Long-Term Relationships in the US Truckload Freight Industry

Adam Harris and Thi Mai Anh Nguyen*

April 13, 2023

Click here for the latest version.

This paper provides evidence on the scope and incentive mechanisms of long-term relationships in the US truckload freight industry. In this setting, shippers and carriers engage in repeated interactions under contracts that fix prices but leave scope for inefficient opportunism. We show that shippers use the threat of relationship termination to deter carriers from short-term opportunism. Carriers respond to the resultant dynamic incentives, behaving more cooperatively when their potential future rents are higher. While shippers and carriers often interact on multiple lanes, we find evidence that shippers' incentive schemes do not take advantage of this multi-lane scope for certain classes of carriers.

The importance and ubiquity of informal interfirm relationships is widely recognized. As the economics, management, and sociology literatures have documented, where contracts do not exist or are incomplete, interfirm relationships are governed by nebulous notions of goodwill, trust, and reciprocity.[1] A wide range of theoretical work has elucidated various reasons why such informal arrangements might exist, what form they might take, and how they might be sustained.[2] A budding empirical literature studies these relationships. This paper contributes to that empirical literature by studying long-term relationships in the US truckload freight industry. The central role

*Harris: Department of Economics, Massachusetts Institute of Technology, 50 Memorial Dr, Bldg E52, Cambridge, MA 02142, USA (email: asharris@mit.edu). Nguyen: Department of Economics, Massachusetts Institute of Technology, 50 Memorial Dr, Bldg E52, Cambridge, MA 02142, USA (email: anhng@mit.edu). This research was made possible by the many people in the truckload freight industry who generously shared their time, insights, and data with us. We are especially grateful to Steve Raetz of CH Robinson—whose expertise we refer to throughout the paper—for both the TMC data set, as well as his invaluable feedback on our work. We are also very grateful to Angi Acocella, Chris Caplice, Glenn Ellison, Bob Gibbons, Tom Hubbard, Stephen Morris, Nancy Rose, Tobias Salz, and Mike Whinston for their advice and comments. We also thank all participants in the MIT Industrial Organization Lunch, the MIT Organizational Economics Lunch, the MIT Freight Lab, and the IIOC Rising Stars session for their comments and feedback. We acknowledge the support of the George and Obie Shultz Fund. This material is also based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1745302. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

[1]Early examples include Macaulay (1963).

[2]For an overview, see Malcomson (2010).

of informal long-term relationships, along with the existence of detailed microdata, makes this setting particularly well-suited for studying these relationships.[3] Exploiting this microdata, we ask (1) what is the mechanism governing these relationships, and (2) what is the scope—both temporal and spatial—of that mechanism.

We answer these questions using transaction-level data from a transportation management system (TMS) used by shippers to manage their relationships with carriers. The data records every interaction within these relationships, as well as the requests for proposals (RFPs) through which relationships are formed. Moreover, it provides explicit rankings of the carriers entered into the TMS by the shipper. These rankings summarize both the status of the relationship and the shipper's intended play. The observability of shipper's play and carriers' responses enables us to shed light on both the mechanism governing relationships and the scope of that mechanism.

We first use this data to establish two key facts about the interactions between shippers and carriers. First, the spot market creates a temptation for carrier deviation; however, carriers do not behave as opportunistically as we might expect if they were playing static best response. Second, shippers control relationship termination, a power that can potentially be used to punish carrier opportunism. This hypothesized punishment mechanism could explain carriers' apparent resistance to opportunistic behavior.

Next, we propose a parsimonious model consistent with these facts in which a punishment mechanism governs the shipper-carrier relationship. From the model, we derive testable predictions about qualitative features of the optimal incentive scheme and carriers' dynamic responses. While the model is intentionally minimal, it serves an important role: guiding our approach to econometric challenges and possible alternative mechanisms in our empirical analysis.

Building on the model, we return to our main questions, empirically testing and quantifying (1) the incentive mechanisms and (2) their temporal and spatial scope.

First, to assess the question of *temporal scope*, we examine carrier behavior in the contract's final weeks. We use mass RFP events as a plausibly exogenous source of relationship termination. We find evidence of endgame effects, with carriers reducing their tendency to accept loads by 10-18 percentage points after learning of the contract period's imminent end. We argue that these findings indicate that carriers are highly responsive to dynamic incentives before the contract's final weeks and that the incentive mechanism's temporal scope is limited to within rather than across contract periods.

Second, to assess the question of *spatial scope*, we quantify the degree to which relationship status on one lane is conditioned on the carrier's performance on the same lane versus on other

[3] About 80% of total industry volume is arranged through long-term shipper-carrier relationships, which, as we will describe in Section 1, are largely informal.

lanes.[4] As we would expect carriers' cost structure to differ based on their size and asset ownership, we naturally estimate shippers' incentive schemes separately for different carrier types: large asset-based carriers, small asset-based carriers, and brokers. This both provides a more nuanced answer to the question of scope and allows us to speak to the mechanisms at play. We find that large asset-based carriers face a harsh single-lane punishment scheme; a rejection on one lane can reduce the expected duration of the relationship on that lane by up to 52 loads. Brokers, in contrast, face a multi-lane punishment scheme.

Third, we quantify carriers' responses to the resultant dynamic incentives by estimating how carrier acceptance responds to own-lane volume, an exogenous proxy for the continuation value of the relationship.[5] The ordering of carrier types by their estimated responses to own-lane volume aligns with the strength of the estimated punishment schemes. Large asset-based carriers are most responsive to volume: doubling volume increases their acceptance probability by 10pp. As this would not be true for a carrier playing static best-response, we take this as strong additional evidence that, as suggested by our finding of endgame effects, carriers respond to dynamic incentives.

Finally, having presented empirical evidence consistent with a punishment mechanism, we discuss the potential role of learning. The fact that demotions tend to be permanent in our data refutes learning as the sole mechanism underlying the observed dynamics of shipper-carrier interactions. However, a combination of learning and network adjustments could explain patterns within relationships involving small asset-based carriers, who show less flexibility in adjusting their network of truck movements than do large asset-based carriers and brokers.

Our paper relates to the empirical literature on long-term informal relationships. While this literature is relatively recent, the last two decades have seen the development of a rich body of empirical evidence on the nature and value of these relationships. Different dynamic mechanisms have been explored, including dynamic enforcement (Brugues, 2023), reputation and learning (Macchiavello and Morjaria, 2015), and adaptation (Barron et al., 2020; Gil et al., 2021). The value and effects of these mechanisms can be quantified by exploiting exogenous variation in spot rates (Macchiavello and Morjaria, 2015) or changing prospects of future interactions (Gil and Marion, 2013). Our paper builds on both the conceptual and methodological insights of this literature to study shipper-carrier relationships in the US trucking industry. While most of our empirical evidence points toward dynamic enforcement, we also find suggestive evidence that, depending on carriers' size and asset ownership, other mechanisms may be at play.

Our paper also relates to the literature on the trucking industry. Early papers in this literature include Rose (1985, 1987), Hubbard (2001), Baker and Hubbard (2003, 2004), and Masten (2009). Since these papers, technological improvements have generated rich transaction-level data

[4] The term "lane" refers to an origin-destination pair.

[5] This use of volume as a proxy for the expected future relationship value is in keeping with Gil and Marion (2013).

on shipper-carrier relationships. Such data has been used in the transportation and logistics literature to analyze carriers' load acceptance, a key measure of performance in the trucking industry.[6] For example, Scott et al. (2017) find that carriers' tendency to accept offers within relationships is positively correlated with the volume and consistency of timing of these offers. Using the same data set as us, Acocella et al. (2020) find that, when shippers maintain high rates during a market downturn, carriers do not reciprocate with higher acceptance during a later market upturn. To the best of our knowledge, our paper is the first to dissect the dynamic mechanisms underlying shipper-carrier relationships.

Relative to these literatures, we make two main contributions. *First*, we study long-term informal relationships in an important yet understudied industry, the US truckload freight industry. Revenue in this industry was $700 billion in 2015, equivalent to about 4% of US GDP. Moreover, two aspects of this industry—(i) *fixed-rate contracts* and (ii) *on-path termination*—differentiate it from other settings in which relational contracts have been studied.[7] On the one hand, the lack of flexible monetary transfers prevents us from using workhorse models of long-term informal relationships (MacLeod and Malcomson, 1989; Baker et al., 2002; Levin, 2003), which rely on relational bonus to support optimal stationary contracts. On the other hand, the fact that relationship termination occurs on-path in our setting allows us to directly estimate the incentive contract, which is nonstationary. Our *second* contribution is that, by exploiting a unique data opportunity, we directly test widely held assumptions in the literature on long-term informal relationships. In particular, we show empirically that relationships do not necessarily exist at the firm-to-firm level as predicted by the multi-market contact literature (Bernheim and Whinston, 1990) and assumed in other empirical studies (Gil et al., 2021). To our knowledge, our paper is the first to test the assumption of relationship scope. Since the relationships studied previously do not feature on-path termination, it would not be possible to test whether incentive power is pooled across all relationship sub-parts (e.g. products or markets). The fact that we observe key aspects of the relationship—its status, the agent's performance, the agent's outside option, and the firm's termination strategy, at a sub-relationship level (in our case, the lane level) permits us to perform such a test.

The paper proceeds as follows. Section 1 describes the US for-hire truckload freight industry, its market institutions, and the structures and norms within which long-term shipper-carrier relationships operate. Section 2 describes our data. Section 3 establishes two key facts that motivate the model we present in Section 4—a repeated principal-agent game in which the shipper uses an incentive contract to deter the carrier's opportunism. Guided by the testable predictions of this model, Section 5 empirically quantifies the strength and scope of dynamic incentives within

[6]For an excellent review of this work, see Acocella and Caplice (2023).

[7]Rationalizing these unique features of the truckload freight setting is beyond the scope of our paper. We will instead take these features as given, allowing us to focus on other aspects of the incentive contracts.

shipper-carrier relationships for three different types of carriers. Section 6 discusses alternative and complementary mechanisms to that of our model. Section 7 offers conclusions about our findings, their implications, and future related research.

1 Setting

We begin by describing our setting: the US for-hire truckload freight industry. This is an economically important industry in which informal interfirm relationships play a central role. We describe the distinguishing features the industry, as well as the market institutions relevant to our analysis.

1.1 The US for-hire truckload freight industry

The freight trucking industry plays a uniquely important role in the US goods economy. In 2015, trucks carried 72% of domestic shipments by value.[8] US trucking firms had revenues of more than $700 billion in 2015, equivalent to nearly 4% of US GDP in 2015.[9]

Within the freight trucking industry, services are differentiated by the contractural relationships between shippers and carriers, by the size of shipments, and by the equipment required. In this paper, we focus on for-hire truckload carriers supplying dry-van services. We will explain each of these terms in turn: First, a *for-hire* carrier is one who sells his services to various different shippers. This is in contrast to a private-fleet carrier, who is vertically integrated with a single shipper. Second, a *truckload* carrier accepts only large shipments that fill all or nearly all of a trailer. Truckload service is "point-to-point": A truckload shipment has a single origin and a single destination. While a truckload carrier must plan his network of truck movements efficiently to minimize empty miles, his problem is far simpler than the optimization problem faced by a less-than-truckload carrier, who aggregates smaller shipments to fill the trailer. Finally, a freight truck consists of a tractor unit, which contains a heavy-duty towing engine and a driver cab, and a cargo trailer, which holds goods being hauled by the tractor. Some common trailer classes include refrigerated, flatbed and tanker. By far the most common trailer type is the *dry van*, used for hauling boxes or pallets of dry goods not requiring refrigeration. We will focus exclusively on dry van truckload services supplied by for-hire carriers. This is the largest subsegment of the trucking industry and one in which carriers' business model and logistical challenges are easy to understand.

[8] See Bureau of Transportation Statistics Freight Facts and Figures 2017.

[9] *American Trucking Trends*, American Trucking Association, 2015.

1.2 Truckload carriers: Assest-based carriers and brokers

Carriers providing truckload services can be divided into two types: asset-based carriers and brokers. An asset-based carrier owns and operates trucks, which he uses to transport goods.[10] A broker, on the other hand, has no trucks; instead, when a broker accepts a load from a shipper, he subcontracts in a spot arrangement with an asset-based carrier to transport the goods.

Among asset-based carriers, there is enormous heterogeneity in fleet size. On one extreme, the largest US carriers each operate more than 10,000 trucks. On the other extreme, as of December 2020, there were 317,791 registered carriers operating only a single truck.[11] While these tiniest owner-operator carrier will not feature in our analysis, heterogeneity in carrier size—and thus capacity, will play an important role. Our empirical analysis will distinguish between small asset-based carriers (100 trucks or fewer) and large asset-based carriers (more than 100 trucks).

1.3 Market institutions

In the US for-hire truckload freight market, shippers and carriers arrange loads through two primary market institutions: a spot market and (largely informal) long-term relationships.

Typically, about 20% of loads are arranged through the spot market.[12] The dominant spot market platform is organized by DAT Solutions; this online "load board" is a simple post-and-search marketplace that facilitates matches of shippers with loads and carriers with trucks.

The remaining 80% of US truckload transactions are arranged through long-term relationships between shippers and carriers. While these relationships are formalized by contracts, the contracts are highly incomplete. A contract defines liability for lost or damaged goods and establishes the rate the shipper will pay the carrier for each load on the lane. However, it imposes few other restrictions on the parties and does not obligate the shipper and carrier to behave cooperatively toward one another.[13] In particular, the contract does not obligate the carrier to accept any loads offered by the shipper under the terms of the contract. If the carrier rejects loads, the contract does not give the shipper any legal recourse.

The dominance of long-term relationships in this industry suggests that they offer benefits not enjoyed in spot arrangements. Such benefits could take several forms: First, shippers and carriers who interact repeatedly may benefit from a familiarity with each other's facilities and processes, which can improve functions like loading and payments. Second, arranging loads through

[10]To avoid confusion, we will, throughout the paper, refer to the shipper using she/her pronouns and the carrier using he/him pronouns.

[11]FMCSA, Motor Carrier Management Information System (MCMIS).

[12]`medium.com/@sambokher/segments-of-u-s-trucking-industry-d872b5fca913`

[13]The fact that these agreements are informal and not legally binding is widely understood in the industry. For instance, Melton Truck Lines Senior Vice President Dan Taylor wrote "The 'rate agreements' and 'load commitments' for the most part have no contractual obligation or penalties on either party." (See Taylor (2011).)

a long-term relationship might save on costs associated with searching and haggling in a thin spot market.[14] Such costs are likely non-negligible, as demand for transportation services is dispersed across space and time. Third, and closely related, because spot-market demand on a particular lane at a particular time might be scarce, carriers may prefer the more consistent demand from contracted shippers, which facilitates a stable, cost-effective network of truck movements for the carrier.

1.4 Managing relationships: The routing guide

A shipper frequently has contracts with several different carriers on a particular lane. These various carriers, are not, however, equal in status. The shipper explicitly ranks the carriers in a catalog called the *routing guide*. This ranking specifies the order in which carriers are sequentially offered each load that the shipper has on this lane.[15][16]

To illustrate this sequential offering process (sometimes called a *waterfall*), Table 1 gives an example of an (anonymized) routing guide for a shipper Z on the lane from City X to City Y. When Z has a load at City X that she wants to ship to City Y on a particular date, she first offers the load to the *primary carrier*, in this case, A. If A accepts, then A carries the load and receives $1230. If A rejects, then the load is offered to B. If B rejects, the load is offered to C, and so on. If the routing guide is exhausted without a carrier accepting, the shipper will typically turn to the spot market to try to find a carrier to accept the load.

The rationale for the shipper maintaining a routing guide with multiple carriers who have the right to reject loads, rather than a single carrier for whom acceptance is obligatory, is that 100% acceptance by a single carrier is unlikely to be efficient. The demand of a shipper over time is random and, therefore, cannot be perfectly predicted by a carrier.[17] This means that when the

[14] The role of transaction costs in driving the tendency towards contractual arrangements is a well established idea. For related studies in the context of the trucking industry, see Hubbard (2001) and Masten (2009).

[15] For a more detailed discussion of the routing guide and related features of truckload operations, see Section 4 of Caplice (2007).

[16] As we discuss in the next section, the process of sequentially offering loads is automated by software called a transportation management system (TMS). The TMS that allows each carrier only a short amount of time to respond to an offer. A typical response window might be fifteen minutes. This rapidity suggests that the shipper does not have a strategic incentive to rank a carrier higher just because that carrier is in high demand by other shippers; so little time passes between offers that a lower-ranked carrier is unlikely to be "snatched up" by another shipper while higher-ranked carriers are responding to their offers. This means that a shipper's static best response is to rank the carriers according to her preference over the carriers. Thus, rejections by top-ranked carriers are generally undesirable for the shipper. In Table 1, for instance, the fact that the shipper chose to rank A above B indicates that she prefers paying $1230 for service from A to paying $1327 for service from B. Furthermore, the fact that she ranked B above C, despite the fact that C has markedly lower contract rate, suggests that B provides the shipper with superior service in some dimension other than rate (e.g. quality, reliability). More generally, differences in non-rate characteristics may also be important to the shipper.

[17] While the timing of loads is random, the demand of a shipper is typically more consistent than the demand of a single consumer in some other transportation industries, e.g. the taxi or ride-hail industry.

Table 1: Example routing guide: Shipper Z, lane City X - City Y (on June 1, 2018)

Order	Carrier	Rate	Type
1	A	\$1230	Primary carrier
2	B	\$1327	Backup carriers
3	C	\$1095	
4	D	\$1450	

shipper offers a load on a particular date, the carrier's trucks may be poorly positioned for carrying this load; doing so may be very costly or infeasible.[18]

There are two features of the routing guide that play an important role in the dynamics of the shipper-carrier relationship:

First, the shipper has discretion to alter the ranking of carriers in the routing guide at any time. Indeed, though Table 1 gives the ordering of carriers on June 1, 2018, the routing guide for the same lane two weeks later is substantially different, with these four carriers ordered C, A, B, D. Why might such a change occur? The shipper might use the power to reorder the routing guide strategically to incentivize carrier cooperation. If a carrier (e.g., Carrier A) were behaving opportunistically, rejecting contract loads in favor of taking higher-paying loads in the spot market, the shipper could punish the carrier by downgrading him to a lower position in the routing guide. Being downgraded diminishes the carrier's future rents from the relationship, as he will now receive fewer offers on this lane. This possibility of punishment via reorganization of the routing guide is the mechanism at the heart of our model and empirical analysis.

Second, at the end of a contract period, the shipper holds a request for proposals (RFP) to determine the set of carriers, their rates, and their initial positions in the new routing guide. In an RFP, a shipper need not award the primary position to the lowest-bidding carrier; non-rate characteristics can be taken into account. While this is intuitively similar to a scoring auction, the way these RFPs are carried out in practice is far more complicated than the formal auctions that have been studied theoretically and empirically in a wide range of economic settings. After a shipper receives carriers' initial bids, multiple rounds of negotiation between the shipper and the various carriers jointly determine carriers' final routing guide positions and rates.

[18] One might think that the GPS-equipped devices now installed on most trucks would seem to offer an opportunity for contracts that obligate carriers to accept unless some verifiable conditions are met, e.g. unless all of the carrier's trucks are more than 100 miles from the load pickup location. However, any such simple criterion on current truck locations is unlikely to be satisfactory. First, a load is typically offered and accepted several days before the load's pickup time. At the time the acceptance decision is made, only current truck locations are verifiable, but what is relevant to the carrier's ability to pick up a load is of course the location of his trucks at the pickup time. Second, any such simple condition on truck locations would fail to account for both network-related issues and the carrier's obligations to other shippers. A carrier may have a truck near the load pickup location, but that truck may be needed to fulfill an obligation to another shipper.

2 Data

We use transaction-level data from the transportation management system software used by shippers to manage their relationships with carriers. The data records every interaction within these relationships. To proxy for carriers' outside option, we use a measure of the going rate for freight services in the spot market from DAT, the gold-standard provider of such spot market data.

2.1 Shipper-carrier microdata

Our analysis is made possible by the fact that shippers use a transportation management system (TMS) to manage their relationships with carriers and to automate the waterfall of tenders.[19] The shipper enters carriers' rates and ranks into the TMS, and then, for each load, prompts the TMS to sequentially send electronic offers to the carriers. For each load sent through the TMS, the software records the details of the load, all offers that are made, and whether each is accepted or rejected. These records for one particular TMS software provider, called TMC, are the source of our microdata.[20]

The microdata covers the period from September 2015 through August 2019. In all, the data set includes 1,074,172 loads and 2,130,125 tenders (i.e. offers). 71% of loads are accepted by the first carrier to which they are offered. All loads in the data set have a haul distance of at least 250 miles.[21] The mean distance is 692 miles with a standard deviation of 440 miles. The average per-mile contract rate is $1.85 with a standard deviation of $0.51.

Shipper-carrier relationships and networks The microdata includes 40 shippers with at least 500 loads. The median shipper has 8,094 loads with, on average, 192 active lanes and 53 active carriers each year.

On many lanes, shippers and carriers interact infrequently. For example, the median lane of the median shipper has only one load per month. However, among the top 10% of lanes for the median shipper, each has, on average, a load for every four days. Such variation in frequency of interactions will be important for our test of carriers' response to dynamic incentives.

Multilane interactions between a shipper and a carrier are also common. The top five and top ten carriers of the median shipper deliver, respectively, 58% and 73% of her loads. Relatedly, it is common for a carrier to serve as a shipper's primary carrier on multiple lanes. For example, the top five carriers of the median shipper hold primary status on an average of 21 lanes each. There is thus significant potential for strategic exploitation of multilane interactions: a shipper might condition

[19] In this industry, an offer of a load to a carrier is commonly referred to as a *tender*.

[20] TMC is a division of CH Robinson, a third-party logistics firm.

[21] For shorter-distance hauls, the prevailing market institutions are somewhat different. These loads are therefore excluded from our analysis.

a carrier's primary position on one lane on his behavior on another lane. This question of the scope of the incentive mechanism—whether shippers exploit multilane interactions to create cooperative incentives—is one of the key questions this paper addresses.

2.2 Spot rate data

We will use data on the average rate for truckload services in the spot market to capture the relevant outside option—the alternative opportunities available to shippers and carriers outside of their long-term relationships. This data comes from DAT Solutions, the leading provider of data on truckload spot markets. For our sample period, the data set gives us seven-day trailing average spot rates for a set of narrowly-defined lanes that cover the continental United States.[22]

Across all lanes and dates, the overall mean spot rate per mile is \$1.68 with a standard deviation of \$0.60. The first quartile, the median, and the third quartile are \$1.26, \$1.53, and \$1.93, respectively. A notable feature of the data is persistent differences in rates across lanes; a regression of spot rates on a set of lane fixed effects has an R^2 of 0.78, indicating that across-lane differences are large relative to within-lane variation. In later empirical analysis, we pool observations across lanes for the purpose of estimating the strategies of shippers and carriers. To make for appropriate comparisons across lanes, we will use residualized, rather than raw, spot rates, partialling out lane fixed effects. For the time series of average monthly spot rates over our sample period, see Figure 1 in the next section.

3 Two Key Facts

In this section, we use our shipper-carrier microdata, together with the data on spot rates, to establish two key facts which suggest that the shipper-carrier relationship can be thought of as a repeated principal-agent game in which an incentive mechanism deters carrier opportunism.

3.1 Fact 1: Temporary spot-contract rate differences create temptation for carriers

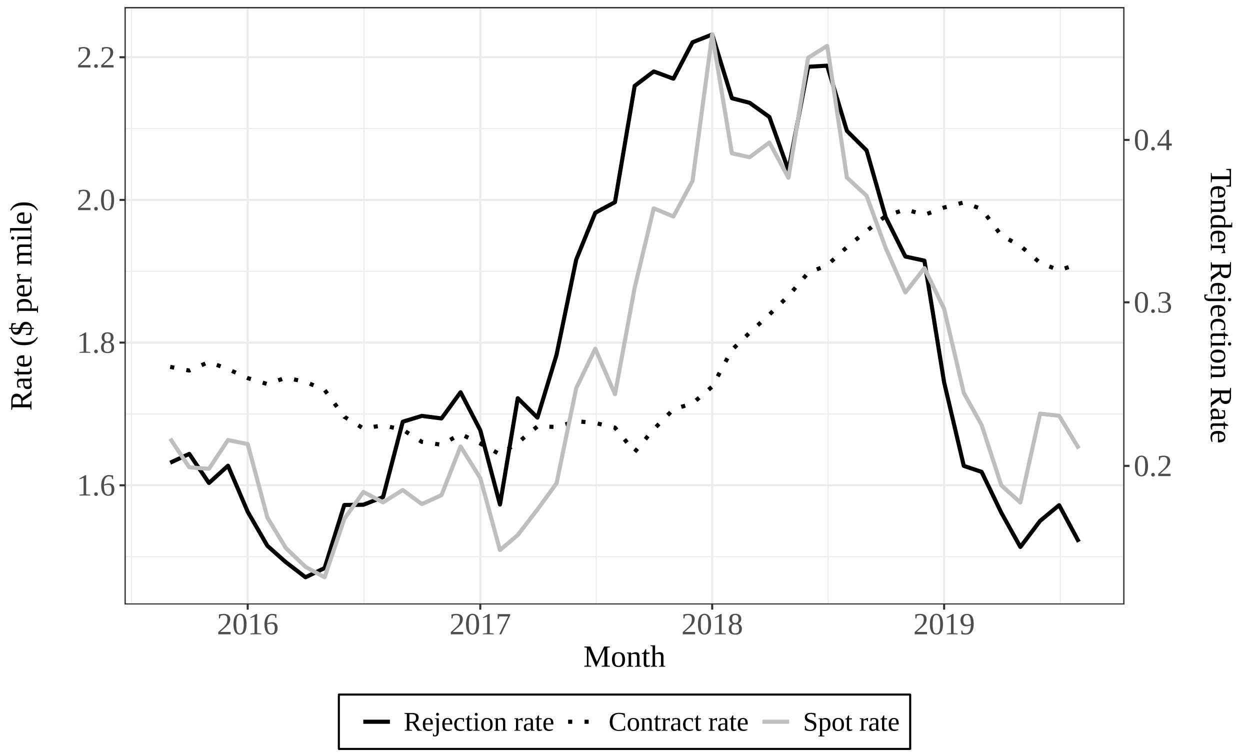
We begin by arguing that when spot rates are higher than contract rates, carriers are tempted by short-term opportunism. As such opportunism is detrimental to the shipper, a moral hazard problem exists within the shipper-carrier relationship.

Figure 1 illustrates the potential for carrier opportunism by depicting two key aggregate trends in our relationship microdata and aggregate data on spot rates. First, there are periods in which

[22]Each lane is defined by a pair of key market areas (or KMAs). The continental US is partitioned into 135 KMAs, so there are 135^2 KMA-to-KMA lanes.

spot rates are significantly higher than contract rates. Second, these periods coincide with a high proportion of rejections by carriers.

Figure 1: Aggregate trends: National averages of rejection, contract, and spot rates



Notes: The monthly rejection rate is constructed from the TMS microdata as the fraction of loads rejected by the first carrier in the routing guide. The average monthly contract rate is constructed from the TMS microdata. The average monthly spot rate is constructed from the DAT data, on the same set of lanes covered by the TMS microdata. The rejection rate, contract rate, and spot rate are all volume-weighted averages.

Figure 1 shows considerably greater intertemporal variation in spot rates than in contract rates. While spot rates were generally lower than contract rates in the first two years of our sample, an aggregate demand shock in late 2017 and early 2018 resulted in a sharp increase in spot rates.[23]

These spot market premia create the potential for short-term opportunism. Recall that a carrier in a long-term relationship always has the option to reject loads offered to him by the shipper. Thus, when the spot rate exceeds the contract rate, the carrier may reject contract loads and instead opt to provide service in the spot market. Figure 1 shows evidence consistent with this hypothesis: The period of high relative spot rates coincides with a large increase in the proportion of tenders rejected by primary carriers. This observation strongly suggests that the spot market represents a key outside option for carriers.

Such opportunism by the carrier presents a moral hazard problem. The fact that long-term

[23]Contemporary articles from various trade publications (including Transportation Topics and FreightWaves.com) describe the high spot rates of the 2017-2018 period as being driven by increased spending on e-commerce, booming US industrial production, and the December 2017 corporate tax cut. See, for instance, www.freightwaves.com/news/market-insight/forecasting-2019.

relationships exist in the first place—rather than all transactions being arranged through the spot market–suggests that there is relationship surplus that would be foregone were the carrier to opportunistically choose to service the spot market.[24] Furthermore, the shipper has imperfect monitoring: the shipper cannot distinguish between an inefficient opportunistic rejection and an efficient rejection resulting from the carrier's current cost of service being high.

Yet Figure 1 also gives us reason to believe that some mechanism exists to alleviate the moral hazard problem. When spot rates peak in January 2018, they are on average 20% higher than contract rates. Despite this strong incentive for carriers to reject loads, the majority of loads are still accepted by primary carriers in this month. That many carriers are willing to forgo significant short-term profits suggests that their opportunistic tendencies are restrained by some other force.

One such force could be an incentive scheme in which the promise of future rents helps alleviate the carrier's short-term opportunism. Necessary conditions for such an incentive scheme to be effective are that (i) the shipper has the power to deny the carrier future rents if he behaves opportunistically and (ii) the carrier's future rents from the relationship are sufficiently large. The next subsection establishes the former; the latter is established in Sections 5.1 and 5.3.

Before addressing the role of shippers in the next subsection, we note an asymmetry between shippers and carriers: While carriers might be tempted to short-term opportunism by high spot rates, shippers seem not to be tempted to opportunism by low spot rates. Harris and Nguyen (2023) show that shippers continue offering loads to their primary carriers even when lower-rate alternatives are available in the spot market.[25]

3.2 Fact 2: Shippers control relationship termination

We next use our shipper-carrier microdata to show that shippers control relationship termination and provide suggestive evidence on the form of shippers' termination strategies.

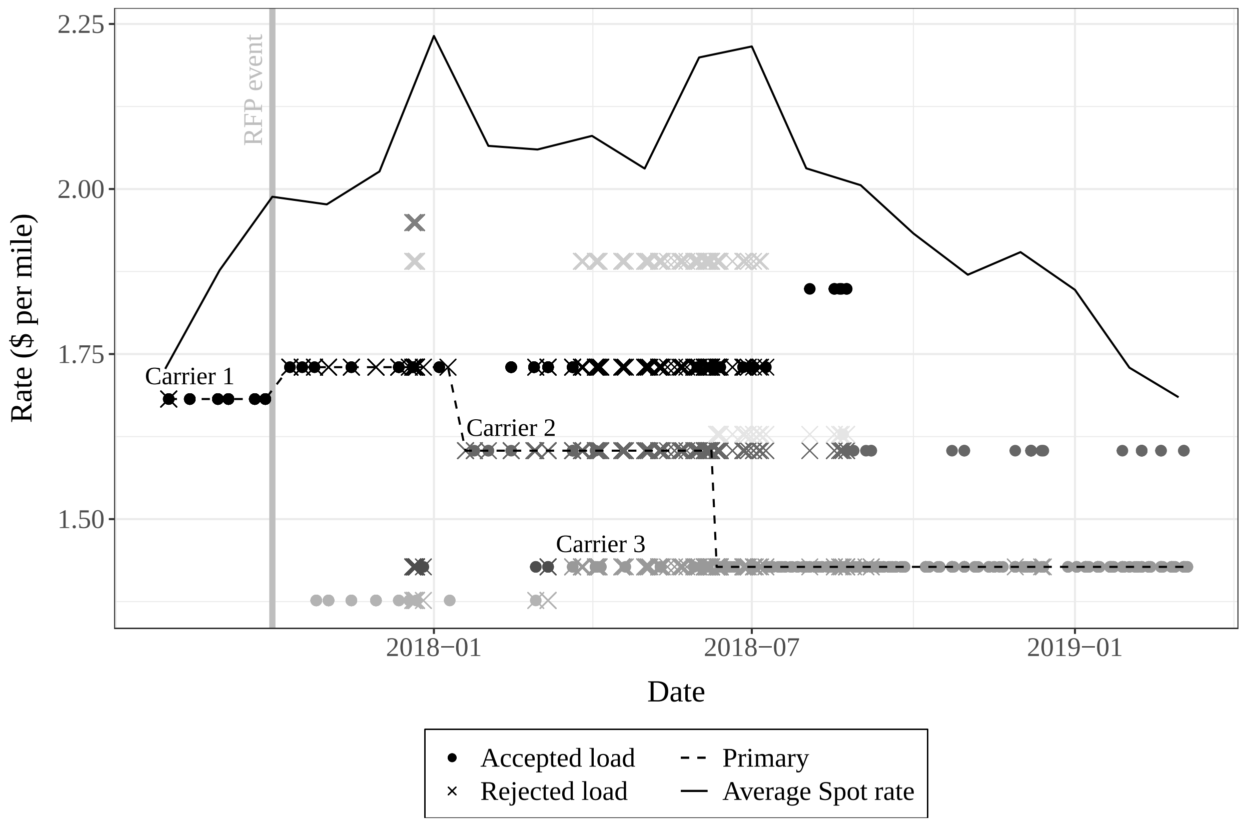
Figure 2 presents an example of a lane history that motivates the way we think about the shipper's decisions. Recall that the chief decision faced by the shipper is that of when and how to change the routing guide. Such changes can be made at any time. Some are the result of RFPs, while others take place within the contract period, i.e. in the time between RFPs. Our analysis will focus on the latter and, in particular, on those changes that replace one primary carrier with another. We refer to such a change as a *demotion* of the current primary carrier.

In the example in Figure 2, Carrier 1 initially holds primary status and accepts most of the

[24] See Section 1.3 for a discussion of possible sources of this surplus.

[25] In addition, industry experts tell us that it is "very rare" for a shipper to go directly to the spot market before the routing guide when spot rates are low. Perhaps because shipping represents only a small component of the operations of shippers, who are usually non-transportation firms (e.g. manufacturers or retailers), taking advantage of short-term opportunities to reduce shipping costs is not a priority. Shippers allow day-to-day shipping decisions to be automated by the TMS and make strategic decisions only on a medium-term basis.

Figure 2: An example: Tenders for Shipper X, City Y - City Z



Notes: Each point represents a tender: circles represent tenders that are accepted while crosses represent tenders that are rejected. Each carrier is indicated by a different color. The dotted black line indicates the rate of the primary carrier at each point in time. Carrier 1, Carrier 2 and Carrier 3 each serve as primary carrier for this shipper and lane for some subset of the period from October 2017 to May 2019. The tender data comes from the TMS microdata. The average monthly spot rate on the same lane is constructed from the DAT data set.

tenders that are offered to him. Around early October 2017, the shipper holds an RFP for this lane; Carrier 1 retains his primary status and gets a rate increase of about 5 cents per mile. However, over the next three months, a period of high spot rates, Carrier 1 rejects many of the loads offered to him. In January 2018, Carrier 1 is demoted from primary status and replaced by Carrier 2. Over the next five months, Carrier 2 rejects most of the loads offered to him. Ultimately, he too is demoted in favor of Carrier 3, who maintains primary status for the rest of the sample period.

This figure, which illustrates patterns that are common to many lanes, motivates two key conclusions about shipper-carrier relationships:

First, while shippers have almost unlimited discretion in what kind of routing guide changes they make, in practice, they do not switch primary carriers frequently; rather, a shipper maintains a primary carrier for a time before ultimately—and usually permanently—demoting that primary carrier.[26] From this observation, it seems appropriate to think of the shipper-carrier relationship in

[26]See Table 5. We find that in more than 95% of instances where a carrier is demoted from primary status, he never regains primary status on the lane in our sample period (2015-2019). While there is a truncation issue here (a demoted

terms of the following kind of principal-agent model: the shipper controls relationship termination and, at each point in time, decides between continuation and (permanent) termination.

Second, a clear pattern on this lane is that a series of rejections by the primary carrier often is followed by a demotion. This pattern is documented more systematically in Section 5.2. This evidence is consistent with an incentive mechanism where the shipper generates dynamic incentives for the carrier by conditioning relationship continuation on acceptance.

The evidence in this subsection indicates that shippers have the power to terminate relationships. Whether the threat of termination is effective in deterring carrier opportunism will be addressed in Sections 5.1 and 5.3, both of which show strong evidence of carriers responding to dynamic incentives.

4 A model of the incentive contract

In this section, we develop a model of long-term shipper-carrier relationships that will serve as a theoretical framework for our empirical analysis. For tractability, we focus on the relationship between a shipper and a primary carrier, abstracting away from the existence of backup carriers. Features of our model are motivated by the two facts established in Section 3:

Fact 1. Temporary spot-contract rate differences create deviating temptation for carriers,

Fact 2. Shippers control relationship termination and can use this power to generate an incentive scheme.

The model also generates testable predictions for our empirical analysis.

Relationship parameters A tuple $(\psi, \eta_1, \eta_2, p, \delta, F, G)$ summarizes the key characteristics of a relationship. Here, ψ is the relationship-specific gain to the shipper from transacting with the carrier; $\eta = \eta_1 + \eta_2$ is the relationship-specific gain to the carrier from transacting with the shipper. Some of these gains are publicly observed (η_1), such as the consistency of timing of shippers' requests, which helps the carrier's planning. Some are privately observed (ψ, η_2), such as the quality of on-road communication or efficiency of loading and docking. In addition, p is the contract rate; $\delta \in (0, 1)$ is the discount factor, measuring the frequency of shipments.[27] F is

carrier may regain primary status after the end of our sample period), it nevertheless seems clear that demotions are typically permanent. While it is possible that shippers are actually employing reward-punishment cycles like those described by Green and Porter (1984) or Li and Matouschek (2013), the cycles would have to be very long. We observe four years of data—quite a long period of time when one considers that the typical time between consecutive loads on a lane is on the order of a few days—, and yet we almost never observe a demoted carrier regaining primary status. Hereafter, we assume that demotion is permanent and use the terms demotion and termination interchangeably.

[27] An underlying assumption is that the shipper's demand for shipments is perfectly inelastic with respect to spot rates. See Harris and Nguyen (2023) for evidence that shippers do not reduce load offers within the routing guide when spot rates are low.

the distribution of the carrier's cost of servicing a shipment on the contracted lane; and G is the distribution of the spot rate on that lane.

Let $\tilde{p}_t$ and c_t denote, respectively, the spot rate and the carrier's cost draw in period t. The shipper's period-t payoff is $u_t = \psi - p$ if she is served by the contracted carrier and $u_t = -\tilde{p}_t$ if she is served by the spot market. The carrier's period-t payoff is $v_t = \eta + p - c_t$ when serving the contracted shipper and $v_t = \tilde{p}_t - c_t$ when serving the spot market (on the same lane). If the carrier chooses to remain idle (or serve a different lane), he gets zero payoff in that period. That is, c_t captures the opportunity cost of servicing the contracted lane in period t. This means that the cost distribution F captures both the contracted lane's average alignment with the rest of the carrier's network and the day-to-day variation of such alignment.

When $\psi + \eta > 0$, shipments fulfilled within relationships generate surplus over spot transactions. In this case, it is never efficient for the carrier to reject the shipper's offer in order to serve the spot market on the same lane. However, requiring the carrier to always accept the shipper's load is also not efficient, since the carrier's (opportunity) cost in some periods might be very high.[28] The inability of the shipper to distinguish between rejections due to high cost draws and rejections due to high spot rates represents a potential source of inefficiency in this setting, one that the shipper may hope to alleviate using the threat of relationship termination.

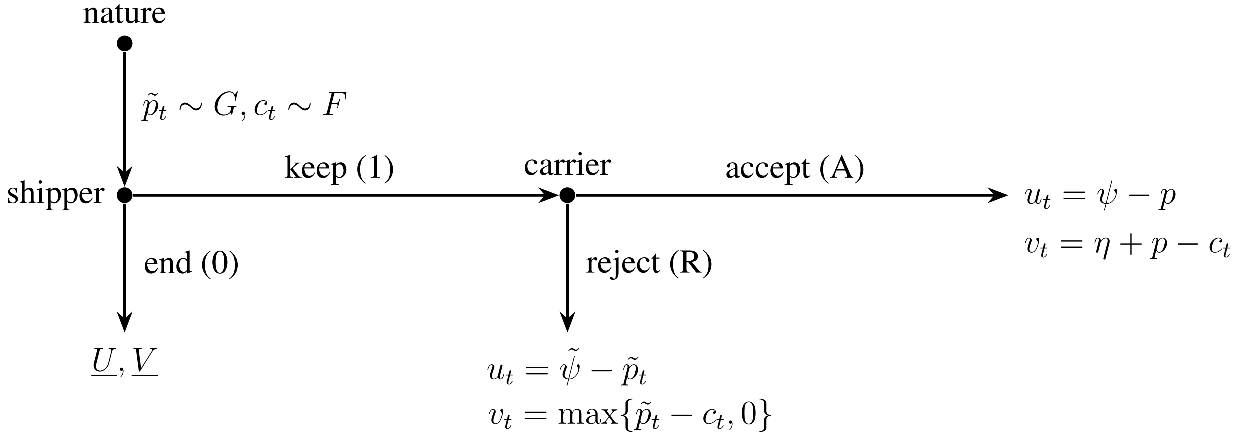
Timing and informational assumptions In period $t = 0$, the shipper holds an RFP to select a primary carrier. This RFP also reveals to both the shipper and carrier the characteristics $(\psi, \eta_1, \eta_2, p, \delta, F, G)$ of their relationships.[29] From period 1 onward, the shipper and the primary carrier interact repeatedly, with spot rates and costs being drawn independently and identically over time. The stage game is summarized in Figure 3. In each period t, a spot rate $\tilde{p}_t$ is drawn from G and publicly observed, and a cost draw c_t is drawn from F and privately observed by the carrier. The shipper then decides whether to "keep" the carrier as the primary carrier or "end" their relationship. If the relationship is maintained, the carrier chooses whether to accept (A) or reject (R) the shipper's load in that period. If he rejects, then he can either serve the spot market or remain idle. Otherwise, if the relationship ends, both sides resort to the spot market for future transactions; the shipper gets expected payoff of $\underline{U}$ and the carrier gets expected payoff of $\underline{V}$.

To generate clean predictions, we focus on the simplest class of shipper's incentive schemes, those that condition only on the carrier's decision in the last period. Denote such an incentive scheme by $\sigma_0 : \{A, R\} \to [0, 1]$, where $\sigma_0(d_{t-1})$ is the probability that the shipper maintains the relationship following decision d_{t-1} of the carrier in the last period. We interpret $\sigma_0(A)$ as the level

[28] Specifically, when $c_t > (\psi + \eta) + \tilde{p}_t$, the efficient outcome is that the carrier remains idle or serves a different lane.

[29] See Harris and Nguyen (2023) for modeling assumptions under which relationship characteristics are revealed via the auction.

Figure 3: The stage game



of rewards following cooperation and $1-\sigma_0(R)$ the level of punishment following noncooperation. We examine the carrier's optimal response to such scheme. Since the shipper conditions only on the last period's decision, we can focus on the carrier's stationary play, $\sigma_1 : \text{supp}(G) \times \text{supp}(F) \to [0,1]$, where $\sigma_1(\tilde{p}_t, c_t)$ is the probability that the carrier accepts the offered load given spot rate $\tilde{p}_t$ and cost draw c_t.

Model predictions Next, we derive testable predictions on the carriers' optimal stationary play and shipper's incentive scheme.

Proposition 1. *(Carrier acceptance) Suppose that $\eta + p \in$ Supp(G).*[30] *The carrier's optimal response takes a threshold form: accept if and only if $\overline{p} \geq \max\{\tilde{p}_t, c_t\}$, where*

$$\overline{p} = \eta + p + \frac{\delta}{1-\delta}(V(A) - V(R)),$$

and $V(d_t)$ is the carrier's expected payoff following $d_t \in \{A, R\}$. Moreover, the following comparative statics hold:

i) (Dynamic incentives) If $\sigma_0(A) > \sigma_0(R)$, then $\overline{p} > \eta + p$. That is, the carrier accepts more often than their static best response.

ii) For a fixed incentive scheme with $\sigma_0(A) > \sigma_0(R)$ and unobserved characteristics (ψ, η_2, F),

$$\frac{\partial \overline{p}}{\partial \delta} \geq 0, \quad \frac{\partial \overline{p}}{\partial \eta} \geq 1, \quad \textit{and} \quad \frac{\partial^2 \overline{p}}{\partial \delta \partial \eta} \geq 0.$$

[30] That is, there are periods in which spot transactions are more attractive than the contracted offer and periods in which the contracted offer is more attractive than spot transactions.

iii) For fixed carriers' parameters (η_1, η_2, F) and δ, and for every $\sigma_0(A) > \sigma_0(R)$,

$$\frac{\partial \overline{p}}{\partial \sigma_0(A)} \geq 0 \quad \textit{and} \frac{\partial \overline{p}}{\partial \sigma_0(R)} \leq 0.$$

Proof. See Appendix A.1. □

Intuitively, the shipper can use her control over relationship continuation as an incentive scheme to induce dynamic incentives for the carrier to accept more loads. The strength of such dynamic incentives depends on both the value of current and future loads to the carrier, as well as the levels of rewards and punishments induced by the shipper's incentive scheme.

Proposition 2. *(Shipper's single-lane incentive scheme) Suppose that shipper's match-specific gain is sufficiently large, $\psi \geq p - \mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq p]$. The optimal single-lane incentive scheme for the shipper satisfies*

i) (Maximum rewards) $\sigma_0^*(A) = 1$ for any parameter values of $(\psi, \eta_1, \eta_2, \delta, F, G)$,

ii) (Soft punishment) $\sigma_0^*(R) \in (0, 1)$ for some parameter values of $(\psi, \eta_1, \eta_2, p, \delta, F, G)$.

Proof. See Appendix A.2.1. □

The incentive scheme affects the shipper's payoffs via two channels: the effect on carrier's acceptance probability (the incentive-inducing effect) and the direct effect on the probability of ending the relationship (the regime-switching effect). Since the shipper faces no tradeoff between these two effects when deciding on the reward, the maximum reward is guaranteed relationship continuation. In contrast, harsher punishment increases the carrier's acceptance probability but also increases the likelihood of relationship termination when the carrier rejects. It is possible that this tradeoff is not resolved by extreme punishment, but rather by soft punishment.

While Proposition 2 assumes that the shipper's incentive scheme operates at the lane level, Example 1 considers the possibility of a broader scope, operating across multiple lanes within the shipper-carrier relationship.

Example 1. *Suppose that the shipper and the carrier interact on two lanes $\ell = 1, 2$, each characterized by $(\psi, \eta_1^\ell, \eta_2^\ell, p^\ell, F, G)$. Let $F \sim \alpha U(0, 1) + (1 - \alpha)\delta_K$ for some large K. That is, a cost draw is with probability α distributed as a standard uniform random variable and with probability $(1-\alpha)$ equal to some $K \gg 1$. Let $G \sim U(0, 1)$, $\alpha = 0.75$, $\delta = 0.9$, and $\psi = 0.3, p = 0.6$. Focus on multi-lane incentive schemes that map the average multi-lane rejections $\frac{1}{2}\sum_{\ell=1}^{2} \mathbf{1}\{d_{t-1} = R\}$ of the carrier in the last period to a probability of the shipper keeping the relationship (on both lanes). Denote by $\hat{\sigma}_0$ the shipper's optimal multi-lane incentive scheme and by $\sigma_0^\ell : \{A, R\} \to [0, 1]$ the shipper's optimal single-lane incentive scheme for $\ell = 1, 2$.*

i) If $\eta_1^1 + \eta_2^1 + p^1 = \eta_1^2 + \eta_2^2 + p^2 = 0.65$*, the shipper is better off using* $\hat{\sigma}_0$ *than* (σ_0^1, σ_0^2)*.*

ii) If $\eta_1^1 + \eta_2^1 + p^1 = 0.65$ *and* $\eta_1^2 + \eta_2^2 + p^2 = 0.8$*, the shipper is worse off using* $\hat{\sigma}_0$ *than* (σ_0^1, σ_0^2)*.*

Proof. See Appendix A.2.2. □

Intuitively, treating relationships at the shipper-carrier level rather than the shipper-carrier-lane level has two benefits. First is an incentive-pooling effect: multi-lane termination offers stronger dynamic incentives than single-lane termination. Second is an allocative-efficiency effect: if the relationship continuation conditions only on the average rejection, the carrier can selectively reject on lanes with high cost draws; this, in turn, increases the value of the relationship to the carrier and, thus, the carrier's incentive to cooperate. However, exploiting multi-lane interactions for dynamic incentives is not always straightforward. As shown in Example 1 (ii), using a simple scheme to pool incentives across heterogeneous lanes can backfire, making the shipper worse off than just using the optimal single-lane incentive schemes.

Empirical challenges In testing the predictions generated by this model, the econometrician only observes a component of carrier's gain (η_1), the contract rate (p), the frequency of interactions (δ), the realized spot rate ($\tilde{p}_t$) and the carrier's decision (d_t) in each period. Our empirical approaches focus on addressing endogeneity issues due to the unobservability of other relationship-specific characteristics (ψ, η_2) and carrier's cost distribution F. We will focus on addressing three channels: First, these unobservable characteristics affect the carrier's decisions (d_t). Second, they affect the contract rate (p) via the RFP process. Third, they potentially correlate with the frequency of interactions (δ). The last of these objects plays an important role in our empirical analysis, serving as a potential shifter for the continuation value of the relationship.

5 Empirical Evidence

In this section, we present empirical evidence on two key questions. First, what is the scope—both spatial and temporal—of the incentive mechanisms that govern shipper-carrier relationships? And, second, how do carriers respond to the dynamic incentives created by these mechanisms?

Our analysis of these questions proceeds in three subsections. In Section 5.1, we begin by presenting empirical evidence on the temporal scope of the relationship using carrier acceptance decisions in the weeks preceding the end of the contract period. This exercise also provides preliminary evidence that carriers respond to dynamic incentives. Next, in Section 5.2, we present empirical evidence on the scope of the relationship in the spatial dimension by estimating shippers' demotion strategies. Finally, in Section 5.3, we provide further evidence that carriers respond to dynamic incentives using carrier behavior throughout the relationship. In the latter two subsections, our analysis accounts for an additional dimension of carrier heterogeneity, analyzing behavior separately for three different types of carriers—large asset-based carriers, small asset-based carriers, and brokers.[31]

5.1 Empirical Evidence: Carrier behavior at the end of the contract period

This subsection presents evidence that (i) carriers respond to dynamic incentives and (ii) the scope of the incentive scheme carriers face is within, not across, contract periods.

We showed in Proposition 1(i) that a carrier facing dynamic incentives has $\bar{p} > p + \eta$, meaning that he is *ceteris paribus* more likely to accept load offers than a carrier with only static incentives. If an exogenous event were to unexpectedly eliminate the carrier's dynamic incentive, then we would expect a decrease in his probability of acceptance. In our data, *mass RFP events* act as a natural experiment which results in such a shock to dynamic incentives.

While we have established that a relationship sometimes ends because the shipper demotes the carrier, it may also end because the shipper holds a new RFP and selects a different primary carrier. Suppose an RFP is held and the primary carrier learns that he has "lost" the RFP, so he will soon lose his primary position. This alters his dynamic incentives. Typically, about four or five weeks pass between the announcement of the RFP outcome and the enactment of the new routing guide that results from that RFP. This means that the carrier experiences a one-month "lame duck" period in which he knows that, after the end of the month, there is no prospect of future relationship surplus. During this lame duck period, we might expect to observe *endgame effects*, where the carrier's tendency to accept loads is diminished. Observing such endgame effects for losing carriers would strongly support the notion that prior to the last few weeks of the contract period, $\bar{p} > p + \eta$; that is, carriers' future relationship surplus induces a cooperative response.

Even for a carrier who "wins" the RFP and will maintain the primary position in the next contract period, dynamic incentives may still be altered in the last few weeks of the current contract period. If a shipper's incentive scheme were conditioned only on carrier behavior *within a contract period*, then a winning carrier's dynamic incentives would be greatly lessened by imminent end of the contract period; he would, in effect, get a "free pass," knowing that the slate will be wiped clean at the start of the next contract period. Observing this kind of endgame effect for winning carriers would therefore not only provide further evidence of a response to dynamic incentives, but would also speak to the *scope of the incentive mechanism* in the time dimension.

While we might worry that the timing of an RFP is not exogenous, we address this concern by restricting our attention to *mass RFP events*, where the shipper holds RFPs simultaneously on

[31] See Section 1.2 for a discussion of these three carrier types.

at least 30 lanes. We think it is unlikely that poor carrier performance on one lane will affect the shipper's decision of when to hold an RFP on such a large set of lanes.[32]

To study these hypothesized endgame effects, we estimate a linear probability model

$$\begin{aligned}\text{Accepted}^{\ell}_{sct} =& \beta_0 + \beta_1 \left(\text{spot rate}^{\ell}_{t} - \text{contract rate}^{\ell}_{sct}\right) \\ &+ \sum_{k=1}^{18} \alpha_k \mathbb{1}\{k \text{ weeks until end of contract}\} + \epsilon^{\ell}_{sct} \end{aligned} \tag{1}$$

regressing an indicator for the primary carrier's acceptance of a tender on a set of dummies for the number of weeks until the end of the contract period (when new rates are enacted), along with the deviation profit (the difference between the spot and contract rates), which captures the carriers' short-run incentives.[33] The pattern of week fixed effects $\{\alpha_k\}$ over time will provide insight into the proposed end-of-contract effects. As we are interested in the potential endgame effects for both losing (lame duck) carriers and winning carriers, we estimate (1) separately for these two groups of carriers. The estimated coefficients $\{\hat{\alpha}_k\}$ on the weeks-to-end-of-contract dummies, along with 95% confidence intervals, are plotted in Figure 4.[34]

In the final month of the contract period, losing primary carriers (solid lines) significantly reduce tender acceptance, with carriers 17pp less likely to accept tenders in the last week (as compared with a baseline rate of 71%). These economically and statistically significant endgame effects provide strong evidence that—prior to learning the RFP outcome—carriers respond to dynamic incentives. The large magnitude of these endgame effects suggests that $\bar{p} \gg p + \eta$, which would results from either large relationship rents or harsh shipper rejection penalties.

The estimated coefficients for winning carriers (dashed lines) similarly show a decline (albeit a smaller one) in tender acceptance after the RFP outcomes are announced. This happens despite the fact that the winning carriers now know they will continue to be primary carriers. This is consistent with the winning carrier anticipating that the slate will be wiped clean at the start of the new contract period. Once his primary status in the next contract period is secured, the carrier's incentive to perform well in the current contract period is much weakened. This strongly suggests that the scope of the incentive mechanism is *within* rather than *across* contract periods.

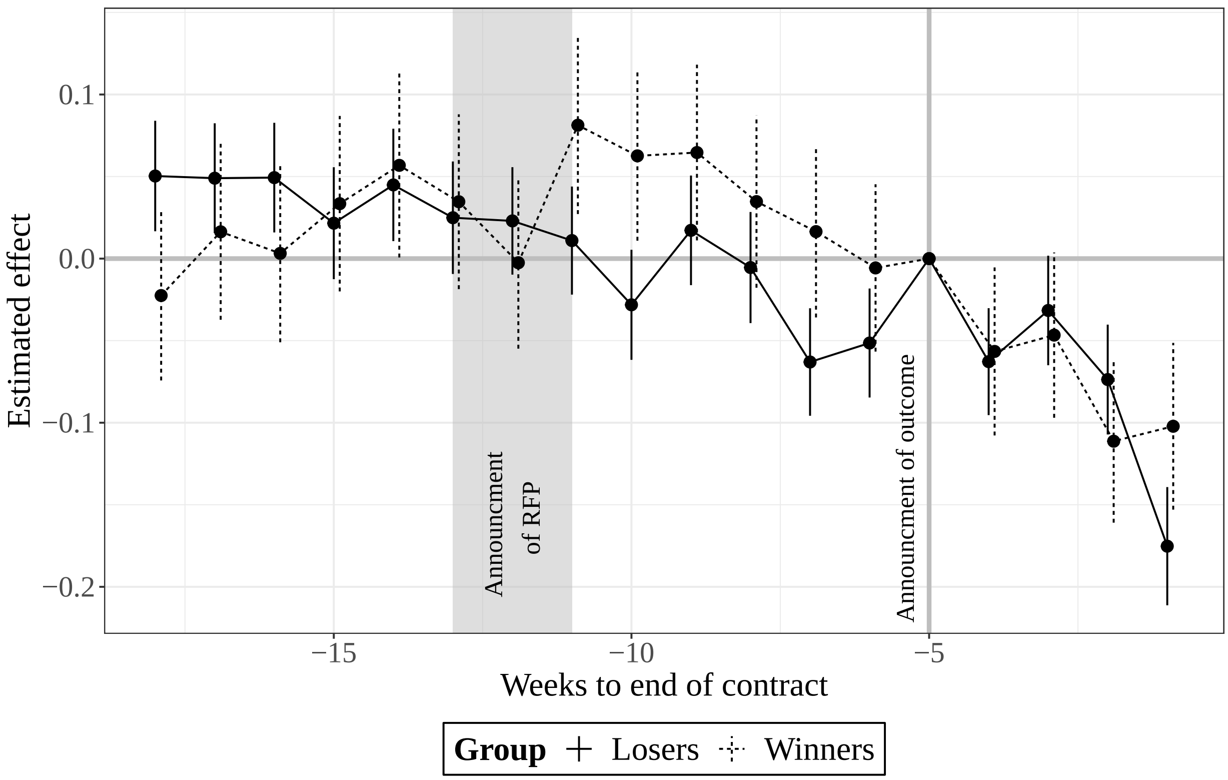
Another feature of Figure 6, however, does hint at the possibility of an across-contract period incentive mechanism. We see that, 9 weeks - 11 weeks before the end of the contract period, the carriers who go on to win the RFP are slightly more likely to accept tenders. It is common for 6-8 weeks to pass between the shipper informing carriers of an upcoming RFP and announcing

[32]This approach is intuitively similar to the "mass layoff" approach used to address worker selection issues in the labor literature.

[33]Notation: s indexes shippers, c indexes carriers, t indexes tenders, and ℓ indexes lanes.

[34]The omitted level is $k = 5$, i.e. we normalize $\hat{\alpha}_5 = 0$.

Figure 4: End-of-contract effects on tender acceptance for winning and losing carriers



Notes: This figure plots the estimated coefficients $\{\alpha_k\}$ from equation (1) with normalization $\alpha_5 = 0$. This choice reflects the fact that the RFP outcome is typically announced about 5 weeks before the end of the contract period. Also labeled in this plot is the approximate timing of the shipper announcing to carriers that an RFP will be held (6-8 weeks prior to the announcement of the RFP outcome).

the winner of the RFP. Thus, by 11 weeks before the end of the contract period, carriers are likely aware that an RFP is imminent. If a carrier believed the RFP outcome to be conditional on his acceptance decisions, this would create an extra dynamic incentive to accept, possibly resulting in a kind of *window-dressing effect*. Indeed, Table 7 (in Appendix C) shows that, conditional on the current primary carrier winning the RFP, current-period acceptances have a positive effect on his new contract rate; the effect on the probability of winning the RFP, however, is negligible.

From the evidence presented in this subsection, we draw two conclusions: First, the temporal scope of the incentive mechanism is within the contract period. Second, the evidence on endgame effects strongly supports the hypothesis that carriers respond to dynamic incentives. This second conclusion, however, comes with a caveat: this evidence on dynamic incentives is limited to a selected subset of relationships (those with mass RFP events) and to a selected time period (the last 18 weeks of the contract period). To further support our conclusion, Section 5.3 will present additional evidence of carriers responding to dynamic incentives based on acceptance decisions across and throughout all relationships in our sample.

5.2 Empirical Evidence: Shippers' Strategies

Having established the temporal scope of the incentive mechanisms, we now assess its spatial scope, as well as its magnitude. To do this, we study shippers' decisions to demote primary carriers to determine whether and how much such demotion decisions on lane ℓ are conditioned on the carrier's performance on all lanes within the shipper-carrier relationship or only on the carrier's performance on lane ℓ. Rather than assuming that this spatial scope is the same across all shipper-carrier relationships, we estimate the shipper's demotion strategy separately for large asset-based carriers (large ABCs), small asset-based carriers (small ABCs), and brokers.[35] While we find that the incentive mechanism for brokers is at the firm-to-firm level, we find that the incentive mechanism for large ABCs is at the narrower lane level. For small ABCs, we do not find evidence of punishment.

At a high level, our approach is a simple one: We estimate the following linear probability model of the shipper's demotion strategy:

$$\begin{aligned}\text{Demotion}_{sct}^{\ell} = \gamma_0 &+ \gamma_{\text{Rej}(\ell)}\text{Rejection rate}_{sct}^{\ell} + \gamma_{\text{Rej}(-\ell)}\text{Rejection rate}_{sct}^{-\ell} \\ &+ \gamma_{\text{Rej}(\ell)\times\text{Rej}(-\ell)}\text{Rejection rate}_{sct}^{\ell} \times \text{Rejection rate}_{sct}^{-\ell} \\ &+ \gamma_X X_{sct}^{\ell} + \gamma_{\text{Rej}(\ell)\times X}\text{Rejection rate}_{sct}^{\ell} \times X_{sct}^{\ell} + \epsilon_{sct}^{\ell},\end{aligned} \tag{2}$$

where $\text{Demotion}_{sct}^{\ell}$ is an indicator for primary carrier c being demoted from primary status on lane ℓ between load t and load $t+1$, and X_{sct}^{ℓ} is a vector that includes the time-invariant relationship characteristics, along with the spot rate at the time of load t.[36] The critical regressors are $\text{Rejection rate}_{sct}^{\ell}$ and $\text{Rejection rate}_{sct}^{-\ell}$, which capture, respectively, the frequency of rejections by carrier c on lane ℓ and on all other lanes on which c is the primary carrier of shipper s. The coefficients on these variables and their interaction speak to the spatial scope of the incentive mechanism.

Defining rejection rates In contrast to the parsimonious model in Section 4, our empirical analysis allows shippers to have memories longer than one load. We adopt a functional form that allows the shipper's strategy to condition on a rejection rate index that summarizes the entire history of rejections in the relationship, though potentially giving greater weight to more recent rejections

[35]To define carrier types, we use an NMFTA crosswalk to convert the Standard Carrier Alpha Code (SCAC) identifiers in our microdata to US DOT codes. We then match US DOT codes to carriers' DOT registration for the year 2020. This method matches 90% of carriers in our data set to a fleet size variable and a broker/non-broker indicator. The latter two groups are for carriers that have divisions operating both a brokerage and an asset-based business.

[36]Recall that by demotion, we mean a change to the lane-ℓ routing guide within a contract period that results in carrier c losing his primary position and being replaced by a new primary carrier. Since this definition is limited to changes within contract periods, any change in primary carrier that coincides with a change in rates (i.e. an RFP) on lane ℓ is not considered a demotion in our analysis.

than less recent ones. For a shipper s, lane ℓ, and carrier c, this index takes the following form:

$$\text{Rejection rate}^{\ell}_{sct} = \frac{\sum_{k=0}^{t-1} \alpha^{\text{days}(t-k,t)} \text{Rejection}^{\ell}_{sct-k}}{\sum_{k=0}^{t-1} \alpha^{\text{days}(t-k,t)}}, \tag{3}$$

where $\text{Rejection}^{\ell}_{sct}$ is an indicator for carrier c rejecting a load t from shipper s on lane ℓ; days $(t-k,t)$ indicates the number of days between load $t-k$ and load t; and $\alpha \in [0,1]$ is a daily decay rate.[37] The other-lanes rejection rate $\text{Rejection rate}^{-\ell}_{sct}$ is defined analogously using acceptance/rejection decisions by carrier c on lanes other lane lane ℓ on which c is the primary carrier for s.

Relationship characteristics In estimating the shipper's strategy, we control for relationship characteristics relevant to the payoffs of the shipper and/or carrier.

First and foremost, X^{ℓ}_{sct} includes the *log of average monthly volume*, which is a proxy for δ, the frequency of interactions between shipper s and carrier c and lane ℓ. We measure monthly volume as the number of loads tendered by the shipper on that lane in an active month.[38]

Second, X^{ℓ}_{sct} includes two measures of the *inconsistency of load timing*: The first measures the degree to which the number of weekly loads varies from week to week. The second measures the degree to which the distribution of loads across days of the week varies from week to week. We treat these measures of load consistency as a component of the carrier's match-specific gain (η_1).[39] Intuitively, if the timing of loads is more consistent, it is easier for the primary carrier to plan his network of truck movements around the expected timing of offers. As discussed in Section 1.2, in the context of the trucking industry, such network planning is important for reducing wasteful expenditures on fuel and labor.[40]

Finally, X^{ℓ}_{sct} includes the difference between the prevailing spot rate for lane ℓ at time t and carrier c's contract rate on lane ℓ. This captures the carrier's short-run incentive for deviation.

Identification strategy In estimating equation (2), we face a potential identification challenge stemming from the fact that the shipper's match-specific value (ψ) may shape the shipper's optimal strategy but is unobserved and therefore omitted. Several variables on the right-hand side of (2) are endogenous and are likely correlated with ψ. First, the rejection rate is the result of the carrier's endogenous response to the shipper's strategy. If the shipper's strategy is shaped by the omitted ψ, then the rejection rate will also be a function of ψ and therefore correlated with the regression error

[37]Note that the special case $\alpha \downarrow 0$ corresponds to the single-load memory restriction imposed in the model.

[38]By not conditioning on the identity of the accepting carrier, this measure avoids potential endogeneity issues. Moreover, since the primary carrier is the first carrier to receive an offer for each load, this measure approximates the expected number of offers the primary carrier receives per month during the relationship.

[39]See Appendix B for a detailed description of how the two inconsistency measures are constructed.

[40]This network-planning explanation for carriers valuing consistent timing of load offers is widely accepted in the truckload industry.

ϵ^ℓ_{sct}. Second, the contract rate, which enters X^ℓ_{sct}, is an endogenous outcome of the RFP process and is likely to be positively correlated with the shipper's match-specific value.

To address these endogeneity concerns, we use an instrumental variables approach that exploits exogenous variation in spot rates. First, we instrument for past acceptance/rejection decisions using the spot rates at the time at which each acceptance/rejection decision was made. To that end, we construct an index of past spot rates analogous to the construction of the rejection rate index:

$$\text{Past spot index}^\ell_{sct} = \frac{\sum_{k=0}^{t-1} \alpha^{\text{days}(t-k,t)} \tilde{p}^\ell_{t-k}}{\sum_{k=0}^{t-1} \alpha^{\text{days}(t-k,t)}}. \tag{4}$$

This index serves as an instrument for the rejection rate. We likewise construct Past spot index$^{-\ell}_{sct}$, an index of past spot rates on other lanes, which serves as an instrument for the other-lanes rejection rate. Second, we instrument for the contract rate using the spot rate at the time of the RFP in which the contract rate was established.[41] The idea is that at the RFP stage, the current spot rate serves as a competitive pressure on proposed contract rates.[42]

Using past spot rates as instruments is attractive because variation in spot rates is plausibly exogenous for our purposes. While endogenous factors at the market level do determine spot rates, the industry is competitive enough that no one shipper or carrier has significant power to influence them. To satisfy the exclusion restriction, however, past spot rates must not directly affect the shipper's strategy. An identifying assumption is therefore that only the period-t spot rate affects the shipper's period-t demotion decision, something which aligns with industry experts' descriptions of typical demotion decision processes. According to these experts, shippers track carrier performance and make demotion decisions using a scorecard that records various performance aspects, including rejection history, but does not record spot rate history.[43]

Estimation We jointly estimate the parameters (α, γ) by GMM. The parameters γ are identified by the standard 2SLS moments. To identify α, we include a set of additional instruments, the prevailing spot rate on lane ℓ during each of the last five weeks, which, under the exclusion restriction, should be conditionally independent of the outcome.[44] For computational efficiency, we implement this GMM estimation via a nested algorithm. For a given value of α, the inner step uses

[41] To be more precise, we use Spot rate$^\ell_{sct}$ − Spot rate$^\ell_{sc0}$ to instrument for Spot rate$^\ell_{sct}$ − Contract rate$^\ell_{sc}$.

[42] Figure 1 shows that the average contract rate tends to adjust with spot rates, though with some lag. This supports the idea that spot rates create competitive pressure on contract rates.

[43] Industry experts with whom we discussed these issues include Steve Raetz, Director of Research and Market Intelligence at CH Robinson; other members of Steve's team; and Chris Caplice and Angi Acocella of the MIT Center for Transportation and Logistics.

[44] These five lagged spot rates are instruments for the past acceptance/rejection decisions over the last five weeks. Under the functional form assumption in (3), however, individual rejection decisions enter into the shipper's strategy only through the rejection rate index. Thus, at the true α, these instruments for the individual acceptance/rejection decisions should be uncorrelated with the error term.

2SLS to obtain estimates of the linear parameters γ, while the outer loop searches for the value of α that minimizes the GMM objective function.

Table 2: Estimation of shipper's strategy

	All carriers		Large asset-based carriers		Small asset-based carriers		Brokers	
	(OLS)	(GMM)	(OLS)	(GMM)	(OLS)	(GMM)	(OLS)	(GMM)
Rejection rate	0.00842	0.00233	0.00823	0.084	0.012	0.000847	0.00642	-0.0151
	(0.000483)	(0.00467)	(0.000815)	(0.0337)	(0.00127)	(0.0123)	(0.00103)	(0.011)
Other lane rejection rate	-0.00547	-0.0166	-0.00241	-0.0408	0.000477	0.0364	0.00281	-0.0754
	(0.00086)	(0.0111)	(0.00136)	(0.032)	(0.00211)	(0.0463)	(0.00149)	(0.023)
Rejection rate $\times$ Other lanes rejection rate	0.0188	0.0743	0.0137	0.018	-0.00557	0.0102	0.00652	0.257
	(0.00142)	(0.0216)	(0.00218)	(0.0428)	(0.00357)	(0.0914)	(0.00264)	(0.0549)
Volume	-0.00689	-0.00684	-0.00438	-0.0137	-0.00415	-0.000986	-0.00805	-0.0106
	(0.000137)	(0.000576)	(0.000282)	(0.00218)	(0.000358)	(0.00224)	(0.000229)	(0.000589)
Inconsistency (loads / week)	0.0173	0.0341	0.0235	-0.0534	0.0122	0.00311	0.0154	0.0328
	(0.000679)	(0.00411)	(0.00176)	(0.0281)	(0.00194)	(0.00713)	(0.001)	(0.00274)
Inconsistency (day of week)	-0.00984	-0.0108	-0.00522	-0.0154	-0.00656	0.00406	-0.0122	-0.0207
	(0.000442)	(0.00134)	(0.000863)	(0.00485)	(0.0012)	(0.00364)	(0.000749)	(0.00177)
Spot rate - contract rate	-0.00224	-0.00947	-0.00346	-0.00264	-0.00362	-0.014	-0.00302	0.00958
	(0.000369)	(0.00199)	(0.00066)	(0.00587)	(0.000845)	(0.00859)	(0.000769)	(0.0066)
Rejection rate $\times$ Volume	-0.00759	-0.00747	-0.0105	0.00927	-0.00803	-0.0192	-0.0089	0.00788
	(0.000285)	(0.00163)	(0.000478)	(0.00451)	(0.000717)	(0.00637)	(0.000596)	(0.00223)
Rejection rate $\times$ Inconsistency (loads / week)	-0.0141	-0.0504	-0.0171	0.0746	-0.00792	0.00527	-0.0147	-0.0621
	(0.00109)	(0.00929)	(0.00219)	(0.037)	(0.00313)	(0.0109)	(0.00189)	(0.00875)
Rejection rate $\times$ Inconsistency (day of week)	-0.00709	0.00279	-0.0116	-0.0103	-0.0093	-0.0478	-0.0064	0.0257
	(0.00101)	(0.00464)	(0.00169)	(0.0129)	(0.00287)	(0.0128)	(0.00196)	(0.00809)
Rejection rate $\times$ (Spot rate - contract rate)	0.000754	0.0208	0.00195	-0.0287	0.00347	0.0252	0.00358	-0.0333
	(0.000769)	(0.00712)	(0.00112)	(0.0191)	(0.00199)	(0.0249)	(0.00187)	(0.0225)
N	680,229	680,229	173,787	173,787	67,508	67,508	250,197	250,197

Notes: Standard errors are in parentheses. For ease of interpretation, covariates that are interacted with rejection rate (other lanes rejection rate, volume, inconsistency, and spot rate minus contract rate) are normalized to have mean zero. The GMM specification jointly estimates α and the linear coefficients presented in this table. An outer loop searches over values of α, while, for a given α, an inner step estimates the linear coefficients by 2SLS and computes the 2SLS objective function. For each carrier type, the OLS specification takes the value of α estimated in the GMM specification as given and estimates the linear coefficients by OLS.

For the different carrier types, we estimate daily decay rates ranging from $\hat{\alpha} = 0.9742$ to $\hat{\alpha} = 0.9992$. This means that the shipper puts between 2.3% and 54.3% less weight on a rejection one month ago than on a rejection today. Our estimates of the linear parameters $\hat{\gamma}$ are reported in Table 2. For each carrier type, the parameter estimates for our main specification are the GMM estimates in the second column. For contrast and to illustrate the endogeneity problem described above, we also report in the first column the OLS estimates of the parameters γ.[45] We discuss the GMM estimates and their interpretation below.

Spatial scope and magnitude of the incentive mechanism The results in Table 2 show substantial heterogeneity in shipper strategies across carrier types. For brokers, we see that $\hat{\gamma}_{\text{Rej}(\ell)\times\text{Rej}(-\ell)}$

[45] These are the OLS estimates of γ with rejection rate measures constructed using the GMM estimate of the daily decay rate.

is positive and statistically significant, indicating that multi-lane punishment like that in Example 1 in the model section. In contrast, for large ABCs, we see that $\hat{\gamma}_{\text{Rej}(\ell)}$ is positive and signficant, while both $\hat{\gamma}_{\text{Rej}(-\ell)}$ and $\hat{\gamma}_{\text{Rej}(\ell)\times\text{Rej}(-\ell)}$ are insignificant, indicating that shippers use a single-lane incentive mechanism for this type of carrier. Example 1, which demonstrated the role of lane heterogeneity in combining incentives across lanes, suggests a possible explanation for this difference across carrier types. Since an ABC's costs and match-specific benefits on a lane depend in large part on the lane's alignment with the rest of the carrier's network, we would expect large ABCs to have significantly more heterogeneity across lanes than brokers.

For all carrier types, the estimated coefficients suggest that the shipper's punishment scheme is soft, rather than harsh. For large ABCs, for instance, our estimate $\hat{\gamma}_{\text{Rej}(\ell)}$ is positive and very statistically significant, indicating that shippers punish rejections with an increased probability of demotion. At first glance, however, this coefficient may appear very small, as it indicates that an increase in the rejection rate from 0% to 100% increases the probability of demotion between load t and load $t+1$ by only 8.4 percentage points. However, this coefficient should be interpreted in light of the fact that Rejection rate$^{\ell}_{sct}$ is a persistent state variable; since $\hat{\alpha} \gg 0$, a rejection of one load results in a sustained increase in the probability of demotion for many periods to come.

To get a sense of the economic significance of the estimated degree of punishment, we run a simple simulation to illustrate the effect of a rejection on the expected duration of a relationship.[46] The results show that if a large asset-based carrier rejects the first offer for the relationship, the expected relationship duration is 224 loads, as compared to 276 loads if he accepts the first offer. This 52 load effect difference is economically large. The mean per-load payment from shipper to carrier is \$1129, so a rejection early in the relationship may cost the carrier as much as \$60,788 in revenue. The prospect of such a loss is likely to create meaningful incentives for carrier cooperation.[47] Nevertheless, we conclude that punishment is soft, not harsh.

5.3 Empirical Evidence: Carriers' Acceptance

To conclude our empirical evidence, we use carrier behavior throughout the relationship to build upon the evidence presented in Section 5.1, bolstering the claim that carriers respond to dynamic

[46]Using the estimated shipper's strategy and the mean tender acceptance probability, we simulate relationships under two scenarios. In the first scenario, the carrier accepts the first load of the relationship; in the second, he rejects the first load of the relationship. Each load $t > 1$ is accepted with probability 0.71, the average primary carrier acceptance rate. Based on each acceptance/rejection, we update the rejection rate and compute the probability of demotion between each load t and $t+1$. For each scenario, we run 10 million simulations.

[47]For several reasons, this is likely an upper bound estimate of the effect of a rejection. First, we ran this simulation using the estimated demotion strategy for large ABCs, the carrier type facing the harshest punishment. Second, note that the form of the rejection rate means that the first acceptance/rejection decision has a long-term effect on the rejection rate than later acceptance/rejection decisions. Third, this simulation does not account for RFPs; instead, it assumes that a relationship continues indefinitely until the primary carrier is demoted. This explains the fact that the duration of a typical simulated relationship is far longer than that of the typical relationship in the data.

incentives and quantifying the magnitude of this response. To do this, we estimate the response of carriers' tendency to accept loads to relationship characteristics. We show that carriers respond strongly to lane volume, which would not be true of a carrier playing static best response.

As we did for shippers' strategies in the previous subsection, we estimate a linear probability model,

$$\begin{aligned}\text{Accepted}^{\ell}_{sct} = {} & \beta_0 \text{ controls}^{\ell}_{s} + \ \beta_{\text{volume}}\text{volume}^{\ell}_{s} \\ & + \beta_{\text{inc.}}\text{inconsistency}^{\ell}_{s} + \beta_{\text{volume}\times\text{inc.}}\text{volume}^{\ell}_{s} \times \text{inconsistency}^{\ell}_{s} \\ & + \ \beta_{\text{spot}}(\text{spot - contract})^{\ell}_{t} + \beta_{\text{volume}\times\text{spot}}\text{volume}^{\ell}_{s} \times (\text{spot - contract})^{\ell}_{t} + \epsilon^{\ell}_{sct}, \qquad (5)\end{aligned}$$

regressing an indicator for carrier c accepting load t from shipper s on lane ℓ on controls, along with a set of lane and relationship characteristics. Our choice of functional form—in particular, the inclusion of interactions between volume and other characteristics—reflects insights from Proposition 1.

Identification strategy In estimating (5), we again face two identification challenges stemming from the fact that a component of match-specific gain (η_2), as well as the carrier's cost distribution (F), is unobserved and thus omitted.

First, as in the previous subsection, we face the problem that the contract rate is an endogenous object likely correlated with the omitted variable.[48] We again address this issue by instrumenting for the contract rate using the spot rate at the time of the RFP in which the contract rate was established.

Second, while our use of average volume as a proxy for the frequency of future interactions (and thus for the discount factor δ) is in keeping with the empirical literature on relational contracting (e.g. Gil and Marion (2013)), we would face a potential identification challenge if this measure of volume were correlated with the unobserved component of carrier's match-specific gains, η_2. Such correlation might arise either because of a *selection effect* or an *investment effect*. The selection effect would result if carriers with better match-specific value or cost were systematically more likely to be primary carriers on higher-volume lanes.[49] The investment effect would result if the carrier could improve his match-specific value or cost by undertaking lane-specific investment; he would have greater incentive to undertake such investment on a higher-volume lane.[50]

[48]Since contract rates are established through an RFP process, a carrier with high match-specific value would tend to submit a lower bid, so we expect upward bias in the OLS estimates of β_{spot} and $\beta_{\text{volume}\times\text{spot}}$.

[49]There are several ways this selection could arise. First, carriers with high η could be more likely to submit bids for RFPs on high-volume lanes. Second, in choosing winners of RFPs, shippers might be more likely to select high-η carriers from among the bidders on high-volume lanes.

[50]Note that this investment effect story involves a violation of the timing assumption of the model, which imposes that (η_2, F) are fixed and known at time $t = 0$.

In either case, a positive correlation between volume and the unobserved match-specific value would induce bias in our key parameter of interest, β_{volume}. To address the potential bias resulting from the selection effect, we include two sets of fixed effects that absorb variation in the carrier's unobserved match-specific gain and cost: the first are shipper-carrier fixed effects; the second are carrier-origin-destination (or carrier-origin-destination-year) fixed effects, where origin and destination are defined as Census regions. By including the shipper-carrier fixed effects, we absorb a variety of potential shipper-carrier specific components of the unobserved value, including, for instance, relationship-specific knowledge and integration of payment or communication systems.[51] By including the carrier-origin-destination(-year) fixed effects, we absorb key geographic components of the carrier's match-specific value, namely the compatibility of a particular (broadly defined) route with the rest of the carrier's network.

For each carrier type, the first column in Table 3 reports the OLS estimates of Equation (5) while the second column reports the IV estimates. The fourth column reports estimates for the main specification, IV with shipper-carrier and carrier-origin-destination-year fixed effects included. For comparison, the third column reports IV estimates for a specification like the main specification, but with carrier-origin-destination, rather than carrier-origin-destination-year fixed effects.

Response to dynamic incentives The results for the main specification indicate that β_{volume} is positive and signficant for all carrier types, which would not be true of a carrier playing static best responses.[52] We interpret this result as strong evidence that carriers respond to dynamic incentives. We also observe striking differences in the magnitude of this response across carrier types: large ABCs exhibit a very strong dynamic response (doubling volume increases acceptance probability by 10pp), whereas brokers (2.3pp) and small ABCs (1.9pp) show weaker responses. Notably, this ordering of carrier types by responsiveness to dynamic incentives aligns precisely with the relative strength of same-lane dynamic incentives created by the incentive schemes each type faces.[53]

As described above, the appropriateness of our average volume measure as a proxy for the discount factor δ is potentially threatened by the selection and investment effects, which, if present, would result in an upward bias in $\hat{\beta}_{\text{volume}}$. However, comparing the IV estimates to the IV-FE1 and IV-FE2 specifications, we see that, for ABCs, the inclusion of fixed effects actually *increases* $\hat{\beta}_{\text{volume}}$. This speaks against the hypothesized selection story; however, the fact that the inclusion

[51]Note, however, that when these fixed effects are included, we estimate the parameters by exploiting variation in characteristics *across* different lanes of the same shipper-carrier pair. This approach would not produce meaningful estimates of the effects of these characteristics if shippers employed multilane punishment strategies; if that were the case, a carrier's acceptance/rejection decisions on lane ℓ might respond to characteristics of other lanes. However, our analysis indicates that this is not a concern for ABCs.

[52]For such a carrier, acceptances would depend only on the current spot rate and threshold $\underline{p} = \eta_1 + \eta_2 + p$.

[53]Recall that results in the previous subsection show that large ABCs face a relatively harsh own-lane punishment scheme, while brokers face a multi-lane scheme and small ABCs face a negligible degree of punishment.

Table 3: Estimation of carriers' acceptance

	All carriers				Large asset-based carriers			
	(OLS)	(IV)	(IV-FE1)	(IV-FE2)	(OLS)	(IV)	(IV-FE1)	(IV-FE2)
Volume	0.0112	0.00558	0.0195	0.0130	-0.000580	0.0418	0.138	0.143
	(0.000539)	(0.000666)	(0.000970)	(0.00133)	(0.00134)	(0.00441)	(0.0335)	(0.0379)
Spot rate - contract rate	-0.234	-0.292	-0.251	-0.182	-0.150	-0.127	0.0320	-0.0205
	(0.000991)	(0.00405)	(0.00365)	(0.00337)	(0.00223)	(0.00875)	(0.0426)	(0.0234)
Inconsistency (loads / week)	-0.0868	-0.0833	-0.0632	-0.0671	-0.251	-0.260	-0.224	-0.168
	(0.00211)	(0.00226)	(0.00230)	(0.00241)	(0.00655)	(0.0102)	(0.0262)	(0.0186)
Inconsistency (day of week)	-0.0256	-0.0210	-0.0220	-0.0170	-0.0400	0.0259	0.0938	0.0847
	(0.00135)	(0.00147)	(0.00173)	(0.00164)	(0.00327)	(0.00787)	(0.0316)	(0.0304)
Volume × (Spot rate - contract rate)	0.0412	0.327	0.304	0.336	0.0414	-1.004	-1.683	-1.650
	(0.000815)	(0.0168)	(0.0217)	(0.0275)	(0.00211)	(0.0945)	(0.440)	(0.470)
Volume × Inconsistency (loads / week)	-0.0391	-0.0307	-0.0232	-0.0254	-0.103	-0.178	-0.142	-0.133
	(0.00131)	(0.00149)	(0.00139)	(0.00143)	(0.00402)	(0.00923)	(0.0245)	(0.0261)
Volume × Inconsistency (day of week)	-0.0553	-0.0376	-0.0377	-0.0373	-0.0792	-0.0915	0.103	0.0484
	(0.00143)	(0.00185)	(0.00184)	(0.00179)	(0.00351)	(0.00556)	(0.0319)	(0.0176)
< 7 days since promotion	-0.0786	-0.0888	-0.0324	-0.0302	-0.136	-0.0960	0.0164	0.0308
	(0.00168)	(0.00189)	(0.00162)	(0.00160)	(0.00428)	(0.00757)	(0.0176)	(0.0208)
Fixed effects								
Shipper × carrier			X	X			X	X
Carrier × region × region			X				X	
Carrier × region × region × year				X				X
N	868014	868014	867741	867152	176569	176569	176524	176438

	Small asset-based carriers				Brokers			
	(OLS)	(IV)	(IV-FE1)	(IV-FE2)	(OLS)	(IV)	(IV-FE1)	(IV-FE2)
Volume	-0.0276	-0.0246	0.0152	0.0269	0.0258	0.0368	0.0368	0.0335
	(0.00181)	(0.00184)	(0.00272)	(0.00286)	(0.000835)	(0.00147)	(0.00121)	(0.00115)
Spot rate - contract rate	-0.323	-0.154	-0.140	-0.0979	-0.359	-0.417	-0.360	-0.271
	(0.00361)	(0.00884)	(0.00783)	(0.00847)	(0.00179)	(0.00591)	(0.00355)	(0.00377)
Inconsistency (loads / week)	-0.164	-0.129	-0.132	-0.0862	-0.0352	-0.0321	-0.0317	-0.0322
	(0.0104)	(0.0113)	(0.0110)	(0.0116)	(0.00250)	(0.00285)	(0.00341)	(0.00340)
Inconsistency (day of week)	-0.0812	-0.0551	-0.0104	0.00586	0.00400	-0.0536	-0.0245	-0.0121
	(0.00562)	(0.00662)	(0.00914)	(0.0103)	(0.00201)	(0.00607)	(0.00377)	(0.00324)
Volume × (Spot rate - contract rate)	-0.0560	-0.0680	0.0522	0.0257	0.0792	0.520	0.342	0.238
	(0.00263)	(0.0126)	(0.0124)	(0.0153)	(0.00155)	(0.0429)	(0.0306)	(0.0289)
Volume × Inconsistency (loads / week)	-0.0507	-0.0237	-0.0592	-0.0426	-0.0201	-0.0311	-0.0112	-0.00972
	(0.00514)	(0.00565)	(0.00546)	(0.00576)	(0.00177)	(0.00227)	(0.00225)	(0.00219)
Volume × Inconsistency (day of week)	-0.0695	-0.0362	-0.0263	-0.0125	-0.0227	-0.0476	-0.0310	-0.0227
	(0.00531)	(0.00622)	(0.00752)	(0.00801)	(0.00218)	(0.00348)	(0.00294)	(0.00260)
< 7 days since promotion	-0.117	-0.111	-0.0452	-0.0465	-0.0302	-0.0391	-0.00953	-0.00793
	(0.00638)	(0.00650)	(0.00539)	(0.00541)	(0.00240)	(0.00284)	(0.00238)	(0.00222)
Fixed effects								
Shipper × carrier			X	X			X	X
Carrier × region × region			X				X	
Carrier × region × region × year				X				X
N	69240	69240	69212	69127	289599	289599	289509	289300

Notes: Standard errors in parentheses. Controls include distance and distance squared. Standard errors are in parentheses. For ease of interpretation, the covariates that are interacted with volume (inconsistency, contract rate, and spot rate) are normalized to have mean zero. Specifications IV-FE1 and IV-FE2 include shipper × carrier fixed effects; specification IV-FE1 includes carrier × origin region × destination region fixed effects; and specification IV-FE2 includes carrier × origin region × destination region × year fixed effects. For the latter two sets of fixed effects, regions are determined using US Census regions of loads' origin and destination locations. The small difference in sample size across the specifications with and without fixed effects reflects the fact that singleton observations are dropped in the specifications with fixed effects.

of fixed effects changes the estimates for ABCs so substantially suggests that there is substantial heterogeneity in costs (F) and match-specific values (η_2) which is absorbed by these fixed effects. For brokers, in contrast, the inclusion of fixed effects changes the estimates very little; this suggests that brokers are much less heterogeneous in costs and match-specific values than ABCs, which accords with the fact that brokers do not need to undertake network planning.

What of the hypothesized investment effect? While the inclusion of the fixed effects may not adequately address this issue, estimates in Table 3 do speak to the role of specific investment for each carrier type. Included in the regression is an indicator for whether a load t occurs within 7 days after carrier c's promotion to primary status on lane ℓ. Suppose that, after winning the primary position, the carrier decides to undertake an investment to improve η_2 or F. If this investment takes some time (as would be expected for the carrier to, for instance, change his network to better align with lane ℓ), then we would expect the carrier's tendency to accept to increase after the investment is fully realized. For small ABCs, this is indeed what we see: A small ABC is 4.7pp less likely to accept a load in the first week after promotion. This suggests that investment does play a role in the shipper-carrier relationship, at least for small ABCs.

Responses to inconsistency and spot rates In addition to shedding light on responses to dynamic incentives, the estimates in Table 3 also provide insight into two other key aspects of carrier behavior and how these differ by carrier type.

First, we see that inconsistency (a component of η_1) does indeed affect carriers' acceptance decisions. In particular, lanes with more inconsistency in the number of loads per week have lower acceptance probabilities, with this effect being stronger on higher-volume lanes (matching the predictions of Proposition 1). The measure of inconsistency in the timing of loads within the week is seemingly of lesser importance.

Second, responses to variation in spot rates differ substantially across carrier types, with brokers being far more sensitive than ABCs. This is intuitive, as the profit margins of brokers, who subcontract loads to the spot market, are much more closely tied to spot rates than those of ABCs, who transport loads using their own physical assets.

6 Discussion

In this section, we discuss alternative mechanisms of the interactions between shippers and carriers. These mechanisms will complement the punishment mechanism modeled in Section 4 to explain the richness of the empirical findings in Section 5.

Table 4 provides a summary of our key empirical findings.

On the one hand, the punishment mechanism can explain most of these findings. First, brokers

Table 4: A recap of empirical findings

Evidence	own-lane	multi-lane	response to volume	heterogeneity	network adjustment
Brokers	-	✓	moderate	-	-
Large asset-based	✓	-	strong	✓	-
Small asset-based	-	-	mild	✓	✓

face multi-lane punishment and thus respond moderately to own-lane volume. Second, large ABCs face single-lane punishment and thus respond strongly to own-lane volume. Third, small ABCs face no punishment and thus have the weakest response to volume. Finally, the differences in the spatial scope of relationships for brokers and asset-based carriers could be explained by the fact that brokers face less heterogeneity in gains and costs across lanes, making effective incentive pooling easier to achieve for brokers.

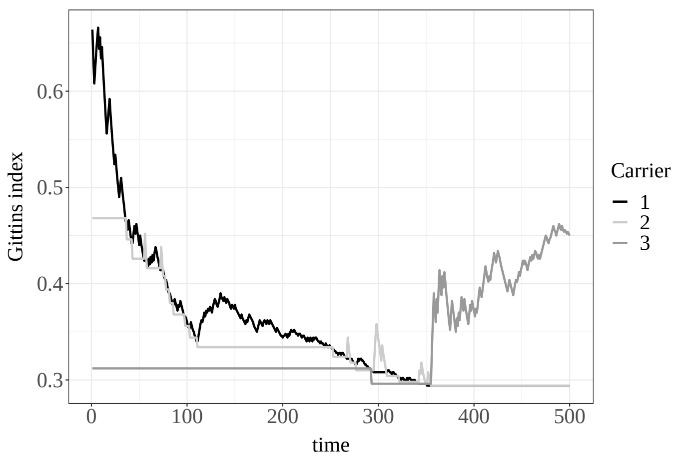
On the other hand, the heterogeneity across relationships of asset-based carriers could give rise to another dynamic mechanism—learning. We argue that while learning cannot be the only mechanism, it serves as a complementary mechanism in explaining our evidence on asset-based carriers, particularly small ones.

A learning model for asset-based carriers Suppose carriers have unobserved characteristics, such as idiosyncratic gains or costs from the relationship. In this case, past rejections would be indicative of future rejections, thereby affecting the shipper's expected value from maintaining the carrier's primary status. Notice that being ranked first on the routing guide ensures the primary carrier greater consistency in the timing of offered loads, facilitating network planning and load fulfillment. Thus, the opportunity cost of maintaining a primary carrier who is likely to reject offers is the higher acceptance probability that the first backup carrier *would have* were he to be primary. This is one reason why the shipper might prefer a primary carrier with higher acceptance probability. From the perspective of a primary carrier, learning by the shipper that conditions on past rejections would also create dynamic incentives for the carrier to accept more loads.

Assumption 1. *Each carrier has a permanent tendency to accept loads and the shipper holds independent priors over carriers' acceptance tendencies.*

Learning cannot be the only mechanism However, we observe that demotions are generally permanent, a fact which rules out learning as the only mechanism. Under Assumption 1, the potential learning mechanism described above can be simplified to a bandit problem with independent arms, a common model in the literature on learning. By choosing a carrier to be the primary carrier and receive the first offer in each period, the shipper gradually learns the carrier's tendency to accept loads as a primary carrier. The solution of the shipper's dynamic optimization problem is as follows: each period, she chooses as primary the carrier with the highest Gittins index, which captures both the exploitation and exploration value of choosing a carrier over the outside option.[54] Given that the tendency to accept loads should be independent across carriers once conditioned on observed characteristics, a carrier's Gittins index evolves only when he is chosen as primary.

Figure 5: A simulated learning path



Notes for Figure 5: In this example, the shipper's prior is overly optimistic about Carriers 1 and 2, and overly pessimistic about Carrier 3. Thus, the Gittins indices of the former two carriers are generally decreasing, both because of the initial overoptimism and because the decrease in informational values as they are chosen. The shipper makes many switches between these two carriers before she starts to experiment with Carrier 3, at which time the Gittins index of Carrier 3 evolves.

Table 5: Probability of repromotion

	Returns to primary	*Never again primary*	*Return probability*
Panel A: All carriers			
Full sample (2015-2019)	405	8701	4.45%
First half (2015-2017)	246	4044	5.73%
Panel B: Large asset-based carriers			
Full sample (2015-2019)	106	1378	7.14%
First half (2015-2017)	65	655	9.03%
Panel C: Small asset-based carriers			
Full sample (2015-2019)	38	647	5.55%
First half (2015-2017)	18	180	9.09%
Panel D: Brokers			
Full sample (2015-2019)	116	4125	2.74%
First half (2015-2017)	61	2035	2.91%

Notes for Table 5: This table presents evidence on the frequency of a carrier who is demoted from the primary position on a lane ℓ ever returning to the primary position on lane ℓ. The first column list the number of times in the given subsample that we observe such a return. The second column lists the number of instances where the demoted carrier never again regains the primary position on the lane. One concern with this exercise is that—since our sample is finite—we might not observe a carrier regaining the primary position because this occurs after the end of our sample period. To ameliorate this concern, the second row restricts the sample to carriers that are demoted in the first half of the sample period. In both rows, we also restrict the sample to the set of lanes that only ever have a single primary carrier at any point in time.

Figure 5 provides an example illustrating the evolution of Gittins indices in a learning problem with three carriers. Initially, Carrier 1 is primary. The shipper continues to choose Carrier 1 until her belief about this carrier's tendency to accept loads drops just below that of Carrier 2, at which point she switches to the latter carrier. While the shipper chooses Carrier 2, the Gittins index of Carrier 1 remains the same and well above that of Carrier 3. Thus, the next time the shipper needs to make a switch, she switches back to Carrier 1 rather than switching to Carrier 3. This intuition

[54]The exploitation value of an option refers to the expected payoff of that option given the current beliefs. The exploration value refers to the informational value of an additional observation of that option. See Whittle (1980) and Weber et al. (1992) for details.

generalizes: when learning steps are small, we expect to see the shippers switching from Carrier 1 to Carrier 2 and then back to Carrier 1 ("switch-back pattern") much more often than we see her switching from Carrier 1 to Carrier 2 and then to Carrier 3.

While the learning story predicts the prevalence of the switch-back pattern, a carrier returning to the primary position on a lane after being demoted ("repromotion") is rare in our TMS microdata. Table 5 breaks down the probabilities of repromotion by carrier type and timing of demotion. For an asset-based carrier who is demoted, the probability of ever being repromoted is about 9% for demotions occurring in 2015-2017 and about 6-7% for demotions throughout the sample. The difference between these two statistics further suggests that if a carrier does get repromoted, this occurs only after a long period of time.[55] Note also that the repromotion probability of brokers is even lower, at less than 3%. This is consistent with the lack of heterogeneity in brokers' performance; there is little to learn about individuals drawn from such a homogeneous population.

Learning as a complementary mechanism Though our data rejects learning as the *only* mechanism, learning could serve as a *complementary* mechanism for asset-based carriers. For large ABCs, a combination of punishment and learning could be at play. In particular, learning offers a strong reason for not combining a carrier's performance across all lanes, thus explaining the difference in the spatial scope of relationships between large ABCs and brokers. For small ABCs, a combination of learning and adaptation could be at play. In particular, changes in the characteristics of these carriers during a contract period, as evidenced in their need for network adjustments, invalidate Assumption 1 and thus, the prevalence of switch-back patterns predicted by our learning model.[56] Since we do not observe carriers' networks, this paper does not explore this potential learning-adaptation mechanism for small ABCs in detail; this would, however, be an interesting avenue for future research.

7 Conclusion

In this paper, we ask how informal interfirm relationships work in an economically important setting: the US truckload freight industry. We use a novel transaction-level data set uniquely well-suited to studying informal relationships to provide evidence on the mechanism governing these relationships, as well as the scope of that mechanism.

[55] Conditional on being repromoted, the average time between demotion and repromotion is about 200 days.

[56] Consider our example in Figure 5 but with a new assumption that being a primary carrier requires network adjustments. In this case, Carrier 1 would undo his investment after the shipper first switches away from him to Carrier 2. This means that when the shipper then contemplates switching away from Carrier 2, her value from switching to Carrier 1 is potentially lower than the value of continuing using Carrier 1 at the time the shipper made her first switch. Thus, switch-back patterns need not be prevalent in models that combine learning and adaptations.

We begin by presenting evidence of endgame effects, a phenomenon that suggests that the temporal scope of the incentive mechanism is within contract periods. Next, we estimate the shipper's demotion strategy. The results indicate that while the spatial scope of the demotion strategy is limited to a single lane for large asset-based carriers, shippers employ multi-lane punishment for brokers. Third, we quantify carriers' responses to the dynamic incentives generated by the incentive scheme. Finally, we address alternative, non-punishment mechanisms. Our evidence suggests that, while punishment seems to be the primary mechanism at play, other mechanisms—in particular, specific investment—likely play a role for relationships involving small asset-based carriers.

Taken together, these results provide valuable insight for future empirical work both on long-term relationships and on the trucking industry. First, we use rich microdata to empirically test the common assumption that relationship scope is at the firm-to-firm level. Our findings demonstrate the value of such a test, which we demonstrate to be feasible with increasingly available microdata. Second, by providing a detailed empirical description of the shipper-carrier relationships around which the trucking industry is organized, this paper serves as a key stepping stone to studying other important questions about this macroeconomically vital, yet understudied, industry.

References

Acocella, A. and C. Caplice (2023). Research on truckload transportation procurement: A review, framework, and future research agenda. *Journal of Business Logistics*.

Acocella, A., C. Caplice, and Y. Sheffi (2020). Elephants or goldfish?: An empirical analysis of carrier reciprocity in dynamic freight markets. *Transportation Research Part E: Logistics and Transportation Review 142*, 102073.

Baker, G., R. Gibbons, and K. J. Murphy (2002). Relational contracts and the theory of the firm. *The Quarterly Journal of Economics 117*(1), 39–84.

Baker, G. P. and T. N. Hubbard (2003). Make versus buy in trucking: Asset ownership, job design, and information. *American Economic Review 93*(3), 551–572.

Baker, G. P. and T. N. Hubbard (2004). Contractibility and asset ownership: On-board computers and governance in us trucking. *The Quarterly Journal of Economics 119*(4), 1443–1479.

Barron, D., R. Gibbons, R. Gil, and K. J. Murphy (2020). Relational adaptation under reel authority. *Management Science 66*(5), 1868–1889.

Bernheim, B. D. and M. D. Whinston (1990). Multimarket contact and collusive behavior. *The RAND Journal of Economics*, 1–26.

Brugues, F. (2023). Take the goods and run: Contracting frictions and market power in supply chains. Technical report, Working Paper.

Caplice, C. (2007). Electronic markets for truckload transportation. *Production and Operations Management 16*(4), 423–436.

Gil, R., M. Kim, and G. Zanarone (2021). Relationships under stress: Relational outsourcing in the us airline industry after the 2008 financial crisis. *Management Science*.

Gil, R. and J. Marion (2013). Self-enforcing agreements and relational contracting: evidence from california highway procurement. *The Journal of Law, Economics, & Organization 29*(2), 239–277.

Green, E. J. and R. H. Porter (1984). Noncooperative collusion under imperfect price information. *Econometrica: Journal of the Econometric Society*, 87–100.

Harris, A. and T. M. A. Nguyen (2023). Long-term relationships and the spot market: Evidence from us trucking.

Hubbard, T. N. (2001). Contractual form and market thickness in trucking. *RAND Journal of Economics*, 369–386.

Levin, J. (2003). Relational incentive contracts. *American Economic Review 93*(3), 835–857.

Li, J. and N. Matouschek (2013). Managing conflicts in relational contracts. *American Economic Review 103*(6), 2328–51.

Macaulay, S. (1963). Non-contractual relations in business: A preliminary study. *American sociological review*, 55–67.

Macchiavello, R. and A. Morjaria (2015). The value of relationships: evidence from a supply shock to kenyan rose exports. *American Economic Review 105*(9), 2911–45.

MacLeod, W. B. and J. M. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica: Journal of the Econometric Society*, 447–480.

Malcomson, J. M. (2010). Relational incentive contracts. *Handbook of Organizational Economics*.

Masten, S. E. (2009). Long-term contracts and short-term commitment: Price determination for heterogeneous freight transactions. *American Law and Economics Review 11*(1), 79–111.

Rose, N. L. (1985). The incidence of regulatory rents in the motor carrier industry. *The RAND Journal of Economics*, 299–318.

Rose, N. L. (1987). Labor rent sharing and regulation: Evidence from the trucking industry. *Journal of Political Economy 95*(6), 1146–1178.

Scott, A., C. Parker, and C. W. Craighead (2017). Service refusals in supply chains: Drivers and deterrents of freight rejection. *Transportation Science 51*(4), 1086–1101.

Taylor, D. (2011). Trucking industry updates (03/22/2011).

Weber, R. et al. (1992). On the gittins index for multiarmed bandits. *The Annals of Applied Probability 2*(4), 1024–1033.

Whittle, P. (1980). Multi-armed bandits and the gittins index. *Journal of the Royal Statistical Society: Series B (Methodological) 42*(2), 143–149.

A Proofs

A.1 Carriers' acceptance

Denote by V the average discounted expected utility of the carrier from the relationship. Denote by $V(A)$ and $V(R)$ the average discounted expected utilities of the carrier at the beginning of period t following $d_{t-1} = A$ and $d_{t-1} = R$, respectively. We have

$$\begin{aligned} V &= \mathbf{E}_{\tilde{p}_t, c_t}\left[\max\left\{(1-\delta)(\eta + p - c_t) + \delta V(A), (1-\delta)(\tilde{p}_t - c_t) + \delta V(R), \delta V(R)\right\}\right] \\ &= \delta V(R) + (1-\delta)\mathbf{E}_{\tilde{p}_t, c_t}\left[\max\left\{\eta + p - c_t + \frac{\delta}{1-\delta}(V(A) - V(R)), \tilde{p}_t - c_t, 0\right\}\right], \end{aligned} \tag{6}$$

where

$$V(A) = \sigma_0(A)V + (1 - \sigma_0(A))\underline{V}, \tag{7}$$

$$V(R) = \sigma_0(R)V + (1 - \sigma_0(R))\underline{V}. \tag{8}$$

Let $\overline{p} = \eta + p + \frac{\delta}{1-\delta}(V(A) - V(R))$ and $h(\overline{p}) = \mathbf{E}_{\tilde{p}_t, c_t}[\max\{\overline{p} - c_t, \tilde{p}_t - c_t, 0\}]$. When $\tilde{p}_t < \overline{p}$, the carrier's optimal strategy is to accept whenever $c_t < \overline{p}$. When $\tilde{p}_t > \overline{p}$, the carrier optimally rejects regardless of the cost draw. Thus, the probability of acceptance at each level of spot rate is $\Pr(d_t = A|\tilde{p}_t) = \mathbf{1}(\tilde{p}_t < \overline{p})F(\overline{p})$, increasing in the acceptance threshold $\overline{p}$. Manipulating Equations (6), (7) and (8) yields the following fixed point equation of $\overline{p}$,

$$\frac{1 - \delta\sigma_0(R)}{\sigma_0(A) - \sigma_0(R)}(\overline{p} - \eta - p) = \delta(h(\overline{p}) - \underline{V}). \tag{9}$$

Lemma 1. *$h' \in [0,1]$ and $h'' \geq 0$.*

Proof. By the independence of spot rates and cost draws,

$$h(\overline{p}) = G(\overline{p}) \int_0^{\overline{p}} (\overline{p} - c) f(c) dc + \int_{\overline{p}}^{\infty} \int_0^{\tilde{p}} (\tilde{p} - c) f(c) g(\tilde{p}) dc d\tilde{p}.$$

Thus $h'(\overline{p}) = G(\overline{p})F(\overline{p}) \in [0,1]$ and $h''(\overline{p}) = g(\overline{p})F(\overline{p}) + G(\overline{p})f(\overline{p}) \geq 0$. □

Proof of Proposition 1 Notice that if $\eta + p \in \text{Supp}(G)$, there is an option value to having contracted offers. That is, $V > \underline{V}$. This means that when $\sigma_0(A) > \sigma_0(R)$, we have $V(A) > V(R)$, and thus $\bar{p} > \eta + p$. This completes the proof of part (i).

Next, we exploit Equation (9) to generate predictions on how relationship characteristics and the reward-punishment scheme affect the likelihood of the carrier accepting, as captured by threshold $\overline{p}$. Referring to the left-hand-side and the right-hand-side of Equation (9) as LHS and RHS, we have

$$\frac{\partial (LHS - RHS)}{\partial \overline{p}} = \frac{1 - \delta \sigma_0(R)}{\sigma_0(A) - \sigma_0(R)} - \delta h'(\overline{p}) > 1 - h'(\overline{p}) \geq 0,$$

where the last inequality follows from Lemma 1. Note that since $\frac{\partial(LHS-RHS)}{\partial \overline{p}} > 0$ for all $\overline{p}$, Equation (9) has a unique solution $\overline{p}$. Also,

$$\frac{\partial (LHS - RHS)}{\partial \delta} = \frac{-\sigma_0(R)}{\sigma_0(A) - \sigma_0(R)} (\overline{p} - \eta - p) - (h(\overline{p}) - \underline{V})) < 0.$$

and

$$\frac{\partial (LHS - RHS)}{\partial \eta} = -\frac{1 - \delta \sigma_0(R)}{\sigma_0(A) - \sigma_0(R)} < 0.$$

Thus it follows from the implicit function theorem that $\frac{\partial \overline{p}}{\partial \delta} > 0$ and

$$\frac{\partial \overline{p}}{\partial \eta} = \left[1 - \frac{\delta(\sigma_0(A) - \sigma_0(R)) h'(\overline{p})}{1 - \delta \sigma_0(R)} \right]^{-1} \geq 1.$$

Furthermore, notice that $\frac{\partial h'(\overline{p})}{\partial \delta} = h''(\overline{p}) \frac{\partial \overline{p}}{\partial \delta} \geq 0$. It follows that $\frac{\partial^2 \overline{p}}{\partial \delta \partial \eta} \geq 0$. This completes the proof of part (ii).

Finally, we prove part (iii) of Proposition 1. Rewrite Equation (9) as follows

$$(1 - \delta \sigma_0(R))(\overline{p} - \eta - p) - \delta(\sigma_0(A) - \sigma_0(R))(h(\overline{p}) - \underline{V}) = 0.$$

Applying the implicit function theorem to the above equation yields

$$\frac{\partial \overline{p}}{\partial \sigma_0(A)} = -\frac{-\delta(h(\overline{p}) - \underline{V})}{1 - \delta\sigma_0(R) - \delta(\sigma_0(A) - \sigma_0(R))h'(\overline{p})} > 0,$$

and

$$\frac{\partial \overline{p}}{\partial \sigma_0(R)} = -\frac{-\delta(\overline{p} - \eta - p - h(\overline{p}) + \underline{V})}{1 - \delta\sigma_0(R) - \delta(\sigma_0(A) - \sigma_0(R))h'(\overline{p})} < 0.$$

The first inequality follows because $h'(\overline{p}) \leq 1$. For the second inequality notice in addition that from Equation (9), $\frac{\overline{p}-\eta-p}{h(\overline{p})-\underline{V}} = \frac{\delta(\sigma_0(A)-\sigma_0(R))}{1-\delta\sigma_0(R)} < 1$. □

A.2 Shipper's strategies

First, we derive the shipper's per-period payoff. Each period in a maintained relationship has three possible outcomes: either the carrier accepts the offered load, the carrier rejects because of a high cost draw, or the carrier rejects because of a high spot rate. Thus, the per-period expected utility of the shipper in the relationship equals

$$\begin{aligned} u &= \underbrace{G(\overline{p})F(\overline{p})(\psi - p)}_{\text{accepted}} + \underbrace{G(\overline{p})[1 - F(\overline{p})](-\mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq \overline{p}])}_{\text{rejected because of high cost}} + \underbrace{[1 - G(\overline{p})](-\mathbf{E}[\tilde{p}_t|\tilde{p}_t > \overline{p}])}_{\text{rejected because of high spot rate}} \\ &= -\mathbf{E}[\tilde{p}_t] + G(\overline{p})F(\overline{p})(\psi - p + \mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq \overline{p}]). \qquad (10) \end{aligned}$$

Notice that the term $\mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq \overline{p}] \leq \mathbf{E}[\tilde{p}_t]$ is the shipper's expected payment were she to be served by the spot market conditional on the carrier being willing to accept the offered load. This term represents a selection effect: the carrier has the largest temptation to reject exactly when his acceptance is most valuable to the shipper.[57] This means that even when $\psi - p > -\mathbf{E}[\tilde{p}_t]$, a relationship that cannot induce sufficiently high level of cooperation may not be worth sustaining for the shipper. The following lemma provides a sufficient condition for the relationship to be worth sustaining for any incentive scheme with $0 \leq \sigma_0(R) < \sigma_0(A) \leq 1$.

Lemma 2. *If $\psi - p + \mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq p] \geq 0$, then $u \geq -\mathbf{E}[\tilde{p}_t]$, that is, the shipper is better off offering to the carrier first than going directly to the spot market.*

Proof. Recall that

$$u = -\mathbf{E}[\tilde{p}_t] + G(\overline{p})F(\overline{p})(\psi - p + \mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq \overline{p}]),$$

[57]Note that acceptances require both a low spot rate and a low cost draw. However, under the assumption that cost draws are independent of spot rates, as in our model, the hypothetical expected payment in the spot market of an accepted load does not depend on the cost draw being low.

where $\mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq \overline{p}]$ represents a selection effect. Define

$$\hat{p} = \inf\{p' \in \text{supp } G : \psi - p + \mathbf{E}[\tilde{p}_t|\tilde{p}_t \leq p'] \geq 0\}.$$

Then $u > \underline{U}$ and $\frac{\partial u}{\partial \overline{p}} > 0$ for all $\overline{p} > \hat{p}$. The shipper should opt out of the relationship if and only if the sustained level of cooperation satisfies that $\overline{p} < \hat{p}$. Under Condition 1 that $\psi - p + \mathbf{E}[\tilde{p}|\tilde{p} \leq p] \geq 0$, we have $\hat{p} \leq p \leq \overline{p}$, so the relationship is worth sustaining. □

We now derive the shipper's average discounted expected utility. Let

$$q = G(\overline{p})F(\overline{p})\sigma_0(A) + (1 - G(\overline{p})F(\overline{p}))\sigma_0(R) \tag{11}$$

denote the probability of maintaining the relationship next period calculated at the beginning of the current period's stage game. The average discounted expected utility U of the shipper in a maintained relationship is

$$U = (1 - \delta)u + \delta(qU + (1 - q)\underline{U}), \tag{12}$$

where $\underline{U} = \mathbf{E}[-\tilde{p}_t]$ is the expected payoff of the shipper going directly to the spot market. Thus,

$$U - \underline{U} = \frac{(1 - \delta)(u - \underline{U})}{1 - \delta q}. \tag{13}$$

For $x \in \{\sigma_0(A), \sigma_0(R)\}$,

$$\frac{dU}{dx} = \underbrace{\left(\frac{\partial U}{\partial u}\frac{\partial u}{\partial \overline{p}} + \frac{\partial U}{\partial q}\frac{\partial q}{\partial \overline{p}}\right)\frac{\partial \overline{p}}{\partial x}}_{\text{incentive-inducing effect}} + \underbrace{\frac{\partial U}{\partial q}\frac{\partial q}{\partial x}}_{\text{regime-switching effect}}, \tag{14}$$

where $\partial U/\partial u, \partial U/\partial q, \partial u/\partial \overline{p}$ and $\partial q/\overline{p}$ are all positive.

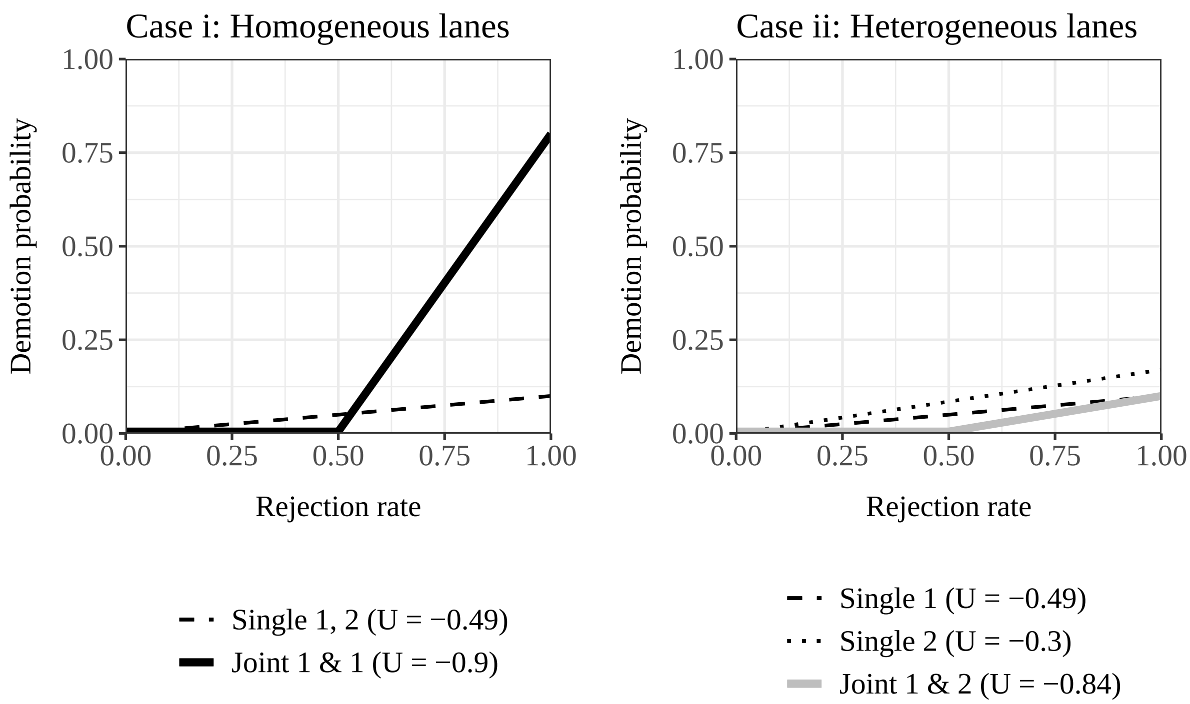
A.2.1 Proof of Proposition 2

From Equation (14), it follows from $\partial \bar{p}/\partial \sigma_0(A) \geq 0$ and $\partial q/\partial \sigma_0(A) \geq 0$ that $\partial U/\partial \sigma_0(A) \geq 0$. Thus, $\sigma_0^*(A) = 1$. To see that the optimal punishment could be soft, refer to the single-lane relationships in Example 1.

A.2.2 Details of Example 1

Under the chosen parameter values, the optimal single-lane punishment on lane 1 and lane 2 are soft. Specifically, when $\eta_1^\ell + \eta_2^\ell + p^\ell = 0.65$, the optimal demotion probability following a rejection is 0.1, and when $\eta_1^\ell + \eta_2^\ell + p^\ell = 0.8$, the optimal demotion probability following a rejection is 0.13. Define a multi-lane rejection rate as the average rejections across all lanes last period, $\frac{1}{2}\sum_{\ell=1}^{2}\mathbf{1}\{d_{t-1}=R\}$, and consider the shipper's incentive scheme that maps last-period average rejections to a probability of maintaining (or ending) both relationships.

Figure 6: Gains from pooling homogeneous lanes and losses from pooling heterogeneous lanes



The left panel of Figure 6 plots the optimal single-lane and multi-lane incentive schemes for case i. In this case, the optimal multi-lane incentive scheme forgives non-concurrent rejections but punishes concurrent rejections harshly. Specifically, if the carrier rejects on both lanes, the probability of demotion in the next period is 0.8. This multi-lane incentive scheme benefits the shipper over the optimal single-lane incentive scheme, $U=-0.9>2(-0.49)$, for two reasons. First, by forgiving non-concurrent rejections, the shipper allows the carrier to attain more allocative efficiency across the two lanes. This, in turn, increases the continuation value of the relationship for the carrier. Second, combining incentives with harsher punishments on joint rejections make these joint rejections low-probability events, and thus, clear signals of noncooperation.

The right panel of Figure 6 plots the optimal single-lane and multi-lane incentive schemes for case ii. The optimal multi-lane incentive scheme in this case takes a similar convex shape as that in case i, but punishes concurrent rejections softly. Moreover importantly, the shipper is strictly worse off by this scheme as compared to using the single-lane optimal incentive scheme, $U=-0.84<(-0.49)+(-0.3)$. This example shows that using a simple incentive scheme that conditions on average rejections (i.e., a common scorecard) across heterogeneous lanes might hurt the shipper.

B Additional empirical details

B.1 Construction of empirical variables

In this appendix, we explain the construction of right-hand side variables used in our analysis but the details of whose construction is omitted from the main text.

Inconsistency (loads / week) Our first empirical measure of the inconsistency of load timing captures a notion of how the number of loads varies from week to week within a month.

Let n^ℓ_{mw} denote the number of loads on lane ℓ in week w of month m. We then construct a measure of how much this count varies from week to week within a month by computing CV^ℓ_m, the coefficient of variation among $(n^\ell_{m1}, n^\ell_{m2}, n^\ell_{m3}, n^\ell_{m4})$. We then get a lane-level measure by averaging CV^ℓ_m over all active months on the lane:

$$\text{Inconsistency (loads / week)}^\ell = \frac{1}{M}\sum_m \text{CV}^\ell_m.$$

Inconsistency (day of week) Our second empirical measure of the inconsistency of load timing captures a notion how the timing of loads within a week varies from week to week (within a month). For instance, if a lane has 50% of its weekly loads on Monday and 50% of its weekly loads on Wednesday every week, the lane's loads would be perfectly consistent according to this measure. If, on the other hand, a lane's weekly loads were randomly allocated across days within each week, this lane's loads would be highly inconsistent according to this measure.

Our construction of this measure is motivated by a Chi-squared Goodness of Fit test of the null hypothesis that the lane's distribution of loads across days of the week is the same every week.

For any week w of month m and any day of the week d, the observed fraction of weekly volume on day of the week d is

$$O^\ell_{mwd} = \frac{n^\ell_{mwd}}{\sum_{d'} n^\ell_{mwd'}}.$$

Across all weeks within month m, the fraction of volume on day of the week d is

$$E^\ell_{md} = \frac{\sum_{w'} n^\ell_{mw'd}}{\sum_{w'}\sum_{d'} n^\ell_{mw'd'}}.$$

Under the null hypothesis, $O^{\ell}_{mwd} = E^{\ell}_{d}$ for all weeks w and days of the week d. The Chi-square statistic measures the deviation from this this null:

$$\chi^{\ell}_{w} = \sum_{d} \frac{\left(O^{\ell}_{mwd} - E^{\ell}_{md}\right)^2}{E^{\ell}_{md}}.$$

Then, we get our lane-level inconsistency measure by averaging across all active months and weeks:

$$\text{Inconsistency (day of week)}^{\ell} = \frac{1}{M} \sum_{m} \frac{1}{W} \sum_{w} \chi^{\ell}_{w}.$$

C Additional empirical evidence

Table 6: Relationship characteristics by carrier type

	Asset-based carriers (Large)	Asset-based carriers (Small)	Brokers
Monthly lane volume	9.14	8.73	6.35
Contract rate - spot rate	-0.125	-0.0878	-0.0432
Inconsistency (loads / week)	1.04	1.12	1.17
Inconsistency (day of week)	0.615	0.546	0.555
Number of lanes	6.91	4.13	11.61

Notes: This table shows differences in relationship characteristics across carrier types. The first four rows give the means of four key shipper-carrier-lane characteristics. The last row indicates the average number of lanes comprising a shipper-carrier relationship.

Table 7: Response of auction outcomes to carrier behavior

	Wins RFP	New contract premium
Rejection rate	-0.0676	-0.430
	(0.0239)	(0.0798)
Observations	1673	373

Notes: Standard errors in parentheses. This table reports estimated coefficients for regressions of the form $y^{\ell}_{scr} = \delta_0 + \delta_1 \text{Rejection rate}^{\ell}_{scr} + \delta_2 \text{Miles}^{\ell} + \epsilon^{\ell}_{scr}$ where Rejection rate$^{\ell}_{scr}$ is the proportion of tenders rejected by carrier c from shipper s on lane ℓ during the contract period and Miles$^{\ell}$ is the length of lane ℓ. In the first column, the outcome variable y^{ℓ}_{scr} is an indicator for whether the primary carrier c wins RFP r and therefore maintains the primary position on lane ℓ. In the second column, the outcome variable y^{ℓ}_{scr} is the new contract rate of carrier c after the RFP (conditional on winning the RFP and maintaining the primary position). In keeping with the event study regressions above, both regression samples are limited to mass RFP events. The sample for the second column is further limited to primary carriers who win the RFP.

Table 8: Performance of promoted carrier relative to demoted carrier

	Large ABCs		Small ABCs		Brokers	
	Acceptance	Contract rate	Acceptance	Contract rate	Acceptance	Contract rate
	-0.0945	-0.00657	-0.00688	-0.0533	0.0831	0.0234
	(0.00589)	(0.00574)	(0.00737)	(0.00727)	(0.00403)	(0.00377)
N	225177	225177	98186	98186	342371	342371

Notes: This table presents estimates from the following regression:

$$y^{\ell}_{sct} = \beta_0 + \sum_{g \in \mathcal{G}} \beta_g \mathbb{1}\{g^{\ell}_{sct} = g\} + \gamma \left(\text{Spot rate}^{\ell}_{t} - \text{Contract rate}^{\ell}_{sct} \right) + \epsilon^{\ell}_{sct}$$

where $\mathcal{G}$ represents a set of three groups of primary carriers: (a) RFP winners who are never demoted, (b) RFP winners who are eventually demoted, and (c) non-RFP winners who are promoted. The estimates listed in the table state the difference $\beta_{(c)} - \beta_{(b)}$. They therefore have the interpretation of how much the expectation of the outcome y changes when the shipper demotes the RFP winner and promotes his replacement. Outcomes are (1) an indicator for carrier c accepting load t and (2) carrier c's contract rate.