untreated observation based on similar values of propensity score; this results in a pseudo-counterfactual to compare against. Matching is done with kNN, where k=1 and observations' nearness is defined by their propensity scores.

One further challenge with matching methods is identifying proper standard errors for the average effect, as discussed in [122, 77]. Uncertainty arises at both steps of this two-stage approach. That is, there are errors introduced with the classification method used to calculate propensity scores and with the matching algorithm. Few methods have been demonstrated to mitigate this issue and measure confidence intervals [141, 146]. Sampling techniques such as bootstrapping as we employ here have had some support from the literature [132, 121]. Moreover, [88] and [76] show that bootstrapping outperforms other methods.

We implement a sampling technique with over- and under-sampling to account for the small sample number of pilot observations. After calculating the propensity score of all observations, we sample 200 of the treated observations with replacement and 200 with replacement of the untreated observations, as discussed in [55]. We match the two sets with kNN and measure the average treatment effect. We repeat this process 100 times, which results in a distribution of 100 average treatment effects. We take the average over these averages and the 95% confidence interval as our resulting causal effects.

Our research questions for this pilot study consider how well indexed contracts influence

[5] In our study, we use the Generalized Estimating Equations (GEE) method with a binomial family to account for the repeated measures of individual carriers' decisions observed throughout the dataset.

primary carriers to maintain their acceptance rates during the tight market and the costs the shipper incurs. As such, the outcomes we measure are the percent change in (1) PAR and (2) accepted price of the load between pre- and post-implementation; in other words, (1) how well primary carriers sustain lane-level acceptance rates during the tight market of 2020 and (2) the spot rate differential (SRD) between pre- and post-implementation.

Our observations are defined by all the loads tendered for a carrier-lane-week number combination. "Treated" observations are loads tendered to Pilot Carriers on Pilot Lanes in the week numbers corresponding to Oct. 2020 to Apr. 2020 (the range of our dataset corresponding to the pilot period). Untreated observations are any Carrier-Lane combination to which the loads are tendered before the pilot indexed contracts are implemented and any Carrier-Lane combinations after implementation that are not tendered to Pilot Carriers on Pilot Lanes. The outcome of interest for each observation is measured as the pre- to post-treatment percent change in (1) PAR and (2) SRD corresponding to that carrier-lane-week.

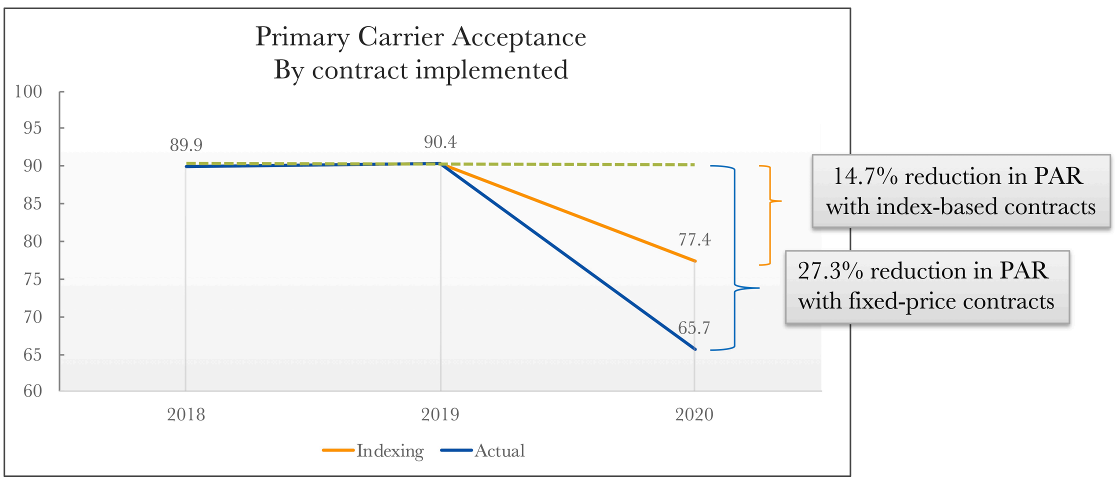
The coefficient results of the logistic regression model, which calculates each observation's propensity score, can be used to understand how important each factor is to matching treated and untreated observations - that is, what factors contribute most to measuring similarities between treated and untreated observations. We find that the most important contributor is whether the carrier is a Pilot Carrier or not. Note that we do not include a Pilot Lane indicator in the propensity score model. This is because we all lanes into their origins, destinations, and distances. By doing so, we provide more potential matches for each treated observation while avoiding multicollinearity in the propensity score regression model.

4.6.5 Causal Inference Results

Controlling for the potential confounding factors and sampling as described in Section 4.6.4, we find that there is a causal effect of indexed contracts on primary carriers' acceptance ratio. Specifically, the average treatment effect is 12.6 with 95% Confidence Interval [12.1-13.1]. That is, if indexed pricing were implemented for all carrier-lanes, the percent change in PAR between fixed price contracts in the soft market and indexed contracts in the following tight market period would have been 12.6 percentage points *higher* than it actually was.

For example, Figure 4-3 depicts the actual PAR degradation AgCorp saw in the tight market of 2020, and what it would have been if it had implemented index-priced contracts. Instead of roughly 90% acceptance in the soft markets of 2018 and 2019 dropping to 65.7% in the tight market (a 27.3% reduction), AgCorp would have seen a decrease to only 77.4% (a 12.9% reduction).

Figure 4-3: Effect of Indexed pricing on PAR



However, we do not find a statistically significant reduction in the accepted cost (SRD) of the loads as a result of the Indexed contracts. These Pilot Lanes are high volatility, low volume, and predominantly moderate cadence lanes with Non-Asset Pilot carriers in Tight markets. Recall from the results in Section 4.5, in order for the shipper to see lower expected costs, the Indexed contract price muse be initialized below the lane benchmark price (i.e., $\alpha < 1$). In this case study, the shipper set the initialization price at each lane's benchmark price. This suggests that the design of the contract is a major contributor to how beneficial it is for the shipper. Moreover, it underscores the importance of our research contributions to practitioners to identify the best contract designs and policies.

4.7 Research Implications and Contributions

In this chapter, we study how a shipper and TL carrier should design and consider implementing index-based contracts into their portfolio of transportation agreements. We demonstrate that during Tight market periods, shippers and carriers can agree to Index-based contracts for High Volatility lanes and Moderate-Low Cadence lanes in Tight freight markets that result in up to 5% lower expected costs for the shipper than the status quo Fixed price contract. This corresponds to an increase in carrier contract compliance (i.e., primary carrier load acceptance rate) of up to 4% while the carriers' expected revenue increases by up to 9%.

To put these numbers into perspective, say a shipper's total annual spend is around $100 million. This is in par with medium sized shippers in the US. Even a 1% reduction in costs would result in $1 million dollars in transportation savings. As one shipper put it, "$1 million dollars saved on transportation is $1 million we can put toward the actual products we sell."

Moreover, shippers may wish to see higher increase in primary carrier acceptance than suggested above and may be willing to increase their expected costs to do so. We show that shippers can increase primary carrier acceptance up to 10% while increasing expected costs to 7-9%. This results in carrier's expected revenue increase of 26-33%. Therefore, we conclude that Index-based truckload freight contracts show promise for both shippers and carriers.

Finally, we isolate and quantify the causal impact of Index-based contracts in Tight market conditions on primary carrier acceptance rates and costs to the shipper in practice. The results underscore the importance of considering the design choices of the contract.

4.7.1 Academic and practical implications

The contributions of this chapter are applicable for both academic and practicing audiences. First, the supplier-buyer contracting literature makes no consideration of the types of contracts we consider here - where supply, demand, and market uncertainty require flexibility in contract compliance for both sides. More specifically, contracts that are non-binding in demand (e.g., loads tendered by a shipper) and in supply (e.g., carrier acceptance of tendered loads). We add to the literature by explicitly considering these contracts, proposing an alternative design, and presenting the ways in which both parties can benefit from these contracts. This serves as a foundation for further experimental and theoretical exploration of these dynamics.

Second, much of the freight transportation contracting literature overlooks the nuances introduced by market dynamics. Simplifying assumptions around the dynamic relationship between spot and contract prices and the nature of these non-binding contracts in practice limit the applicability of the results of previous studies. Moreover, due to difficulty in obtaining proprietary company data, the extant set of literature rarely - if ever - models behaviors and prices on empirical data. Their results often abstract away from real-world challenges. Not only do we build our models and test them on an extensive empirical dataset, but we work with a large US firm to implement index-based contracts to both validate our models and demonstrate the causal effect of them in a real-world setting. Thus, we contribute empirical research to the contract design literature.

The practical contributions of this work stem from the growing interest in index-based contracts. Both buyers and suppliers of truckload transportation seek contract designs to balance the performance and cost trade-offs discussed here. However, the design, strategy, and segmentation choices considered in this research have not been explored to this extent by practitioners. Moreover, our results suggest that certain design decisions - specifically the initialization price - impact the actual effectiveness of the contract. We identify and formulate a set of contract parameters to consider and provide recommendations on how practitioners should approach the design and implementations challenges. We develop models that individual shippers can apply to their businesses to make the performance and cost trade-offs that apply best to their own transportation goals.

The contributions can be generalized to other contexts in which flexible contracts lead to an interest in dynamic pricing. In these settings, we offer a segmentation strategy that applies to demand patterns for any sector, and recommendation for how to construct and design the contracts.

An interesting stream of future research could take an analytical, game theoretic approach to our context of market-based pricing under supply, demand, and market uncertainties. However, the impact of these compounding uncertainties would be difficult to capture and likely lead to intractable solutions. Our empirical approach offers suggestions of the most important aspects to retain in future formulations or related problems.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 5

Summary of Contributions and Extensions

In this thesis, we explicitly address the notion that demand, supply, and macroeconomic uncertainty impact the relationships between shippers and contracted truckload motor carriers. We empirically model these uncertainties and incorporate them into shippers' strategic procurement decisions. We demonstrate under which conditions and relationship forms shippers can experience lower expected costs and higher contracted carrier service levels, and carriers can receive more attractive freight with higher potential revenue than the traditional fixed-price contract.

5.1 Contributions

In this section, we summarize our six research contributions to the field of transportation procurement and operations.

1. Include Uncertainty Using Empirical Models

A key contribution of this thesis is our empirical methodology applied to expand our understanding of the shipper-carrier relationship - and more generally, the buyer-supplier relationship. Using highly detailed microeconomic data covering many shippers and carriers, multiple years, and on lanes across the country, we offer evidence to future research on the incentives and behaviors of these firms. For example, we identify the key factors that predict contracted carrier load acceptance.

Moreover, with traditional analytical approaches, incorporating uncertainty into the problem over-complicates the analysis and can lead to intractable solutions. By explicitly considering the supply, demand, and market uncertainties in this thesis, we contribute more realistic models of shipper and carrier behaviors.

2. Establish Timing of Freight Market Cycles

A well established fact in the TL industry is that the market cycles between periods of over and under supply of trucking capacity (soft and tight markets, respectively). And while the timing of each market period over time has been estimated estimated visually from plots of various freight market indicators by practitioners, no scientific analysis of these fluctuations has been done to determine the precise timing of these cycles. We do so in Chapter 2.

3. Identify Shippers' Actions and Timing to Maintain High Contracted Carrier Acceptance Rates

The results of Chapter 2 suggest that despite long-term relationships between shippers and carriers that span across these market cycles, shippers must be diligent during tight markets. They must maintain short dwell times at drop-off locations, consistent tendering behaviors where possible and, importantly, ensure contract prices are competitive with the going market in order to incentivize contracted carriers to keep load acceptance rates high. Shippers cannot rely on having previously established good performance with a carrier in a soft market and expect that carrier to uphold acceptance rates without continuing that performance through the tight market. Instead, the carrier is beholden to the market constraints and fixed capacity.

Not only does this speak to the shifting power dynamics between shippers and carriers as markets fluctuate, but it demonstrates how the two stages of truckload transportation impact one another. Shippers contract with carriers and the agreements are set for a defined period of time. However, due to the two-sided non-binding nature of these contracts, as the external market conditions change, the contracts can become obsolete.

4. Quantify Carriers' Contract Price Stickiness on Freight and Network Segments

While there is much discussion of alternative contract forms to the standard long-term fixed-price contract, there are benefits to the traditional form. In Chapter 3, we demonstrate for which segments of a shipper's distribution network of lanes, carrier base, and under which market conditions the standard long-term fixed-price contract works well, and where a market-based approach would likely better suit the shipper.

We quantify carriers' contract price stickiness, which is the rate at which they are willing to be pulled from their contracts as the exogenous best-priced alternative option becomes more attractive. In other words, as the spot market price increases relative to the contract price, how likely are they to reject contracted loads and offer their capacity to higher paying shippers.

5. Design Index-based Contracts to (Pareto) Improve the Shipper and Carrier Relationship

We explicitly incorporate the market volatility and resulting uncertainty into the contract design in Chapter 4. We demonstrate how shippers and carriers should considers designing index-based freight contracts as a form of risk sharing. We show that there are designs in which these contracts do in fact result in an improvement for both shippers and carriers. However, an improper design will not have the desired effects. We quantify the potential benefit to both sides - that is, the reduction in expected cost and increase in contracted carrier load acceptance rates for the shipper, and increase in expected revenue for the carrier.

6. Measure the Causal Effect of Index-based Contracts

With empirical modeling approaches, researchers are often limited to making claims of correlations rather than causation. However, we implement a controlled experiment of index-based contracts to measure their causal effect. These results in Chapter 4 underscore the importance of thoughtful design of these contracts for either party to receive the intended benefits.

5.2 Extensions

While the transportation procurement and operations literature is vast, few empirical studies exist. This thesis underscores the complexities of the interactions between shippers and carriers at both stages and the need for more, similar research. Researchers should not miss these opportunities, as the growing trends in data availability are pushing companies to use their data as a competitive advantage.

In addition to expanding on the methodological approaches of this research, there is still much research to be done on the shipper-carrier relationship. New relationship forms have emerged. For example, the “digital only” freight matching platforms such as Uber Freight, Transfix, and Convoy have entered to compete in the freight brokerage space. These platform providers promise better freight for carriers and lower prices for shippers. However, questions remain over whether they are a new carrier type altogether, or if their novelty will fade and there will be a convergence of business models between the traditional freight brokers and these new entrants.

Moreover, the themes and behaviors explored in this thesis and its contributions consider the natural dynamics of freight transportation. However, many of these challenges were brought front and center during the still ongoing Covid-19 global pandemic. In fact, the behaviors we explore and resulting insights are immediately applicable to the extremely tight freight market and irregular recovery patterns that we have experienced over the last two years. Thus, we would be remiss to ignore this as an opportunity to study a natural experiment on the industry. In an effort not to let a good crisis go to waste, as Winston Churchill said[1], a useful stream of further research would explore the impact of Covid-19 on shipper-carrier relationships, how they have evolved, and how shippers and carriers should continue to explore new relationship forms to address the challenges brought about by dynamic freight markets.

[1]Rahm Emanuel, Chief of Staff to President Obama is also credited with this quote

Appendix A

Appendix to Chapter 2

A.1 Sensitivity of Jumping Mean Method

Jumping Mean Method break dates by window length

6	10	12	26	34	52
01-24-16	01-24-16	02-07-16			
05-29-16	05-22-16	05-29-16			
08-21-16			07-24-16	06-12-16	
12-11-16	11-13-16				
01-22-17	01-22-17	01-29-17			
06-04-17	06-04-17	06-25-17	07-23-16		
08-27-17				09-03-17	08-20-17
12-24-17	12-03-17				
01-28-18					
03-04-18	03-04-18				
04-29-18					
06-10-18	06-10-18				
08-05-18		09-02-18			
10-14-18					
01-20-19	01-20-19	01-20-19			

A.2 Regional price premiums

Fixed & per-mile costs: $\hat{\beta}^{Soft}_{base} = \$332.51$; $\hat{\beta}^{Tight}_{base} = \$4.72$; $\hat{\beta}^{Soft}_{dist} = \$1.28/mi$; $\hat{\beta}^{Tight}_{dist} = \$1.36/mi$ All lane distances are over 250 miles (as noted in text), with an average length of haul of 716.4 mi.

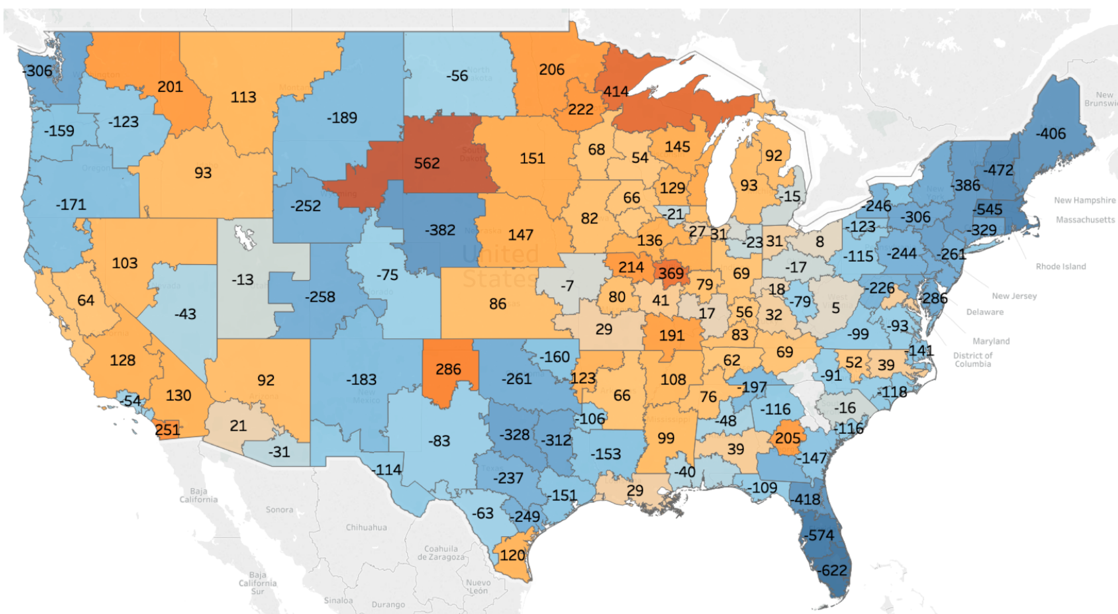


Figure A-1: Origin Region Price Premiums, Soft Period

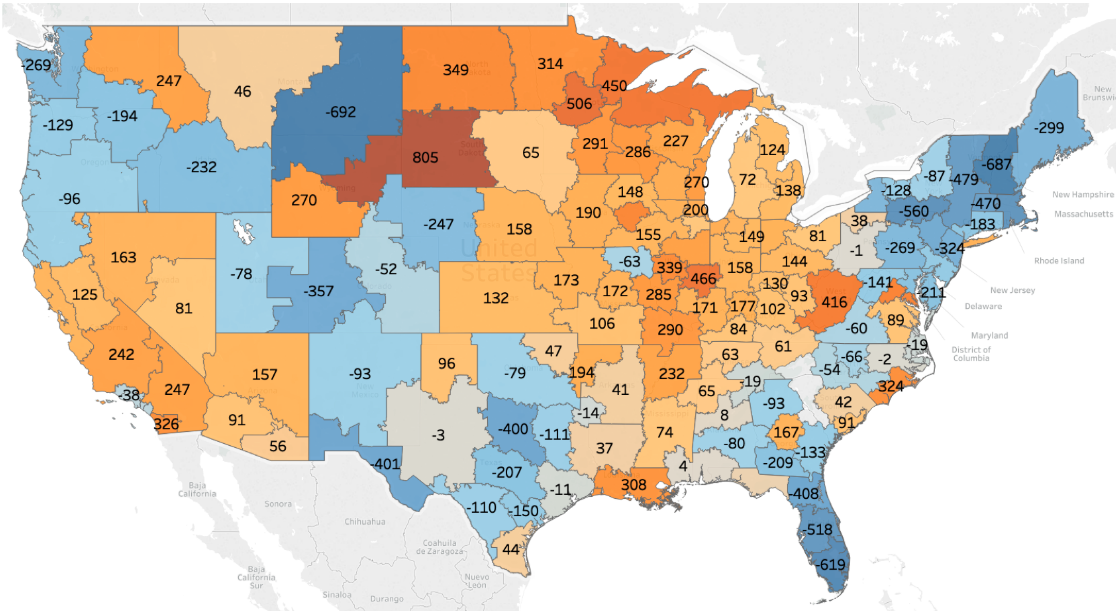


Figure A-2: Origin Region Price Premiums, Tight Period

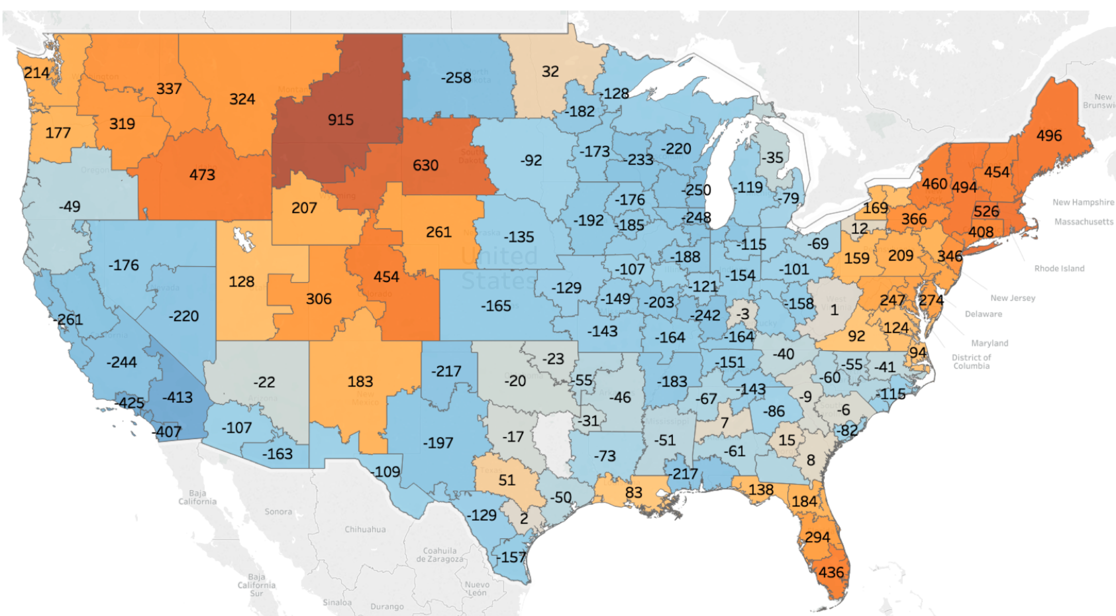


Figure A-3: Destination Region Price Premiums, Soft Period

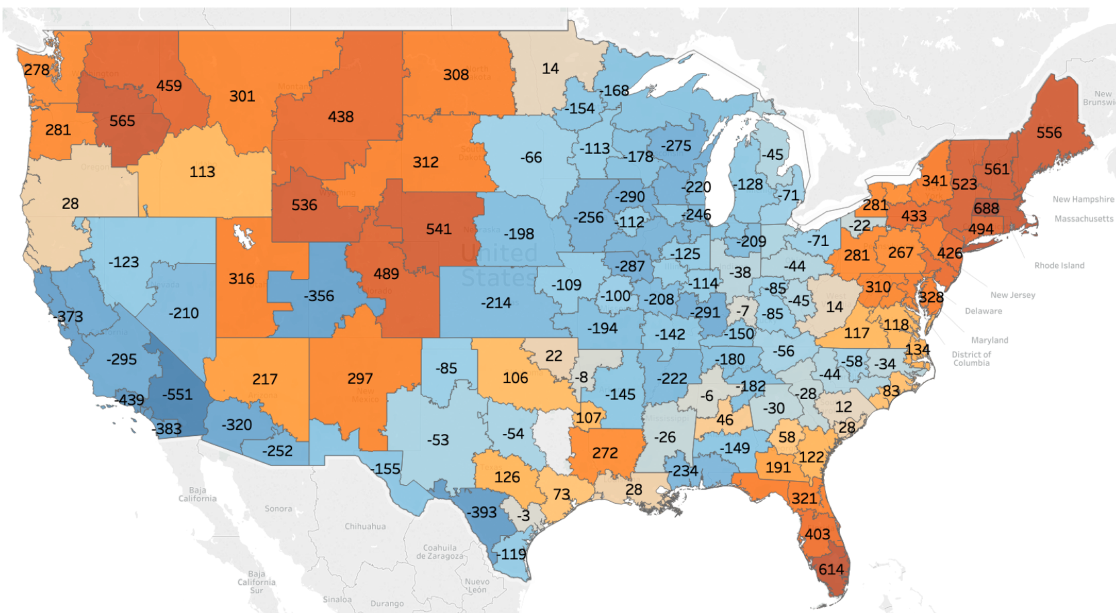


Figure A-4: Destination Region Price Premiums, Tight Period

A.3 Variable summary statistics

Variable	Definition	N	Mean		St. Dev.	
			Soft	Tight	Soft	Tight
Asset-based carriers						
PAR	Average weekly primary carrier acceptance ratio, as a percentage	503	85.9	69.0	20.6	32.2
MRD	Market rate differential: contract rate percentage above or below lane benchmark prices	503	3.06	-6.12	20.91	22.25
Tendering volatility	Coefficient of variation of weekly tendered volume	503	0.429	0.422	0.258	0.261
TLT	Tender lead time (days) between load tendered to primary carrier and when it needs to be picked up	503	4.47	4.61	2.87	3.03
Origin dwell	Amount of time (hours) between when driver arrives at origin empty to leaving with full load	503	2.14	2.22	1.88	2.28
Destination dwell	Amount of time (hours) between driver arrival at destination with full load to leaving empty	503	1.96	2.05	1.09	1.27
Offer cadence	Number of weeks (as a percentage of a 52-week year) in which loads are tendered to the primary carrier	503	50.40		33.21	
Shipper size	Log, shipper total annual volume	503	3.56		0.649	
Automotive	Indicator of shipper industry	159	-		-	
F&B/CPG	Indicator of shipper industry	159	-		-	
Manufacturing	Indicator of shipper industry	76	-		-	
Paper & Packaging	Indicator of shipper industry	89	-		-	
Other	Indicator of shipper industry	20	-		-	
Carrier fleet size	Log of the number of trucks	503	2.52		0.975	

A.4 Variable summary statistics

Variable	Definition	N	Mean		St. Dev.	
			Soft	Tight	Soft	Tight
Non-asset providers						
PAR	(above)	159	88.3	69.9	18.5	32.7
MRD	(above)	159	1.49	-3.52	19.15	23.7
Tendering volatility	(above)	159	0.448	0.423	0.268	0.282
TLT	(above)	159	4.16	4.24	2.72	2.51
Origin dwell	(above)	159	1.77	1.86	0.916	1.06
Destination dwell	(above)	159	1.70	0.846	1.82	1.03
Offer cadence	(above)	159	41.73		20.84	
Shipper size	(above)	159	8.26		1.46	
Automotive	(above)	41	-		-	
F&B/CPG	(above)	53	-		-	
Manufacturing	(above)	32	-		-	
Paper & Packaging	(above)	28	-		-	
Other	(above)	5	-		-	

Appendix B

Appendix to Chapter 3

Table B.1: GEE model results: Spot Rate Differential, asset and non-asset carriers, soft and tight markets

Spot Rate Differential	Asset carriers Soft market	Asset carriers Tight market	Non-asset carriers Soft market	Non-asset carriers Tight market
Constant	2.0032	1.2945	0.634	1.0894
	(0.882)	(1.121)	(0.837)	(1.451)
SRD:	0.4862**	0.6254*	0.3237	0.2436
($< -50\%$)	(0.236)	(0.382)	(0.431)	(0.466)
SRD:	0.6307***	0.7545**	0.5789**	0.4886
$[-50\%, -45\%)$	(0.233)	(0.332)	(0.292)	(0.846)
SRD:	0.3642*	0.1127	0.2974	0.0262
$[-45\%, -40\%)$	(0.202)	(0.272)	(0.425)	(0.415)
SRD:	0.3445	0.242	0.6138**	0.0223
$[-40\%, -35\%)$	(0.233)	(0.290)	(0.279)	(0.432)
SRD:	0.467***	-0.0195	0.5006*	-0.1933
$[-35\%, -30\%)$	(0.183)	(0.245)	(0.267)	(0.436)
SRD:	0.3231*	-0.1767	0.4178*	0.4921
$[-30\%, -25\%)$	(0.195)	(0.233)	(0.239)	(0.444)

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

Table B.2: GEE model results: Spot Rate Differential, asset and non-asset carriers, soft and tight markets, continued

Spot Rate Differential	Asset carriers Soft market	Asset carriers Tight market	Non-asset carriers Soft market	Non-asset carriers Tight market
SRD: $[-25\%, -20\%)$	0.1934 (0.196)	0.028 (0.209)	0.0378 (0.211)	0.7571*** (0.253)
SRD: $[-20\%, -15\%)$	0.2504 (0.200)	-0.1155 (0.198)	0.0674 (0.218)	0.2454 (0.269)
SRD: $[-15\%, -10\%)$	0.2193 (0.171)	-0.1083 (0.177)	0.0619 (0.204)	-0.2322 (0.406)
SRD: $[-10\%, -05\%)$	0.0146 (0.151)	0.1091 (0.179)	0.1712 (0.238)	-0.1569 (0.266)
SRD: $[-05\%, 0\%)$	-0.0314 (0.150)	0.2073* (0.124)	-0.1154 (0.256)	0.3754* (0.202)
SRD: $[0\%, 05\%)$	omitted	omitted	omitted	omitted
SRD: $[05\%, 10\%)$	0.0264 (0.1360)	-0.2026 (0.149)	-0.0804 (0.395)	-0.0055 (0.250)
SRD: $[10\%, 15\%)$	-0.0054 (0.165)	-0.1423 (0.148)	-0.0226 (0.246)	-0.0968 (0.247)
SRD: $[15\%, 20\%)$	0.026 (0.149)	-0.1789 (0.158)	0.1061 (0.309)	-0.2498 (0.361)
SRD: $[20\%, 25\%)$	-0.152 (0.297)	-0.1171 (0.182)	0.5717** (0.267)	-0.8346*** (0.340)
SRD: $[25\%, 30\%)$	0.0913 (0.208)	-0.2481 (0.191)	-0.2945 (0.436)	-0.5186 (0.385)
SRD: $[30\%, 35\%)$	-0.3266 (0.287)	-0.1904 (0.189)	0.7472*** (0.308)	-0.5058 (0.371)
SRD: $[35\%, 40\%)$	-0.1179 (0.300)	0.266 (0.201)	-0.0413 (0.638)	-0.5918** (0.286)
SRD: $[40\%, 45\%)$	-0.2114 (0.325)	-0.1257 (0.225)	1.8206*** (0.695)	-0.7198** (0.314)
SRD: $[45\%, 50\%)$	-0.286 (0.299)	0.0457 (0.217)	1.1126 (0.800)	-0.6647** (0.332)
SRD: $[> 50\%)$	0.0406 (0.275)	-0.1655 (0.207)	0.9579*** (0.320)	-0.8166*** (0.284)

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

Table B.3: GEE model results: Lane Distance (measured in Travel Days), asset and non-asset carriers, soft and tight markets

Travel days	Asset carriers Soft market	Asset carriers Tight market	Non-asset carriers Soft market	Non-asset carriers Tight market
1 day	-0.0152	0.0594	-0.3804**	0.3756*
	(0.167)	(0.168)	(0.192)	(0.212)
2 days	omitted	omitted	omitted	omitted
3 days	0.2134	-0.0518	0.0866	0.6837***
	(0.220)	(0.162)	(0.268)	(0.206)
4 days	0.3755	-0.3351	-0.2664	0.8952**
	(0.260)	(0.244)	(0.363)	(0.409)
>4 days	0.5188	-0.0202	0.3404	0.2598
	(0.402)	(0.373)	(0.309)	(0.268)

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

Table B.4: GEE model results: Lane Tendering Consistency, asset and non-asset carriers, soft and tight markets

Consistency variable	Asset carriers Soft market	Asset carriers Tight market	Non-asset carriers Soft market	Non-asset carriers Tight market
Cadence: 25%	-0.3690*** (0.105)	-0.2296** (0.119)	-0.4965*** (0.195)	-0.3933 (0.249)
Cadence: 50%	-0.1291** (0.063)	-0.1145 (0.088)	-0.1757 (0.141)	-0.0943 (0.152)
Cadence: 75%	omitted	omitted	omitted	omitted
Cadence: 100%	-0.1195 (0.093)	-0.0611 (0.077)	0.243 (0.198)	0.199 (0.200)
Volatility: Up to 10%	0.9023*** (0.243)	0.9203*** (0.219)	0.6908*** (0.169)	0.6318* (0.373)
Volatility: (10-25%]	0.4429*** (0.139)	0.3591*** (0.106)	0.2136 (0.219)	0.5434*** (0.161)
Volatility: (25-50%]	omitted	omitted	omitted	omitted
Volatility: (50-75%]	-0.3142*** (0.065)	-0.3789*** (0.098)	0.0508 (0.132)	0.2485** (0.110)
Volatility: (75-100%]	-0.5076*** (0.086)	-0.6222*** (0.133)	-0.050 (0.177)	0.1623 (0.232)
Volatility: (100-125%]	-0.8049*** (0.119)	-0.8206*** (0.133)	-0.1754 (0.254)	0.1213 (0.210)
Volatility: (125-150%]	-0.7579*** (0.116)	-0.8419*** (0.140)	0.0278 (0.285)	0.1355 (0.220)
Volatility: (150-200%]	-0.6011*** (0.127)	-0.8065*** (0.182)	-0.0155 (0.304)	0.1349 (0.238)
Volatility: Over 200%	-0.4359*** (0.143)	-0.746*** (0.175)	0.1467 (0.313)	0.2419 (0.266)

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

Table B.5: GEE model results: Surge Volume, asset and non-asset carriers, soft and tight markets

Surge Category	Asset carriers Soft market	Asset carriers Tight market	Non-asset carriers Soft market	Non-asset carriers Tight market
Within Mean	0.1757** (0.089)	0.1009 (0.084)	0.0181 (0.251)	0.2668* (0.146)
Mean to 10% Surge	omitted	omitted	omitted	omitted
10-20% Surge	-0.0203 (0.108)	-0.1837** (0.094)	0.1271 (0.294)	0.124 (0.150)
Over 20% Surge	-0.1559** (0.079)	-0.2334*** (0.083)	-0.1387 (0.265)	-0.0099 (0.120)

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

Table B.6: GEE model results: Carrier Fleet Size, asset carriers, soft and tight markets

Fleet size No. tractors	Asset carriers Soft market	Asset carriers Tight market
Log Tractor Count	-0.0109 (0.079)	-0.0088 (0.103)

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

Table B.7: GEE model results: Shipper Fixed Effects, asset and non-asset carriers, soft and tight markets

Shipper fixed effects variable	Asset carriers Soft market	Asset carriers Tight market	Non-asset carriers Soft market	Non-asset carriers Tight market
Shipper size (log monthly volume)	0.4033** (0.177)	0.284 (0.271)	0.5908** (0.27)	0.3054 (0.406)
Vertical: Automotive	0.655*** (0.216)	0.8121*** (0.261)	-0.3843 (0.279)	0.2414 (0.451)
Vertical: F&B/CPG	0.9596*** (0.194)	0.9327*** (0.280)	-0.4351 (0.266)	0.1938 (0.439)
Vertical: Paper & Packaging	omitted	omitted	omitted	omitted
Vertical: Manufacturing	0.5557 (0.381)	0.5942* (0.336)	-0.8417* (0.468)	0.5001 (0.531)
Vertical: Other	-0.2443 (0.602)	0.9402 (0.636)	1.3812*** (0.467)	-0.2445 (0.684)

Note: robust standard errors reported in parentheses
significance level: *0.1; **0.05; ***0.01

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix C

Appendix to Chapter 4

C.1 Segmentation

Figure C-1: Lane Segments: Volume

Lane Volume (#loads/yr)

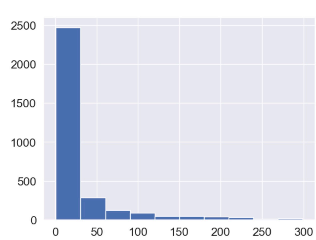
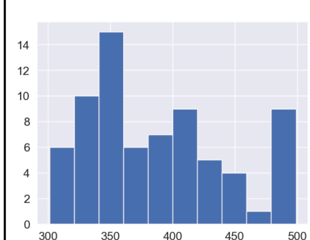
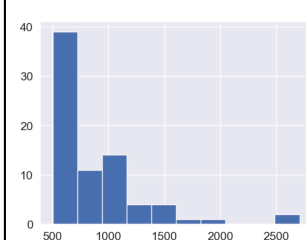
Segment	**Low**	**Moderate**	**High**
Count	3177	72	76
Mean	26.17	385	884.99
StDev	50.19	56.48	430.88
Min	1	301	502
50%	6	372.5	709
Max	299	499	2706
# lanes			

Figure C-2: Lane Segments: Volatility

Lane Volatility (w-w % change tendered volume)

Segment	**Low**	**Moderate**	**High**
Count	57	1414	1828
Mean	0.108	0.638	1.81
StDev	0.066	0.214	0.635
Min	0	0.20	1
50%	0.121	0.644	1.70
Max	0.195	0.999	6.20

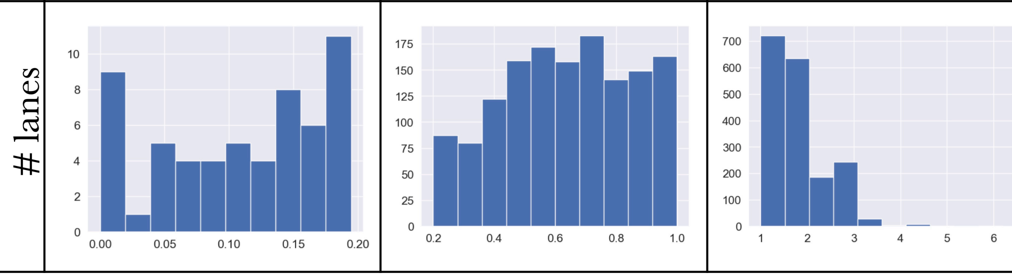
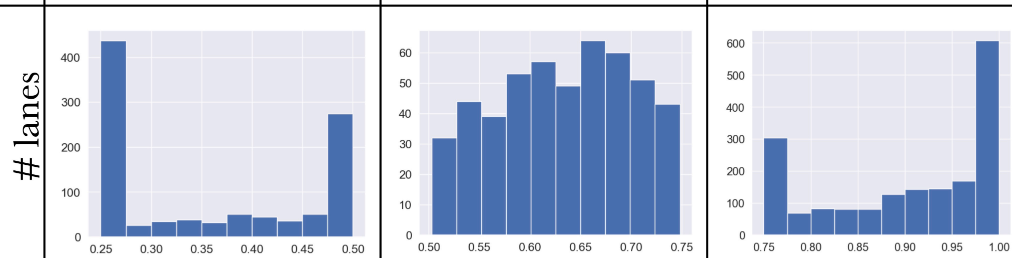


Figure C-3: Lane Segments: Cadence

Lane Cadence (% previous 4 wks with tendered volume)

Segment	**Low**	**Moderate**	**High**
Count	1024	492	1809
Mean	0.357	0.634	0.904
StDev	0.107	0.066	0.092
Min	0.25	0.503	0.75
50%	0.33	0.643	0.927
Max	0.5	0.748	1



C.2 Approaches for Modeling with Imbalanced Data

C.2.1 Pre-processing

The first common pre-processing technique is over-sampling. Either minority class samples are repeated and added to the training dataset or synthetic observations are created through interpolation of neighboring observations in the minority class [47, 72]. The other common pre-processing approach is under-sampling, where observations of the majority class are removed until class proportions are balanced [82] Researchers have shown that a combination of over-sampling and under-sampling can be an effective method to deal with imbalanced data.

However, there are a number of issues that my result from such sampling techniques. First, the optimal class ratio may not be 1:1; [158] finds that the optimal ratio depends on the data characteristics. Further, over-sampling may unintentionally amplify noise contained by the minority class, and if a synthetic data simulation technique is chosen, the best method to do so may not be trivial. Over-sampling has also been found to overfit training data [47]. On the other hand, under-sampling may lead to unintentionally removing majority class data points containing important information. However, [57] find that the drawbacks of under-sampling outweigh those of oversampling.

C.2.2 Cost-sensitive Learning

Each classification model attempts to either maximize a likelihood estimator (MLE) or minimize a loss function to determine each observation's class. It has been shown that adjusting the class weight in these MLE or loss function equations can improve minority class predictions for imbalanced data. It is common to assign weights by inverse class frequency [78, 156] or by a smoothed, inverse square root of class frequency [102, 99, 52]. In some model specifications, class weighting reduces to - or can even outperform - resampling techniques [84, 165, 126, 7].

C.2.3 Scoring metric

Finally, some metrics commonly used to determine model performance are inadequate for imbalanced data classification problems.

Accuracy, for example, does not capture how well a biased model predicts the minority class (here, rejected loads). Measures such as recall and precision are similarly poor measures of the minority class. Recall is the fraction of all truly positive observations that are correctly classified: $Recall = TP/(TP + FN)$, where TP is the true positive predictions, or the number of observations predicted to be in the positive class that are actually in the positive class, and FN is the false negative predictions, or the number of observations predicted to be in the negative class but are actually in the positive class. Precision is the fraction of positive predicted observations that are actually positive $Precision = TP/(TP + FP)$, are good for measuring how well the model predicts the positive (majority, or "accepted") class.

F-scores are a popular choices for imbalanced data, as they are a set of metrics that create a linear combination of precision and recall [80, 86]. Another widely used method that is promising for imbalanced data problems is the receiver operating curve (ROC), which plots FP on the x-axis and TP on the y-axis and its corresponding area under the curve (AUC). These tools allow the user to choose the model parameter values that make the best trade off between misclassification types and do not place more emphasis on one class over the other [27].

Related are cost-sensitive scorers, where rather than including the cost of misclassification in the error term during the model training, those costs are applied to the model's ability to classify each class separately and, weight them based on the known costs of misclassification to achieve a single model score. However, this requires knowledge of costs misclassifying each class.

C.3 Model Hypertuning Process

We split each of our primary carrier datasets (defined by one of four carrier service type {Asset, Non-asset} - market condition {Tight, Soft} combinations into a training and validation set (70% of the primary load tenders) and a test set (30%), stratified so that the proportion of majority ("accepted") to minority ("rejected") class samples in the original, full dataset is maintained in each of the resulting sets. Each model under consideration is defined by a set of hyperparameters. We consider only the set of hyperparameters for each model relating to issues that arise from the imbalanced data problem.

All models take a class weight parameter, which describes how the model weighs prediction errors of each class in its corresponding cost function minimization. The class weight parameter takes value of either "uniform", where there is no re-weighting of classes, or "balanced", where class weights are defined as the inverse of class frequencies so that the minority class is given higher weighting to increase its contribution to total error [84] and [165].

The logistic regression model and the first two regularizations (L1 and L2) do not have additional relevant hyperparameters to tune. The elastic net model includes a regularization parameter, $\lambda \in [0, 1]$, which defines the linear combination of the L1 and L2 penalty terms. We test discrete values of $\lambda \in \{0, 0.1, 0.25, 0.5, 0.75, 0.9, 1\}$, where a value of 0 corresponds to the L2 penalty regularization, and a value of 1 corresponds to the L1 penalty regularization.

The kNN model takes parameters k, the number of neighboring observations to consider in classification prediction, and observation weights, measuring the weight each observation has on the classification decision based on its distance from that observation. We test the set of k values:
$k \in \{5, 10, 25, 30, 35, 40, 45, 50, 60, 75, 100, 125, 200\}$. The observation weights can be either uniform, where all points in each neighborhood are weighted equally, or weighted, where all points are weighted by the inverse of their distance from the observation point.

Finally, the random forest takes parameters for the number of estimators (i.e., single trees) to construct. We test the performance of 10, 50, 100, and 200 estimators. At each node split we allow the algorithm to consider all input features.

For each combination of parameter values and models, we calculate the 5-fold (stratified) cross-validated majority Brier Score, BS^+, minority Brier Score, BS^-, and the standard Brier Score. We choose the set of parameter values that results in the lowest combination of majority and minority Brier Scores. This best parameter combination for each model constitutes the best-tuned, or hypertuned model.

We find that all models are best tuned by a balanced class weight. The best regularization parameter for the elastic net model takes value of $\lambda = 1$ (i.e., the special case corresponding to the L1 regression model). The best trade-off between Brier Scores for the kNN model results in $k = 75$ and uniform observation weights. The best-tuned random forest uses 100 trees. We then apply each best-tuned model to the test data and again calculate the majority and minority Brier Score. We choose the single hypertuned model that results in the best combination of majority and minority Brier Scores on the test data.

While the random forest model does not offer an analytical relationship between input variables and the predicted outcome, we can measure the relative importance of each input feature based on how many important node splits are made in each sub tree for each variable. This is measured by the decrease in Gini impurity score that results from splits on that variable. The most important input feature to correctly predict if a load is accepted is its offered price spot rate differential, followed by lane cadence, origin, destination, carrier service type and size, and lane volatility. These results also correspond to the econometric model results in [3].

Bibliography

[1] Dilip Abreu and David Pearce. A perspective on renegotiation in repeated games. In Reinhard Selten, editor, *Game Equilibrium Models II: Methods, Morals, and Markets*. Springer-Verlag Berlin Heidelberg, 1991.

[2] Angela Acocella, Chris Caplice, and Yossi Sheffi. Elephants or goldfish?: An empirical analysis of carrier reciprocity in dynamic freight markets. *Transportation Research Part E: Logistics and Transportation Review*, 142:102073, 2020.

[3] Angela Acocella, Chris Caplice, and Yossi Sheffi. The end of "set it and forget it" pricing? opportunities for market-based freight contracts. Technical report, MIT Center for Transportation & Logistics, 2021.

[4] Angela Acocella, Sherry Liu, Alex Miller, and Chris Caplice. Should shippers be afraid of ghost freight? Technical report, MIT Center for Transportation & Logistics, 2021.

[5] Nishitha Reddy Aemireddy and Xiyang Yuan. Root cause analysis and impact of unplanned procurement on truckload transportation costs. M. Eng in Logistics Thesis, MIT, 2019.

[6] Edward I. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589–609, 1968.

[7] Takeshi Amemiya and Quang H. Vuong. A comparison of two consistent estimators in the choice-based sampling qualitative response model. *Econometrica*, 55:699–702, 1987.

[8] Nane Amiryan and Sharmistha Bhattacharjee. Relationship between price and performance: An analysis of the us trucking market. M. Eng in Logistics Thesis, MIT, 2015.

[9] Donald Andrews. Tests for parameter instability and structural change with unknown change point. *Econometric Theory*, 61(4):821–856, 1993.

[10] Donald Andrews and Werner Ploberger. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62(6):1383–1414, 1994.

[11] Mikael Apel, Richard Friberg, and Kerstin Hallsten. Microfoundations of macroeconomic price adjustment: Survey evidence from swedish firms. *Journal of Money*, 37:313–338, 2005.

[12] Gerhard Arminger, Daniel Enache, and Thorsten Bonne. Analyzing credit risk data: A comparison of logistic discrimination, classification tree analysis, and feedforward networks. *Computational Statistics*, 12(2):293–310, 1997.

[13] AT Kearney. Annual state of logistics report. *Council of Supply Chain Management Professionals*, 2020.

[14] Peter Austin and Elizabeth Stuart. Estimating the effect of treatment on binary outcomes using full matching on the propensity score. *Statistical Methods in Medical Research*, 26(6):2505–25, 2017.

[15] Jushan Bai. Estimating multiple breaks one at a time. *Econometric Theory*, 13(3):315–52, June 1997.

[16] İsmail Serder Bakal, Nesim Erikip, and Refik Güllü. Value off supplier's capacity information in a two-echelon supply chain. *Annals off Operations Research*, 191(1):115–135, 2011.

[17] Gary A. Ballinger. Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*, 7(2):127–150, 2004.

[18] Ronald Ballou. The accuracy in estimating truck class rates for logistical planning. *Transportation Research - Part A: Policy and Practice*, 25(6):327–337, 1991.

[19] Dawn Barnes-Schuster, Yehuda Bassok, and Ravi Anupindi. Coordination and flexibility in supply contracts with options. *Manufacturing & Service Operations Management*, 4(3):171–207, 2002.

[20] R. Jothi Basu, Nachiappan Subramanian, and Naoufel Cheikhrouhou. Review of full truckload transportation service procurement. *Transportation Reviews*, 35(5):599–621, 2015.

[21] Michael H. Belzer and Stanley A. Sedo. Why do long distance truck drivers work extremely long hours? *Economic and Labour Relations Review*, 29(1):59–79, 2018.

[22] Gerardo Berbeglia, Jean François Cordeau, and Gilbert Laporte. Dynamic pickup and delivery problems. *European Journal of Operational Research*, 202(1):8–15, 2010.

[23] Andrew Bignell. Characteristics of spot-market rate indexes for truckload transportation. M. Eng in Logistics Thesis, Massachusetts Institute of Technology, 2013.

[24] Alan S. Blinder. Why are prices sticky? preliminary results from an interview study. *American Economic Review*, 81:89–96, 1991.

[25] Alan S. Blinder, Elie Canetti, David Lebow, and Jeremy Rudd. Asking about prices: A new approach to understanding price stickiness. *The Journal of Consumer Affairs*, 32(2):424–427, 1998.

[26] BR Williams. What is “index-based pricing” in transportation and logistics and how can it help my organization? Accessed March 29, 2021, 2020.

[27] Andrew P. Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30:1145–1159, 1997.

[28] Leo Breiman. Random forests. *Machine Learning*, 44(1):5–32, 2001.

[29] Glenn W. Brier. Verification of forecasts expressed in terms of probability. *U.S. Department of Commerce Monthly Weather Review*, 78:1–3, 1950.

[30] Iain Brown and Christophe Mues. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, 39(3):3446–3453, 2012.

[31] Xavier Brusset. Choosing a transport contract over multiple periods. *International Journal of Logistics Systems and Management*, 5(3):273–322, 2009.

[32] Xavier Brusset. Multi period contracts in transport under asymmetric information and prior investments. In *Proceedings - Logistics Management Conference*, 2009.

[33] Stephen Burks and Kristen Monaco. Is the u.s. labor market for truck drivers broken. *Monthly Labor Review*, 142(3):1–23, 2019.

[34] G’erard P. Cachon. Supply chain management: Design, coordination, and operation. *Handbooks in Operations Research and Management Science*, 11:227–339, 2003.

[35] G’erard P. Cachon and Martin A. Lariviere. Supply chain coordination with revenue-sharing contracts: Strengths and limitations. *Management Science*, 51(1):30–44, 2005.

[36] G’erard P. Cachon and Serguei Netessine. Game theory in supply chain analysis. *Informs Tutorials in Operations Research*, pages 200–233, 2006.

[37] Erik Caldwell and Bryan Fisher. The impact of lead time on truckload transportation rates. M. Eng in Logistics Thesis, MIT, 2008.

[38] Shutao Cao, Wei Dong, and Ben Tomlin. The sensitivity of producer prices to exchange rates : Insights from micro data. Technical Report 2012-20, Bank of Canada Working Paper, 2012.

[39] Chris Caplice. *Optimization Based Bidding, a New Framework for Shipper-Carrier Relationships*. PhD thesis, Massachusetts Institute of Technology, 1996.

[40] Chris Caplice. Electronic markets for truckload transportation. *Production and Operations Management*, 16(4):423–436, 2007.

[41] Chris Caplice, Francisco Jauffred, Mike Bright, and Michael Disney. Optimal transportation portfolio management at walmart stores. Project Summary, 2011.

[42] Chris Caplice and Yossi Sheffi. Optimization-based procurement for transportation services. *Journal of Business Logistics*, 24(2):109–128, 2003.

[43] Chris Caplice and Yossi Sheffi. Combinatorial auctions for truckload transportation. *Algorithmic Game Theory*, pages 267–300, 2005.

[44] Cass Information Systems. Cass national truckload linehaul index, 2020. https://www.cassinfo.com/freight-audit-payment/cass-transportation-indexes/october-2020, accessed 2020-10-07.

[45] Stephen G. Cecchetti. The frequency of price adjustment: A study of the newsstand prices of magazines. *Journal of Econometrics*, 31:255–274, 1986.

[46] C.H. Robinson. Do 'favored shippers' really receive better pricing and service? White Paper, 2015.

[47] Nitech V. Chawla, Kevin W. Bawyer, Lawrence O. Hall, and W. Phillip Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2003.

[48] Gregory Chow. Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3):591–605, 1960.

[49] Morris A. Cohen and Narendra Agrawal. An analytical comparison of long and short term contracts. *IIE Transactions*, 31(8):783–796, 1999.

[50] Convoy. A new approach to primary freight. Industry White Paper, Accessed April 1, 2021, 2020.

[51] Bob Costello. Trucking economic update. In *2018 ATA Trucking Outlook*, 2018. American Trucking Association.

[52] Yin Cui, Menglin Jia, Yang Song, and Serge Belongie. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

[53] Carlos Daganzo. Transportation costs. chap. 2.3. In *Logistics Systems Analysis*, pages 23–29, 2005. 4th ed.

[54] Mostafa El Habib Daho, Nesma Settouti, Mohammed El Amine Lazouni, and Mohammed El Amine Chikh. Weighted vote for trees aggregation in random forest. In *International Conference on Multimedia Computing and Systems*, pages 438–443, 2014. ICMCS.

[55] Rajeev H. Dehejia and Sadek Wahba. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448):1053–1062, 2002.

[56] Raymond Deneckere, Howard P. Marvel, and James Peck. Demand uncertainty and price maintenance: Markdowns as destructive competition. *The American Economic Review*, 87(4):619–641, 1997.

[57] Chris Drummond and Robert C. Holte. C4.5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling. *Workshop on learning from imbalanced datasets II*, 11:1–8, 2003.

[58] Gary D. Eppen and Anantha V. Iyer. Backup agreements in fashion buying - the value of upstream flexibility. *Management Science*, 43(11):1469–1484, 1997.

[59] Silvia Fabiani, Martine Druant, Ignacio Hernando, Claudia Kwapil, Bettina Landau, Claire Loupias, Fernando Martins, Thomas Matha, Roberto Sabbatini, Harald Stahl, and Ad C. J. Stokman. What firms' surveys tell us about price-setting behavior in the euro area. *International Journal of Central Banking*, 2:3–48, 2006.

[60] Silvia Fabiani, Claire Suzanne Loupias, Fernando Manuel Monteiro Martins, Roberto Sabbatini, et al. *Pricing decisions in the euro area: how firms set prices and why.* OUP USA, 2007.

[61] Silvia LP Ferrari and Francisco Cribari-Neto. Beta regression for modeling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815, 2004.

[62] Jorge I Figueroa-Zúñiga, Reinaldo B Arellano-Valle, and Silvia LP Ferrari. Mixed beta regression: A bayesian perspective. *Computational Statistics & Data Analysis*, 61:137–147, 2013.

[63] William H. Greene. *Econometric Analysis*. Pearson Education, fifth edition, 2003.

[64] Gianfranco Guastaroba, Renata Mansini, and M. Grazia Speranza. Modeling the pre-auction stage: The truckload case. In *Innovations in Distribution Logistics*, pages 219–233. Springer, 2009.

[65] Yunsong Guo, Andrew Lim, Brian Rodrigues, and Yejun Zhu. Carrier assignment models in transportation procurement. *Journal of the Operational Research Society*, 57(12):1472–1481, 2006.

[66] Simon Hall, Mark Walsh, and Anthony Yates. Are uk companies' prices sticky? *Oxford Economic Papers*, 52:425–446, 2000.

[67] Bruce Hansen. Approximate asymptotic p-values for structural-change tests. *Journal of Business and Economic Statistics*, 15(1):60–67, 1997.

[68] Bruce Hansen. The new econometrics of structural change: Dating breaks in u.s. labor productivity. *Journal of Economic Perspectives*, 15(4):117–128, 2001.

[69] James Hardin and Joseph Hilbe. *Generalized Estimating Equations (GEE)*, volume 99. Chapman & Hall, 2012.

[70] Matthew Harding. Can shippers and carriers benefit from more robust transportation planning methodologies? M. Eng in Logistics Thesis, MIT, 2005.

[71] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Data mining, inference, and prediction. In *The Elements of Statistical Learning*, 2016. 2nd ed.

[72] Haibo He, Yang Bai, Edwardo A. Garcia, and Shutao Li. Adasyn: Adaptive synthetic sampling approach for imbalanced learning. *IEEE international joint conference on neural networks (IEEE world congress on computational intelligence*, pages 1322–1328, 2008.

[73] Haibo He and Edwardo A. Garcia. Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9):1263–1284, 2009.

[74] Jan B. Heide and Anne S. Miner. The shadow of the future: Effects of anticipated interaction and frequency of contact on buyer-seller cooperation. *Academy of Management Journal*, 35(2):265–291, 1992.

[75] Chanel M. Helper, Lauren B. Davis, and Wenbin Wei. Impact of demand correlation and information sharing in a capacity constrained supply chain with multiple-retailers. *Computers and Industrial Engineering*, 59(4):552–560, 2010.

[76] Jennifer L. Hill and Jerome P. Reiter. Interval estimation for treatment effects using propensity score matching. *Stat Med*, 25(13):2230–56, 2002.

[77] Jennifer L. Hill, Donald B. Rubin, and Neal Thomas. The design of the new york school choice scholarship program evaluation. In *Research Designs: Inspired by the Work of Donald Campbell*, 1999.

[78] Chen Huang, Yining Li, Chen Change Loy, and Xiaoou Tang. Learning deep representation for imbalanced classification. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[79] Thomas N. Hubbard. Contractual form and market thickness in trucking. *The RAND Journal of Economics*, 32(2):369–386, 2001.

[80] Estabrooks A. Japkowski. A mixture-of-experts framework for learning from imbalanced data sets. In F. Hoffmann, D.J. Hand, N. Adams, D. Fisher, and G. Guimaraes, editors, *Advances in Intelligent Data Analysis*, number 2189 in 1, 2001.

[81] J.B. Hunt. 660 minutes: How improving driver efficiency increases capacity. Industry White Paper, 2015.

[82] Taghi M. Khoshgoftaar, Chris Seiffert, Jason Van Hulse, Amri Napolitano, and Andres Folleco. Learning with limited minority class data. *IEEE Sixth International Conference on Machine Learning and Applications*, pages 348–353, 2007.

[83] Yoo Joon Kim. Analysis of truckload prices and rejection rates. M. Eng in Logistics Thesis, MIT, 2013.

[84] Gary King and Langche Zeng. Logistic regression in rate events data. *Political Analysis*, 9:137–163, 2001.

[85] Roger Kirk. Procedures for the behavioral sciences. In *Experimental design*, 2013. 4th ed.

[86] Sotiris Kotsiantis, Dimitris Kanellopoulos, and Panayiotis Pintelas. Handling imbalanced datasets : A review. *Science*, 30:25–36, 2006.

[87] Martin A. Lariviere. Supply chain contracting and coordination with stochastic demand. In Sridhar Tayur, Ram Ganeshan, and Michael Magazine, editors, *Quantitative Models for Supply Chain Management*, chapter 8, pages 233–268. Springer, 1999.

[88] Michael Lechner. Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. *Journal of the Royal Statistical Society Series A*, 165(1):59–82, 2002.

[89] Chi-Guhn Lee, Roy H Kwon, and Zhong Ma. A carrier's optimal bid generation problem in combinatorial auctions for transportation procurement. *Transportation Research Part E: Logistics and Transportation Review*, 43(2):173–191, 2007.

[90] Tae Whee Lee and Ronaldo Po. Optimizing the use of dedicated and contract transportation assets to maximize total system profit. M. Eng in Logistics Thesis, MIT, 2007.

[91] Kung-Yee Liang and Scott L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986.

[92] Weiwei Lin, Ziming Wu, Longxin Lin, Angzhan Wen, and Jin Li. Cooperative profit random forests with application in ocean front recognition. *IEEE Access*, 5:16568–16575, 2017.

[93] C Lindsey, A Frei, H S Mahmassani, H Alibabai, Y W Park, D. Klabjan, M. Reed, G. Langheim, and T. Keating. Online pricing and capacity sourcing for third-party logistics providers. White Paper, 2015.

[94] C Lindsey and H S Mahmassani. Measuring carrier reservation prices for truckload capacity in the transportation spot market choice experiment. *Transportation Research Record*, 2478:123–132, 2015.

[95] Todd D Little. *The Oxford handbook of quantitative methods, Volume 1: Foundations.* Oxford University Press, 2013.

[96] Claire Loupias and Roland Ricart. Price setting in the french manufacturing sector: New evidence from survey data. *Revue d'Economie Politique*, 116:541–554, 2006.

[97] Claire Loupias and Patrick Sevestre. Costs, demand, and producer price changes. *Review of Economics and Statistics*, 95(1):315–327, 2013.

[98] Chin-Shan Lu. The impact of carrier service attributes on shipper-carrier partnering relationships: A shipper's perspective. *Transportation Research Part E: Logistics and Transportation Review*, 39(5):399–415, 2003.

[99] Dhruv Mahajan, Ross Girshick, Vignesh Ramanathan, Kaiming He, Manohar Paluri, Yixuan Li, Ashwin Bharambe, and Laurens van der Maaten. Exploring the limits of weakly supervised pretraining. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[100] Victor Martínez-de-Albéniz and David Simchi-Levi. A portfolio approach to procurement contracts. *Production and Operations Management*, 14(1):90–114, 2005.

[101] Scott E. Masten. Long-term contracts and short-term commitment: Price determination for heterogeneous freight transactions. *American Law and Economics Review*, 11(1):79–111, 2010.

[102] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26, pages 3111–3119. Curran Associates, Inc., 2013.

[103] Michael H. Morris and Mary L. Joyce. How marketers evaluate price sensitivity. *Industrial Marketing Management*, 17(2):169–176, 1988.

[104] Michael Jay Mulqueen. Creating transportation policy in a network that utilizes both contract carriers and an internally managed fleet. M. Eng in Logistics Thesis, MIT, 2006.

[105] John Ashworth Nelder and Robert WM Wedderburn. Generalized linear models. *Journal of the Royal Statistical Society: Series A*, 135(3):370–384, 1972.

[106] Alexandru Niculescu-Mizil and Rich Caruana. Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning*, 2005.

[107] Raydonal Ospina and Silvia LP Ferrari. A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis*, 56(6):1609–1623, 2012.

[108] Fabrice Lumineau Joanne E. Oxley. Let's work it out (or we'll see you in court): Litigation and private dispute resolution in vertical exchange relationships. *Organization Science*, 23(3):820–834, 2012.

[109] Leslie E Papke and Jeffrey M Wooldridge. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of applied econometrics*, 11(6):619–632, 1996.

[110] Barry Alan Pasternack. Optimal pricing and return policies for perishable commodities. *Marketing Science*, 4(2):166–176, 1985.

[111] Tarciana L Pereira and Francisco Cribari-Neto. Detecting model misspecification in inflated beta regressions. *Communications in Statistics-Simulation and Computation*, 43(3):631–656, 2014.

[112] Todd Perry and Mohamed Bader-El-Den. Proceedings of the companion publication of the 2015 annual conference on genetic and evolutionary computation. In *Imbalanced classification using genetically optimized random forests*, pages 1453–1454, 2015.

[113] Chris Pickett. The coyote curve: A model for mitigating risk and uncertainty in modern supply chain operations. Coyote Logistics, 2018.

[114] Warren Powell. A stochastic formulation of the dynamic assignment problem, with an application to truckload motor carriers. *Transportation Science*, 30(3):195–219, 1996.

[115] Warren Powell, Patrick Jaillet, and Amedeo Odoni. Stochastic and dynamic networks and routing; chap. 3. In *Handbooks in Operations Research and Management Science*, 1995. Vol. 8.

[116] Warren Powell, Yossi Sheffi, Kenneth Nickerson, Kevin Butterbaugh, and Susan Atherton. Maximizing profits for north american van lines' truckload division: A new framework for pricing and operations. *Interfaces*, 18(1):21–41, 1988.

[117] Richard Quandt. Tests of the hypothesis that as linear regression system obeys two separate regimes. *Journal of American Statistical Association*, 55(290):324–330, 1960.

[118] Amelia C. Regan, Hani S. Mahmassani, and Patrick Jaillet. Evaluation of dynamic fleet management systems. In *Transportation Research Record*, pages 176–184, 1998.

[119] Monia Rekik and Sehl Mellouli. Reputation-based winner determination problem for combinatorial transportation procurement auctions. *Journal of the Operational Research Society*, 63(10):1400–1409, 2012.

[120] Lloyd M Rinehart, James A Eckert, Robert B Handfield, Thomas J Page, and Thomas Atkin. An assessment of supplier-customer relationships. *Journal of Business Logistics*, 25(1):25–62, 2004.

[121] Sherri Rose. Consistent estimation of propensity score functions with oversampled exposed subjects. Technical report, Department of Health Care Policy, Harvard Medical School, 2018.

[122] Donald B. Rubin and Neal Thomas. Matching using estimated propensity scores: Relating theory to practice. *Biometrics*, 52(1):249–264, 1996.

[123] Takaya Saito and Marc Rehmsmeier. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10:1–21, 2015.

[124] Schneider National. Market index pricing benefits shippers in all market cycles, 2021. https://vimeo.com/600087431?embedded=true&source=video_title&owner=109806718. accessed 2021-11-15.

[125] Schneider Transportation. Management optimizes packaging manufacturer's spot freight with index pricing model. Industry White Paper, 2019.

[126] A.J. Scott and A.J. Wild. Fitting logistic models under case-control or choice based sampling. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48:170–182, 1986.

[127] Alex Scott. The value of information sharing for truckload shippers. *Transportation Research Part E: Logistics and Transportation Review*, 81:203–214, 2015.

[128] Alex Scott. Carrier bidding behavior in truckload spot auctions. *Journal of Business Logistics*, 39(4):267–281, 2018.

[129] Alex Scott, Christopher Craighead, and Chris Parker. Now you see it, now you don't: Explicit contract benefits in extralegal exchanges. *Production and Operations Management*, 29(6):1467–1486, 2020.

[130] Alex Scott, Chris Parker, and Christopher Craighead. Service refusals in supply chains: Drivers and deterrents of freight rejection. *Transportation Science*, 51(4):1086–1101, 2016.

[131] Alex Scott, Chris Parker, and Christopher W Craighead. Service refusals in supply chains: Drivers and deterrents of freight rejection. *Transportation Science*, 51(4):1086–1101, 2017.

[132] Jasjeet S. Sekhon. The neyman-rubin model of causal inference and estimation via matching methods. *The Oxford handbook of political methodology*, 2:1–32, 2008.

[133] Hugo Simão, Jeff Day, Abraham George, Ted Gifford, John Nienow, and Warren Powell. An approximate dynamic programming algorithm for large-scale fleet management: A case application. *Transportation Science*, 43(2):178–197, 2009.

[134] David Simchi-Levi, Xin Chen, and Julien Bramel. *The Logic of Logistics*. Springer, 3 edition, 2014.

[135] Omar Singh. Perfectly executed routing guides fail 100% of the time. Surge Transportation, 2021.

[136] Ranjit Singh, Ramesh Billa, and Bhupinder Singh Arora. Index based pricing: Managing risk and profitability. *Deloitte*, 2016.

[137] Atmaja Sinha and Rakesh Thykandi. Alternate pricing model for transportation contracts. M. Eng in Logistics Thesis, MIT, 2019.

[138] Jennifer Smith. Trucking's tight capacity squeezes us businesses. *The Wall Street Journal*, August 9 2019.

[139] David Sokoloff and Gaohui Zhang. Predicting and planning for the future: North american truckload transportation. M. Eng in Logistics Thesis, MIT, 2020.

[140] Jiongjiong Song and Amelia Regan. Approximation algorithms for the bid construction problem in combinatorial auctions for the procurement of freight transportation contracts. *Transportation Research Park B: Methodological*, 39(10):914–933, 2005.

[141] Elizabeth Stuart. Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1):1–21, 2010.

[142] Jianyuan Sun, Guoqiang Zhong, Junyu Dong, Hina Saeeda, and Qin Zhang. Cooperative profit random forests with application in ocean front recognition. *IEEE Access*, 5:1398–1408, 2017.

[143] Yanmin Sun, Andrew KC Wong, and Mohamed S. Kamel. Classification of imbalanced data: A review. *International journal of pattern recognition and artificial intelligence*, 23(4):687–719, 2009.

[144] Junichi Takeuchi and Kenji Yamanishi. A unifying framework for detecting outliers and change points from time series. *IEEE Transactions on Knowledge and Data Engineering*, 18(4):482–492, April 2006.

[145] Lester G. Telser. A theory of self-enforcing agreements. *The Journal of Business*, 53(1):27–44, 1980.

[146] Joseph Terza. Simpler standard errors for two-stage optimization estimators. *The Stata Journal*, 16(2):368–385, 2016.

[147] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society - Series B (Methodological)*, 58:267–288, 1996.

[148] Darsono Tjokroamidjojo, Erhan Kutanoglu, and G. Don Taylor. Quantifying the value of advance load information in truckload trucking. *Transportation Research Part E: Logistics and Transportation Review*, 42(4):340–357, 2006.

[149] Mei-Ting Tsai, Jean Daniel Saphores, and Amelia Regan. Valuation of freight transportation contracts under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 47(6):920–932, 2011.

[150] Uber Freight. Uber freight expands procurement solutions with market access. Accessed March 29, 2021, 2021.

[151] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 2 edition, 1947.

[152] Wall Street Journal. Trucking Company Shutdowns Grow as Shipping Market Cools, 2019. https://www.wsj.com/articles/trucking-company-shutdowns-grow-as-shipping-market-cools-11567104495, accessed 2021-10-28.

[153] Bryon C. Wallace and Issa J. Dahabreh. Improving class probability estimates for imbalanced data. *Knowledge and Information Systems*, 41:33–52, 2014.

[154] Chong Wang and Xu Chen. Optimal ordering policy for a price-setting newsvendor with option contracts under demand uncertainty. *International Journal of Production Research*, 53(20):6279–6293, 2015.

[155] Shanshan Wang. *Contract Portfolio Optimization for a Gasoline Supply Chain*. PhD thesis, University of Chicago Booth School of Business, 2011.

[156] Yu-Xiong Wang, Deva Ramanan, and Martial Herbert. Learning to model the tail. In *31st Conference on Neural Information Processing Systems*, 2017.

[157] Kenneth H. Wathne and Jan B. Heide. Opportunism in interfirm relationships: Forms, outcomes, and solutions. *Journal of Marketing*, 64(4):36–51, 2000.

[158] Gary M. Weiss and Foster Provost. Learning when training data are costly: The effect of class distribution on tree induction. *Journal of artificial intelligence research*, 19:315–354, 2003.

[159] Oliver Williamson. *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York, 1975.

[160] Oliver Williamson. Ch. 1 transaction cost economics; ch. 2 contractual man; and ch. 3 the governance of contractual relations. In *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. Free Press, 1985.

[161] Oliver Williamson. Calculativeness, trust, and economic organization. *Journal of Law & Economics*, 36(1):453–486, 1993.

[162] Oliver E. Williamson. Transaction-cost economics: The governance of contractual relations. *The Journal of Law and Economics*, 22(2):233–262, 1979.

[163] Oliver E. Williamson. The theory of the firm as governance structure: From choice to contract. *The Journal of Economic Perspectives*, 16(3):171–195, 2002.

[164] Bingbing Xia, Huiyan Jiang, Huiling Liu, and Dehui Yi. A novel hepatocellular carcinoma image classification method based on voting ranking random forests. *Computational and mathematical methods in medicine*, 2016.

[165] Yu Xie and Charles F. Manski. The logit model and response-based samples. *Sociological Methods & Research*, 17:283–302, 1989.

[166] Yunming Ye, Qingyao Wu, Joshua Zhexue Huang, Michael K. Ng, and Xutao Li. Stratified sampling for feature subspace selection in random forests for high dimensional data. *Pattern Recognition*, 44(3):769–787, 2013.

[167] Bianca Zadrozny and Chrles Elkan. Transforming classifier scores into accurate multiclass probability estimates. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2002.

[168] Min Zhu, Jing Xia, Xiaoqing Jin, Molei Yan, Guolong Cai, Jing Yan, and Gangmin Ning. Class weights random forest algorithm for processing class imbalanced medical data. *IEEE Access 6*, pages 4641–46525, 2018.

[169] Hossein Zolfagharinia and Michael Haughton. The benefit off advanced load information for truckload carriers. *Transportation Research Part E: Logistics and Transportation Review*, 70(1):34–54, 2014.

[170] George Zsidisin, Douglas Voss, and Matt Schlosser. Shipper-carrier relationships and their effect on carrier performance. *Transportation Journal*, 46(2):5–18, 2007.