Drivers of Less-than-Truckload (LTL) Carrier On-Time Performance

by

Benjamin McGrath

and

Chun Wai Leung

Submitted on May 8, 2026 in partial fulfillment of the
Requirements of the Degree of Master of Applied Science in Supply Chain Management
at the Massachusetts Institute of Technology

Abstract

This project analyzes the on-time performance of Less than Truckload (LTL) carriers measured by the differences between their published (or scheduled) and actual transit times. By applying statistical methods including logistic and linear regression over a full year of shipment data, it was determined that shipment attributes, accessorial services, appointments being made, and regional lane characteristics all have significant influences on the likelihood and severity of delays. A set of carrier performance indication tools were developed to enhance C.H. Robinson's carrier-selection decisions with improved transit times given specific shipment characteristics. Ultimately, these findings empowered C.H. Robinson with practical, data-driven insights into the reliability of LTL services.

Capstone Advisor: Dr. Angela Acocella
Title: Research Scientist

Capstone Co-Advisor: Dr. Chris Caplice
Title: Senior Research Scientist

Table of Contents

1. Introduction

C.H. Robinson is one of the world's largest global logistics providers, offering transportation and logistics services to more than 75,000 shippers through a network of over 450,000 contract carriers, including truckload and less-than-truckload (LTL) carriers. According to C.H. Robinson (2025), they are the largest LTL third-party logistics provider in the United States, with LTL operations contributing approximately $3 billion in annual revenue.

LTL logistics involves shipments typically weighing less than 15,000 pounds, where multiple shippers share the capacity of a single truck. This transportation model provides cost efficiency and flexibility for smaller shipments but often results in longer transit times due to increased terminal handling, route circuity, and multiple delivery stops. To address varying customer needs, C.H. Robinson develops customized transportation solutions.

1.1 Motivation

C.H. Robinson procures LTL transportation services from carriers and aligns them with customer demand and shipment requirements. Carrier selection involves balancing several factors, including cost, service quality, and transit times, with transit reliability often serving as the most critical consideration for shippers. However, LTL carriers often adjust their transit data corresponding to their committed published transit time, considering various uncontrollable factors that can affect each carrier differently. The lack of objective and transparent transit data creates uncertainty in carrier selection and procurement decisions.

Actual transit times and on-time performance are influenced by a wide range of operational and external factors. Shipment characteristics such as weight, dimensions, and handling requirements can affect delivery performance, while network-related factors including route design, terminal transfers, and number of stops contribute to transit variability. Additional influences such as seasonality, weather conditions, market disruptions, and the impacts of the COVID-19 pandemic have further increased variability across the LTL industry.

C.H. Robinson believes that improving visibility into actual carrier performance and identifying the factors that drive transit variability will enable both the company and its customers to make more informed carrier selection decisions.

1.2 Problem Statement and Key Questions

LTL carriers transit times do not reliably indicate their on-time performance, and shippers lack the objective data needed to make well-informed choices.

This study seeks to address the following key questions:

- What factors contribute most significantly to deviations between published and actual LTL transit times?

- How can C.H. Robinson improve its data-driven carrier selection process through a deeper understanding of on-time performance?
- How can shippers establish more accurate transit time expectations for customers?

1.3 Project Goals and Expected Outcome

This project evaluates the relationship between carriers' published transit times and their actual executed shipment performance. The primary objective is to analyze shipment, carrier, and lane attributes to determine which factors contribute most significantly to transit time variability and late deliveries. By identifying key performance drivers, the project aims to help C.H. Robinson better understand carrier reliability, optimize carrier portfolio management, and improve customer service.

Deliverables to C.H. Robinson

- A comprehensive analysis of LTL carriers' transit time performance across different attributes
- Identification of trends, correlations, and outliers that contribute to transit time delays
- A structured set of insights to inform carrier portfolio management
- Visualizations that clearly communicate performance metrics and trends

Expected Improvements from C.H. Robinson

The anticipated outcomes for C.H. Robinson include improved visibility into carrier reliability and transit time variability, along with enhanced data-driven insights to support carrier selection decisions. Through the proposed analytical framework, the company can better identify where carriers are consistently performing well or underperforming within the LTL market. These insights enable both C.H. Robinson and its customers to establish more accurate transit expectations and make more informed shipping decisions.

2. State of the Practice

Carrier on-time performance has been a widely studied topic within freight transportation research due to its direct impact on shipper planning and customer satisfaction. Research conducted by C.H. Robinson (2023) in collaboration with Massachusetts Institute of Technology found that LTL carriers achieve an average on-time performance of approximately 71%, with shippers frequently experiencing delays of up to three days beyond published transit times.

The U.S. Department of Transportation (USDOT) also publishes performance data related to freight transportation reliability across major highway corridors in the United States. One commonly used metric is the Buffer Index (BI), which measures transit time variability by comparing the difference between the average transit time and the 95th percentile transit time relative to the average transit time. According to USDOT (2026), the corridor between Chicago and Milwaukee exhibited the highest transit variability in 2023 Q3, with a BI of 0.4438, followed by the Richmond-to-New Haven corridor at 0.2849. USDOT identifies congestion, traffic incidents, infrastructure conditions, and weather as major contributors to unreliable transit performance. Although these external variables are not directly included within this project's dataset, they provide important context when interpreting carrier performance and identifying opportunities to improve transit reliability.

From a methodological perspective, Kent and Wallbank (2021) emphasized the importance of properly structured hypothesis testing within transportation research, while also recommending complementary statistical techniques to evaluate complex datasets. Given the large number of shipment, carrier, and lane attributes included in this study, hypothesis testing plays an important role in isolating the factors that meaningfully effect on-time performance.

Regression-based approaches are central to this project's analytical framework. While linear regression remains common in empirical transportation research, logistic regression has been widely adopted in studies focused on predicting carrier on-time performance. For example, Yin and Rallis (2018) applied logistic regression to pickup and delivery data for LTL shipments and found that shipment size, transit length, and destination shipment volume were among the strongest drivers of carrier performance. Because on-time delivery outcomes can naturally be represented as probabilities, logistic regression is well suited for this study.

3. Methodology

This methodology outlines the analytical framework used to quantify variability between published and actual transit times of LTL carriers. The approach integrates industry-aligned transit-day computation, extensive data preparation, and statistical analysis to isolate specific attributes that influence transit-time performance.

The analysis incorporates two primary datasets. The first dataset contains published transit times, representing the estimated delivery performance provided by carriers or internal planning systems. These values establish the baseline expectation for shipment performance. The second dataset contains actual transit times from executed shipments, measured from origin to destination, and represents the realized carrier performance used throughout the regression analysis. Scenario testing and regression-based techniques are then applied to identify significant shipment attributes and evaluate their relationship to on-time performance and transit variability.

The data analysis process began with data cleaning and preparation to ensure the shipment records were accurate, complete, and suitable for analysis. This step included removing inconsistencies, standardizing variables, and validating shipment information.

After cleaning the data, the study compared carriers' published transit times against actual shipment performance. This comparison allowed the team to identify discrepancies between expected delivery timelines and real-world execution. From this analysis, the Published-Actual Deviation (PAD) metric was developed. PAD measures the difference between a carrier's published transit time and the actual transit time experienced by the shipment. This created a standardized method for evaluating carrier reliability and lateness across different shipments.

Next, hypotheses were developed and tested to determine which shipment characteristics most strongly influenced on-time performance and the severity of lateness. Variables such as appointment deliveries, residential destinations, weight class, and lane characteristics were evaluated to determine their statistical significance. This step helped identify meaningful operational trends and performance drivers within the dataset.

Finally, regression and statistical analyses were conducted to quantify the relationships between shipment variables and carrier performance outcomes. These models provided insight into both the probability of a shipment arriving on time and the expected severity of delay when shipments arrived late.

3.1 Defining Transit Days

Each shipment record includes the dates when freight was picked up by the carrier and delivered to the consignee. Transit time is calculated as the difference between these dates. However, shipment movement may span weekends and carrier-observed holidays, during which pickups, deliveries, and linehaul operations are often suspended. Because carriers' published transit times provided by C.H. Robinson are measured in business days, establishing a consistent definition of operational holidays was necessary to accurately evaluate carrier performance.

More than half (55.7%) of the shipment records provided by C.H. Robinson in the dataset originated from the five carriers with the highest shipment volumes. To develop a standardized business-day calendar, the holiday schedules of these carriers were reviewed. All five carriers observed the following holidays with either limited or fully suspended operations:

- New Year's Day
- Memorial Day
- Independence Day
- Labor Day
- Thanksgiving Day
- Day after Thanksgiving
- Christmas Day

After reviewing the operational schedules across carriers, the project team and C.H. Robinson agreed to use the holiday schedule observed by Carrier C as the standardized baseline for subsequent analysis given the representativeness of its holiday schedule among all carriers. Under this approach, when a holiday falls on a weekend, the adjacent Friday or Monday is treated as a non-operating holiday to maintain consistency in transit-day calculations.

With the holiday schedules defined, raw transit times were converted into business-day based transit durations by excluding weekends and carrier-recognized holidays. By utilizing the Python function `np.busday_count` in combination with a custom holiday dictionary derived from Carrier C, this allowed us to achieve actual business days across the data set.

3.2 Data Collection and Preparation

The dataset was derived from raw 2025 data provided by C.H. Robinson. Each record included origin and destination timestamps, carrier-published transit commitments, Standard Carrier Alpha Code (SCAC) codes, and supporting operational variables.

3.2.1 Data Cleansing

After computing business days between origin departure and destination arrival as described in Section 3.1, a one-day offset was applied to accommodate the industry practice of treating the departure date as “day 0.” Shipments yielding negative or over 15 transit days were deleted. This accounted for 0.23% of the data set.

Certain LTL carriers underwent consolidation or exited the market during the period covered by the data set. We obtained relevant details from C.H. Robinson relating to such market dynamics and adjusted the data set by removing LTL carriers that are no longer in the market, and merged shipment records of LTL carriers that were consolidated into the same group. This is an important step to ensure that outcomes of this study would be relevant to the latest market environment, and applicable for any further studies in the future.

3.2.2 Published-Actual Deviation (PAD)

After establishing accurate actual transit-day calculations, an error term was developed to measure deviations between expected and realized carrier performance. This metric, referred to as the Published-Actual Deviation (PAD), was calculated as the difference between actual transit days and the carrier’s published transit time.

The PAD calculation was applied consistently across all shipments, including those with delivery appointments, allowing appointment and non-appointment shipments to be evaluated within the same analytical framework. This standardized approach enabled direct comparisons across varying shipment conditions and carrier operating environments.

Negative PAD values, representing shipments delivered earlier than the published transit commitment, were censored at zero to prevent early deliveries from offsetting the impact of late shipments during analysis. As a result, PAD provided a stable and interpretable dependent variable suitable for hypothesis testing and regression modeling.

3.2.3 Categorization of Carriers and Defining Transit Lanes

Categorizing carriers and defining transit lanes are essential for any meaningful and generalizable analysis. C.H. Robinson categorized the over 80 unique LTL carriers into regional and national carriers based on their presence across the United States and segmented the country into 15 unique Key Market Areas (KMAs) based on 3-digit ZIP codes. Locations in Canada were segmented into East, West, and Central. Carriers and transit lanes in the dataset were mapped in the same way to allow analysis to be conducted based on the presence of the carriers (i.e., National, Regional, or Canada) and KMA pairs of the transit lane represented by each shipment record.

3.3 Hypothesis Testing

Hypothesis testing was used as an initial statistical tool to evaluate whether specific shipment attributes significantly influenced carrier transit-time performance. This approach provided an efficient method for identifying operational characteristics that warranted deeper investigation within the broader analysis.

The selection of attributes began through discussions with C.H. Robinson to identify variables most relevant to LTL business operations and carrier performance. Hypothesis testing was then applied to determine whether these attributes demonstrated statistically significant relationships with on-time delivery outcomes. Variables identified as significant through this process were carried forward into the regression analysis for further evaluation.

3.4 Statistical Methods for Identifying Transit Time Insights

Shipment attributes from the provided data set can be categorized into the following types of independent variables:

- Binary – a “True” or “False” indicator for various attributes, such as accessorial services, including requirements for appointments, limited access to the location, and the need for using a lift gate for delivery.
- Numerical – integers, such as published transit days of the transit lane.
- Categorical – the most common type of attributes available in the data set, including the weight class of shipments, the market segment of carriers and shippers, and the region and Key Market Area represented by the transit lanes.

Regression models described below were deployed to explore how the above types of independent variables may affect transit times.

3.4.1 Linear Regression

Linear regression was used to evaluate the relationship between shipment attributes and transit time deviation. This method estimates how changes in independent variables influence the dependent variable, allowing the magnitude and direction of each relationship to be quantified. The general linear regression equation is given by:

$$Y = a + bX + \varepsilon$$

In the context of this study, Y represents transit time deviation, X represents the shipment or operational attribute being evaluated, b represents the sensitivity of transit time deviation to changes in that attribute, and ε captures residual variation explained by factors outside the model.

Ordinary Least Squares (OLS) regression was selected due to its interpretability, computational simplicity, and availability within Excel and Python. Linear regression also provides an intuitive interpretation of model outputs by expressing the effects of shipment attributes directly in terms of additional days delayed relative to the published transit commitment.

3.4.2 Logistic Regression

Logistic regression was used to estimate the probability of a shipment arriving on time based on a set of shipment attributes. Unlike linear regression, which predicts a continuous numerical outcome such as transit-day deviation, logistic regression is designed for binary outcomes. In this study, the dependent variable was classified as either on-time or late delivery.

The logistic regression model estimates the probability of an on-time shipment using the following expression:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots)}}$$

In this equation, $\beta_0$represents the intercept term corresponding to the baseline scenario when all explanatory variables are equal to zero. The variables $x_1, x_2, \ldots$ represent shipment attributes included in the model, such as appointment requirements, accessorial services, and weight classifications. The coefficients associated with each variable quantify how those attributes influence the probability of a shipment arriving on time.

Because the model outputs probabilities bounded between zero and one, logistic regression provides an interpretable framework for evaluating carrier reliability and identifying the factors most strongly associated with late deliveries.

3.5 Statistical Analysis of PAD

The Published-Actual Deviation (PAD) metric served as the primary dependent variable throughout this study. Both the magnitude of PAD and whether its value was zero or positive were used to evaluate LTL carrier on-time performance and shipment delays.

Following the discussion in Section 3.4.1, OLS regression was applied to quantify the magnitude of the effects that shipment attributes had on PAD. Logistic regression was then used to determine how those same attributes influenced the probability of a non-zero PAD, representing the likelihood of a shipment delay.

Because many shipment characteristics were categorical, they were converted into dummy (indicator) variables prior to regression modeling. These variables included the 18 origin regions, 18 destination regions, LTL accessorial services, and shipment weight classes. The selected attributes were identified through earlier hypothesis testing and exploratory analysis to ensure the models focused on operationally meaningful drivers rather than arbitrary features. The final set of variables was also reviewed and confirmed by C.H. Robinson.

To reduce multicollinearity, one category from each variable group was intentionally omitted and treated as the baseline reference condition. Under this framework, the regression intercept represents the expected outcome for a baseline shipment profile, while each coefficient measures the relative effect of a specific attribute compared to that reference category. The baseline reference conditions adopted for the regression models are summarized in Table 1.

Attributes for Analysis	Baseline Reference
Accessorial services • Shipment with appointment time • Limited access to the delivery location • Extreme length of shipment • Hazmat shipment • Inside delivery required • Lift gate required for delivery • Residential delivery location	When these accessorial services are not applicable
Prescence of carrier • National • Regional • Canadian	National
Weight class of shipment • 50 – 85 • 85 – 175 • 175 – 500	50 – 85

Table 1. Baseline Reference Condition

3.5.1 Impact on The Number of Days Delayed

OLS regression was run by directly adopting PAD as the dependent variable and the attributes converted into dummy variables as the independent variables. The expected number of days delayed from the published transit time, in other words, the expected value of the PAD, can be derived by the sum of the intercept (a) and the coefficients (b_1, b_2, b_3 ...) of all applicable independent variables representing the explored scenario:

$$E(PAD) = a + b_{Accessorial} + b_{Weight\ Class} \cdots\cdots$$

Since all dummy variables are reduced to 0 for the baseline scenario, the expected value of the PAD, that is, the expected number of days delayed for a shipment, was obtained simply by taking the intercept (a) of the OLS regression result:

$$E(PAD_{Baseline}) = a$$

3.5.2 Likelihood of an On-time Shipment

Similarly, logistic regression was run by taking PAD as the dependent variable and the dummy variables as independent variables. The log-odds (z) were obtained from the logistic regression result by the summation of the intercept (β_0) and the coefficients ($\beta_1, \beta_2, \beta_3$ ...) of all applicable independent variables describing the explored scenario:

$$z = \beta_0 + \beta_{Accessorial} + \beta_{Weight\ Class} \cdots\cdots$$

As described in Section 3.4.2, by exponentiating the coefficients (e^{-z}), the results can be interpreted as odds ratios, indicating how much more or less likely a shipment is to arrive on time compared to the baseline condition. Additionally, these coefficients can be combined to create multiple combinations of a specific type of shipment. The probability of a shipment being delivered by its published transit time can be expressed as follows:

$$P(On\ Time) = \frac{1}{1 + e^{-(\beta_0 + \beta_{Accessorial} + \beta_{Weight\ Class} \cdots\cdots)}}$$

The probability of an on-time shipment for the baseline scenario was obtained by reducing all dummy variables to zero.

$$P(On\ Time) = \frac{1}{1 + e^{-\beta_0}}$$

4. Results and Discussion

This section presents the results of the OLS and logistic regression models conducted to identify the primary factors influencing transit-time variability and shipment lateness within the LTL network. These models were designed to quantify the impact of shipment attributes, regional characteristics, and accessorial services on transit performance. Through these findings, we identify key operational insights and patterns that help explain variability in carrier performance and network reliability.

4.1 Carrier Quadrant Analysis

We begin by putting together the carrier reliability and the severity of delays when shipments miss their delivery window into the quadrant analysis (Figure 1). The y-axis represents the percentage of shipments that are on time, while the x-axis represents the expected number of days late given that a shipment is late. The chart is divided into four quadrants using the median value of each metric, allowing carriers to be segmented relative to the overall network performance.

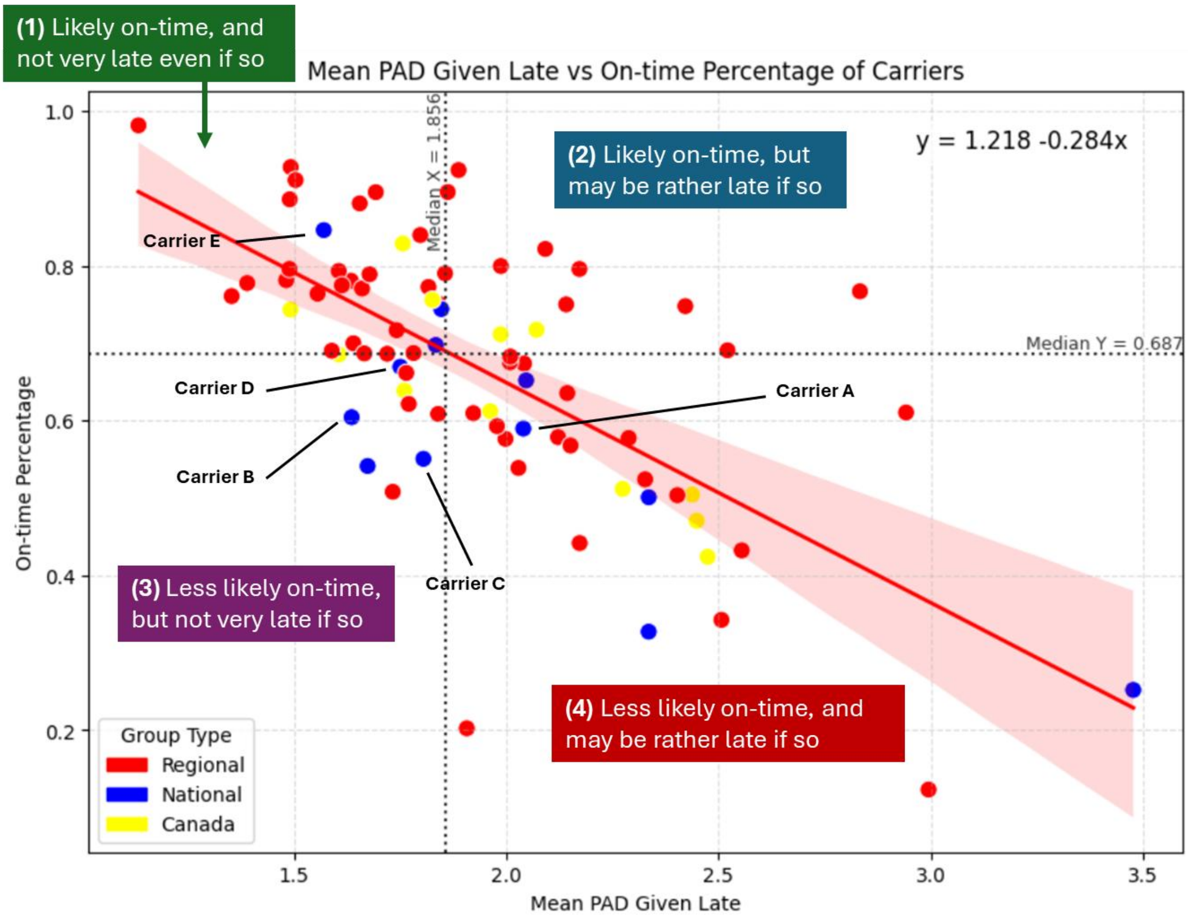

Figure 1. Carrier Quadrant Analysis

Quadrant 1 (top left) represents the strongest performing carriers, such as Carrier E: they have above-median on-time rates and below-median days late, meaning that they are usually on time and, when delays occur, they tend to be minor. Quadrant 2 (top right) includes carriers that maintain above-median on-time rates but above-median days late, indicating that they are generally reliable but experience significant delays when disruptions occur. Quadrant 3 (bottom left) represents carriers, such as Carriers B, C, and D, with below-median on-time rates but below-median days late, meaning that they are frequently late but typically only by a small margin. Finally, Quadrant 4 (bottom right) identifies the poorest performers, such as Carrier A, with below-median on-time rates and above-median days late, indicating carriers that are both frequently late and experience severe delays when shipments miss their delivery window. This framework helps distinguish carriers that occasionally experience operational failures but recover quickly from those whose delays are both frequent and operationally significant. This distinction is important because on-time percentage alone does not fully capture carrier reliability, and understanding both the frequency and severity of delays enables shippers to make more informed carrier selection and service-level decisions.

4.2 Carrier Performance per Transit Lane

To identify the best-performing carrier for any transit lane, we computed the percentage of on-time shipments and the mean and standard deviation of PAD given that a shipment is late, for each carrier within each transit lane. This information is visualized in Figure 2, where C.H. Robinson can toggle the display for any transit-lane pairs. To further facilitate carrier selection, the number of shipments handled by the carrier was presented to show respective market share on each transit lane.

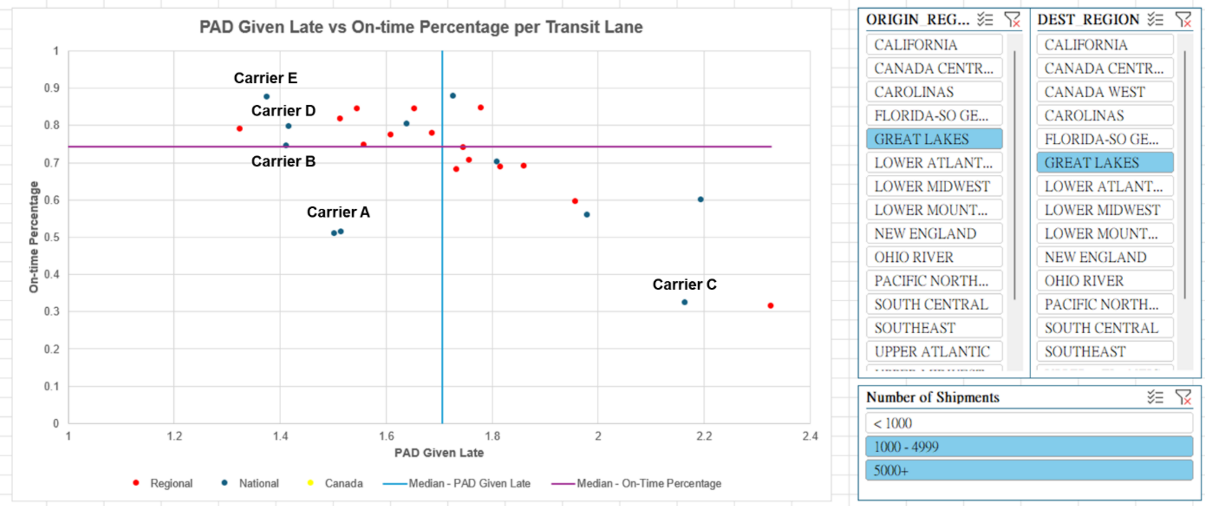

Figure 2. Visualization of Carrier Performance per Transit Lane

Once the optimal carrier is determined for a particular transit lane, C.H. Robinson can utilize the results of the regression analysis to select the most appropriate carrier based on the specified shipment profile.

4.3 Logistic Regression Results

The logistic regression adopts a baseline typical scenario of small weight class shipments without any accessorial service of a national carrier. Results in Table 2 indicate that several operational variables significantly influence the probability of shipment lateness. Although the pseudo R^2 is relatively low, this is expected given the complexity and randomness inherent in real-world logistics systems. The objective of this model is to identify the key drivers of performance. The results successfully highlight which attributes have the strongest directional impact on the outcome, providing actionable insights for decision-making.

Logistic Regression Results			
Dependent Variable: Percentage of On-time Shipments			
No. of Observations: 6,039,065			
Pseudo R-Square: 0.1175			
Variable	**Coefficient**	**z-score**	**P-value**
Constant	1.233	739.655	0.000
Shipment with appointment time	-1.717	-852.124	0.000
Limited access to the delivery location	-0.134	-30.597	0.000
Extreme length of shipment	-0.285	-66.246	0.000
Hazmat shipment	-0.016	-2.721	0.007
Inside delivery required	0.075	9.372	0.000
Lift gate required for delivery	-0.086	-23.713	0.000
Residential delivery location	-0.337	-67.323	0.000
Canadian carrier	-0.261	-33.972	0.000
Regional carrier	0.176	82.868	0.000
Weight Class of 85 – 175	-0.169	-82.618	0.000
Weight Class of 175 – 500	-0.151	-56.944	0.000

Table 2. 2025 Logistic Regression Results

The presence of an appointment has the strongest relationship with being late, with a coefficient of -1.717 ($p < 0.001$), indicating that shipments with appointments are less likely to arrive on-time. Conversely, regional shipments have the most positive coefficient, 0.176 ($p < 0.001$), indicating improved reliability compared to longer-distance shipments. Second to appointments, residential delivery locations show a negative impact on being on time, with a coefficient of -0.337 ($p < 0.001$). This suggests that shipments delivering to residential locations face a significantly higher risk of delay, potentially due to tighter delivery windows. Other accessorial variables also contribute to lateness, with extreme length of shipment (-0.285, $p < 0.001$) and limited access to the delivery location (-0.134, $p < 0.001$) both decreasing the probability of a shipment arriving on time, reflecting added operational complexity within those shipment types. Shipment weight break categories also show moderate effects, with the weight class of 85 – 175 (-0.169, $p < 0.001$) and the weight class of 175 – 500 (0.151, $p < 0.001$) associated with increased lateness risk, potentially due to additional handling or consolidation within the LTL network. Although all variables are statistically significant, the magnitude and direction of the coefficients highlight operational drivers of shipment lateness within the network.

4.4 OLS Regression Results

The OLS regression adopted the same baseline scenario as the logistic regression. It was used to estimate the magnitude of lateness (PAD_GIVENLATE) for shipments that arrived late, providing additional insight beyond the logistic regression, which only predicts the probability of a shipment being late.

OLS Regression Results			
Dependent Variable: PAD given that a shipment is late			
No. of Observations: 2,146,617			
R-Square: 0.032			
Variable	**Coefficient**	**t-value**	**P-value**
Constant	1.542	775.003	0.000
Shipment with appointment time	0.495	238.997	0.000
Limited access to the delivery location	0.078	17.656	0.000
Extreme length of shipment	0.048	10.608	0.000
Hazmat shipment	0.002	0.238	0.812
Inside delivery required	-0.089	-11.105	0.000
Lift gate required for delivery	0.021	5.625	0.000
Residential delivery location	0.185	39.684	0.000
Canadian carrier	0.281	32.626	0.000
Regional carrier	0.004	1.516	0.129
Weight Class of 85 – 175	0.092	41.327	0.000
Weight Class of 175 – 500	0.096	33.473	0.000

Table 3. OLS Regression Results

The results as shown in Table 3 indicate that shipments with an appointment time have the largest effect on the severity of lateness, with a coefficient of 0.495 ($p < 0.001$), suggesting that shipments with appointments tend to experience approximately 0.5 additional days of lateness once a shipment is already late. In contrast, shipments that involve inside delivery show a negative coefficient of -0.089 ($p < 0.001$), indicating that when these shipments are late, they tend to be approximately 0.09 days less late on average compared to the baseline category. Residential deliveries also demonstrate a notable increase in lateness magnitude, with a coefficient of 0.185 ($p < 0.001$), indicating that late residential shipments are delayed by roughly 0.2 additional days. Other accessorial variables such as limited access to the delivery location (0.078) and extreme length of shipment (0.048) also contribute to increased lateness duration, although their effects are more moderate.

The results also found that hazmat shipments and regional carriers are statistically not important in explaining PAD_GIVENLATE. The model explains approximately 3.2% of the variation in lateness ($R^2 = .032$) across 2.15 million late shipments. Together, these results complement the logistic regression analysis by showing not only which factors increase the likelihood of a shipment being late, but also which factors contribute most to the severity of lateness once delays occur.

4.5 Carrier Performance Tool

To combine findings from both OLS and logistic regression analyses, allowing for comparisons between carriers, the Carrier Performance Tool was created. This tool estimates the expected number of days a shipment will be late, conditional on the shipment already being late. The model uses coefficients derived from the OLS regression to quantify how different shipment characteristics influence the magnitude of lateness. Each variable represents a shipment attribute, such as the presence of accessorial services and shipment weight break categories. The constant term represents the baseline expected lateness in days when none of the additional characteristics are present. Each coefficient then adjusts this baseline depending on the attributes of the shipment.

Carrier Performance Tool

	SCAC	Carrier A		Carrier B		Carrier C		Carrier D		Carrier E	
		Coefficients									
		LOGIT	OLS	LOGIT	OLS	LOGIT	OLS	LOGIT	OLS	LOGIT	OLS
Constant	1	1.40	1.59	0.91	1.39	1.11	1.41	1.34	1.54	2.04	1.42
Appointment	1	-1.98	0.61	-1.95	0.65	-1.77	0.53	-1.76	0.33	-1.27	0.36
Limited Access		0.05	0.00	-0.20	0.09	-0.07	0.00	-0.25	0.00	-0.06	0.00
Extreme Length		-0.05	-0.16	-0.74	0.11	0.21	-0.21	-0.14	0.00	-0.55	0.19
Hazmat		-0.14	-0.15	-0.23	0.20	0.37	-0.21	-0.10	0.00	-0.31	0.00
Inside Delivery		0.26	-0.13	-0.07	-0.07	0.09	-0.12	0.14	0.00	-0.11	0.00
Liftgate		-0.16	-0.07	-0.13	-0.06	0.06	-0.12	-0.07	0.00	0.00	0.00
Residential		-0.57	0.34	0.00	0.05	0.25	0.08	-0.25	0.12	0.10	0.00
Weight Class 50-85		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Weight Class 85-175		-0.20	0.08	-0.23	0.04	-0.06	0.07	-0.12	0.06	-0.05	0.05
Weight Class 175-500		-0.26	0.03	-0.18	0.03	0.24	-0.04	-0.04	0.03	0.00	0.05
	Z (LOGIT)	-0.58		-1.03		-0.66		-0.42		0.78	
	P(ON TIME)	**36.01%**		**26.22%**		**33.99%**		**39.62%**		**68.48%**	
E(PAD) days|	**GIVEN LATE**	**2.20**		**2.04**		**1.94**		**1.87**		**1.78**	

Figure 4. Carrier Performance Tool

For example, in the case illustrated for Carrier C, the presence of an appointment (Appointment = 1) does have a statistically significant impact on delays. Conversely, certain shipment characteristics reduce the severity of lateness; for instance, shipments with inside deliveries reduce expected lateness by 0.12 days, while extreme length shipments reduce it by 0.21 days. By combining the constant with the relevant coefficients for a shipment's characteristics, the model calculates the expected lateness for that specific shipment profile, which in this example results in an estimated 1.94 days late given that the shipment is already late.

The other portion of the tool uses logistic regression to estimate the probability that a shipment will be delivered on time based on the characteristics of the shipment and the selected carrier. Similar to the linear regression model used to estimate the expected number of days late, this model applies a set of carrier-specific coefficients derived from historical shipment data. Each coefficient represents how a particular shipment characteristic affects the likelihood of a shipment arriving on time. The tool combines the constant term with the relevant coefficients corresponding to the shipment attributes to produce a probability of on-time delivery between 0 and 1. The probability is calculated using the formula presented in Section 3.4.2. In the example shown for Carrier A, the combination of shipment characteristics results in an estimated 36.01% probability that the shipment will arrive on time.

By combining the OLS and logistic regression models, the tool enables the sponsor company to evaluate both the probability of on-time performance and the expected magnitude of delays, allowing for more informed carrier selection and shipment planning based on the specific characteristics of each shipment. The Carrier Performance Tool also allows the sponsor company to visualize carrier performance under different shipment conditions by simply selecting a carrier and adjusting shipment attributes. Because the coefficients were estimated separately for each carrier, the model captures how different carriers respond to such operational characteristics as residential deliveries, appointment requirements, or shipment size. As a result, the tool provides a flexible “plug-and-play” framework for evaluating expected delay severity across carriers, enabling more informed carrier selection and operational planning.

This analysis shows that on-time performance and the severity of lateness represent two distinct dimensions of carrier performance. A carrier may deliver most shipments on time but still experience significant delays when shipments are late, meaning these metrics should be evaluated separately. In addition, the impact of shipment characteristics and accessorial services varies across carriers.

4.6 Notes on Shipments with Appointments

As discussed in Section 4.3 and 4.4, delivery appointments have the strongest impact on carrier performance. To better evaluate its effect, the appointment (APT) variable was further analyzed. The original APT variable only identified whether a shipment required a delivery appointment and did not account for how the appointment itself influenced the shipment’s ability to arrive on time. In practice, appointments can either provide additional flexibility in delivery timing or introduce scheduling constraints that affect carrier operations.

To address this, an alternative PAD metric was calculated as the difference between the shipment’s actual arrival date and the effective delivery deadline, defined as the scheduled appointment date. Under this framework, when an appointment occurs after the originally published transit commitment, the appointment date becomes the operationally relevant delivery expectation.

After the alternative calculation of the PAD was adopted for shipments with appointments made, carriers with a higher percentage of shipments with appointments was found achieving a much better on-time performance. This is consistent with the fact that over 65% of appointments were made later than the published transit time, as we found in the dataset. It is intuitive that because carriers with a later appointed delivery time are given more time to transport a shipment to their destination, their on-time performance measured by the alternative PAD metric will improve. Appointments give shippers more certainty about delivery time, while not necessarily guaranteeing faster delivery.

5. Recommendations

This project supports the strategic priorities of C.H. Robinson by leveraging data analytics to improve decision-making and customer service within its LTL network. Using regression and logistic modeling techniques, the analysis evaluates how shipment characteristics influence both the likelihood of on-time delivery and the severity of delays. The findings show that on-time performance and lateness are driven by different factors, highlighting the need to augment their current approach to evaluating carrier performance rather than relying on traditional assumptions or perceived performance.

5.1 Management Recommendations

The results provide actionable insights into the key drivers of carrier performance, including lane characteristics, accessorial services, and geographic factors, which vary across carriers. Based on these insights, it is recommended that C.H. Robinson implement a data-driven carrier selection tool enhancing their current methods of carrier selection. Layering this research and modeling into current tools help match execution to ever evolving shipper requirements. Additionally, segmenting carriers by both on-time probability and delay severity can improve routing decisions, while integrating predictive analytics into routing guides can enhance overall network efficiency. By selecting carriers tailored to their shipment requirements, shippers can expect improved on-time performance. From a strategic perspective, this fosters reliable trust and cultivates enduring customer relationships, ultimately supporting greater customer retention for C.H. Robinson.

5.2 Future Work

This study relied primarily on historical shipment data and traditional statistical modeling techniques. While these approaches provided meaningful insight into the relationships between shipment attributes and carrier performance, future research could expand the analytical framework using more advanced machine learning techniques. Models such as Random Forest and XGBoost could better capture complex nonlinear relationships and interactions within the dataset that may not be fully represented by traditional regression methods. The ability to iteratively tune and optimize these models could significantly improve predictive accuracy and provide deeper operational insight into the factors contributing to shipment delays.

Future work could also incorporate real-time operational data sources into the analysis. External variables such as weather conditions, traffic congestion, infrastructure disruptions, and regional market conditions can significantly influence transit reliability but were not included. Integrating these factors could improve model performance and provide a more comprehensive understanding of shipment variability across the LTL network. Additional research could investigate carrier performance across specific shipper industries and shipment profiles. Industries such as healthcare, retail, manufacturing, and food distribution often operate with different service requirements, freight characteristics, and sensitivity to delivery disruptions. Segmenting carrier performance by industry could provide deeper insight into which carriers perform best under specific operating conditions and support more specialized carrier selection strategies. Future analysis could also further explore the relationship between transit reliability and shipment distance traveled. While lane and regional effects were incorporated, a more detailed evaluation of distance-based performance may reveal nonlinear relationships between haul length, network complexity, and delay severity. Understanding how carrier performance changes across short-, medium-, and long-haul shipments could provide additional insight into routing efficiency and operational constraints within the LTL network.

Changes in market structure may also influence carrier performance and warrant additional study. One example is the shutdown of Yellow Corporation in July 2023, which was the third-largest LTL carrier in the United States at the time (CNN, 2023). The redistribution of freight volume following Yellow's exit likely altered carrier capacity, routing dynamics, and transit reliability across the industry. Future research could evaluate whether carrier performance and delay patterns shifted across specific lanes and shipment attributes as remaining carriers absorbed this additional demand.

Finally, future work could extend beyond predictive analytics into prescriptive optimization. By combining predictive models with optimization techniques, organizations could develop decision-support systems capable of recommending optimal carrier assignments and routing strategies based on predicted delay risk. Such systems would allow logistics providers to proactively mitigate service failures, improve network reliability, and support more effective operational decision-making for both carriers and shippers.

References

C.H. Robinson Worldwide, Inc. (2023). *Key Insights to Improve LTL Transit Times.* https://www.chrobinson.com/en-us/resources/blog/key-insights-to-improve-ltl-transit-times/

C.H. Robinson Worldwide, Inc. (2025). *Annual Report on Form 10-K for the Year Ended December 31, 2024.* https://investor.chrobinson.com/Financials/Annual-Reports/default.aspx

CNN. (2023). *99-year-old trucking company Yellow shuts down, putting 30,000 out of work*. https://www.cnn.com/2023/07/31/business/yellow-corp-closing

Kent, J., & Wallbank, C. (2021). *The use of hypothesis testing in transport research.* TRL. https://doi.org/10.58446/rrzh8247

U.S. Department of Transportation, Bureau of Transportation Statistics (2026). *Freight Facts and Figures.* https://data.bts.gov/stories/s/Freight-Transportation-System-Condition-Performanc/vvk5-xjjp/#system-performance

Yin, B., & Rallis, C. W. (2018). *A Study of Shipper Performance in the Less-Than-Truckload Market.* Massachusetts Institute of Technology. http://hdl.handle.net/1721.1/118140