Data-driven Risk Assessment for Truckload Service Providers

By Sriram Kishore Chittella and Marcos Machado Teixeira
Thesis Advisor: Daniel W. Steeneck

Topic Area: Transportation

Summary: This research proposes a methodology for truckload service providers to better understanding the risk and the volatility patterns when dealing with the trucking spot market across the different geographies, time periods and others, allowing them to reduce their risk when setting long-term contracts with shippers. Using three different measures of volatility (coefficient of variation, beta, and month over month percentage change) we developed a framework to assess the company’s risk profile and to distinguish regions and lanes from a risk perspective, arising with insights to improve 3PL companies’ business.

Introduction

Non-asset backed third-party logistics companies provide shippers access to a flexible source of capacity through trucking spot market. However, the increased volatility in the trucking spot market rates is turning the 3PL businesses more risky and complex. To maximize profitability of their business, it is important to 3PL companies to develop an understanding of how the risk and the volatility patterns of the spot market rates vary across different geographies, time, distance and others.

KEY INSIGHTS

1. Multiple measures to measure risk are required as each of them indicate different characteristics of risk. As no one size fits all, a holistic view is crucial.
2. Volatility characterization is geographic. As seen in the analysis, volatility varies geographically across the various regions. Certain regions and lanes are more risky than others, so further sub grouping is necessary.
3. Factors that influence one region does not necessarily influence the other. Each region and each lane behaves uniquely. While a characteristic is true for a region, it may not be true for a lane within that region.
Throughout this research, we worked with one of the largest North American providers of non-asset based logistics services. The company’s goal is to serve shippers wishing to outsource some or all of their transportation needs. In this role, the company negotiates rates with carriers and manages day-to-day operations for its clients, effectively matching loads with carrier availability and capacity. By having a better understanding of the risk and the volatility of the spot market rates, 3PL companies can mitigate market uncertainty.

**Methodology and Dataset**

The original data was provided by our sponsor, and the dataset includes shipment records of dry van equipment, nationwide (US) over the past three years (From October 2011 through September 2014). The data is composed of 1,609,594 transactions during this three years period.

To be able to compare lanes with different distances, we normalized the shipments’ cost by transforming the total cost per shipment into a per mile basis. The fuel cost was isolated from the line-haul cost to avoid any fuel related cost fluctuations.

We aggregated and selected the data for analysis purposes. On top of the segmentation we decided to aggregate the data in terms of origin and destination pairs. The data was aggregated in lanes by 3 digit origin and destination zip codes to improve significance and relevance.

The data set was then further transformed to categorize the loads by region based on the state of origin. This was done by grouping each of the states into 5 categories: West Coast, Rocky Mountain, Midwest, East Coast and Gulf Coast. East Coast is divided into New England, Central Atlantic and Lower Atlantic Regions. District of Columbia is grouped under the Central Atlantic region. This is based on the Department of Energy Petroleum Administration for Defense Districts map.

To assess and evaluate risk in trucking spot market operations, we developed a framework using three different measures to quantify volatility in line-haul rates: coefficient of variation, beta value and average month/month percentage changes. In addition, based on all three measures, we generated risk profiles at national and regional level. The risk profiles are histograms indicating the spread and frequencies of the volatility measures among the sample analyzed. In order to demonstrate the application of the research in supporting long-term contract pricing, we used one lane as example, utilizing histogram of the historical cost per mile prices in this lane. Then, a distribution fit was performed for this histogram, and based on this information we proposed a methodology for pricing this lane.

Finally, using regression analysis we tested which factors, such as GDP Index, seasonality, number of trucks at origin, and commodities seasons (agricultural activity), influence or do not influence the volatility in the spot market rates in a given region.

**Results**

1. **Data Exploration**

First we calculated the monthly average line-haul cost per mile. We call these as the monthly national indices, and this forms the basis of the core analysis for spot market rates volatility.

Graphing the national indices, this indicates that there is both a trend and seasonality in the average cost per mile nationally. Typically, the peaks in the cost rates are occurring in the summer months and the dips are occurring in the winter months. However in the case of 2014, the dip is not prominent in the winter months likely due to harsh nation-wide weather conditions.

Once the data is separated by regions, the regional indices are also generated. Figure 1 shows the monthly regional index.

![Figure 1 – Regional Index (monthly view)](image-url)
2. Volatility and Correlation

To quantify the volatility in the lanes, we calculate beta values, coefficient of variation and average month/month percentage change. First the volatility of each of the regional indices is determined in comparison to the company national index (based on beta). Figure 2 visualizes the volatility of regional indices in comparison with the national index as a function of time to better identify the trends in the volatility.

Figure 2 - Volatility by beta of regional indices (in comparison with national index)

It illustrates that the Rocky Mountain region is more volatile than the rest of the markets that the company operates in. The peaks indicate that the rates are highly volatile in comparison to the national index, moving in the same direction as the national index. On the other hand, the Mid West and Lower Atlantic regions are flat and the beta values consistently hovering around 1 showing that both these regions behave identical to the national index and are stable in comparison to the national index.

The Coefficient of Variation is also calculated based on the monthly values of $\sigma$ and $\mu$ during 2012 and 2013 on a national and regional level. The national CV for this timeframe is 21.4%. Rocky Mountain again is the most volatile region, with a CV of 72%, while Midwest is the least volatile region with a CV of 21.0%.

Next, volatility is calculated by examining the average month over month percentage change in the cost per mile rates during the 2012 and 2013 time horizon based on the monthly rates. The national and regional results are presented below:

- National: 1.51%
- Central Atlantic: 1.69%
- Gulf Coast: 1.21%
- Lower Atlantic: 1.48%
- Midwest: 1.67%
- New England: 4.32%
- Rocky Mountain: 5.17%
- West Coast: 1.30%

3. Risk Profiles

Figure 3 shows the national risk profile based on beta as a measure of risk.

Figure 3 – National risk profile (2012 and 2013 beta)

A higher proportion of lanes fall within the beta value of -1 and 1 indicating that a large number of lanes are less volatile than the overall national index. The end points of -14.1 and 4.7 indicate that there are a small number of highly volatile lanes. National and regional risk profiles were also generated based on coefficient of variation and average M/M % change.

4. Regression Analysis

During the regression analysis, it was observed that each region behaves uniquely, which was already expected based on the other analysis performed during this research. The variable or set of variables that influence(s) one region does not necessarily influence(s) the other. Crop seasons plays a role in some regions (e.g. Gulf Coast and Lower Atlantic) but does not play in others (e.g. Rocky Mountain). The same happens with other factors.
Insights and Conclusion

The methods we developed in this research provide a contribution to truckload business in terms of better understanding the trucking spot market volatility. Applying three different measures of volatility, we were able to distinguish regions and lanes from a risk perspective, which leads us to some important conclusions:

I. Volatility characterization is holistic: As seen in the analysis, it is essential to take a holistic approach when assessing volatility and risk. In this case three measures i.e. coefficient of variation, beta and average month/month percentage change were developed to quantify the volatility in lanes and regions and characterize them. This ensured that though a lane may appear to be “safe” under one type of volatility measure, its “risk” as measured by another type of volatility measure is accurately captured.

II. Volatility characterization is geographic: As seen in the analysis, volatility varies geographically across the various regions. Some regions such as the Rocky Mountain and New England are evidently more volatile than rest of the markets. Hence, it is important to view the regional differences in volatility when assessing the risk faced by the 3PL business. In situations where individual lanes do not follow the regional characteristics, sub-region categorization and individual lane based volatility analysis will help improve the precision.

III. Volatility factors vary geographically: The factors that influence the volatility in one region or lane do not necessarily impact the other regions or lanes. Hence, the factors have to be tested for their influence across the various regions and lanes to understand volatility patterns.

In addition, this research reveals many interesting opportunities for future work. It would be interesting to incorporate other external factors that might influence spot market rates volatility, to identify other potential segmentation in terms of distance, regions, lanes, corridors, and time periods (monthly vs. weekly), and to investigate potential connections between financial markets and line-haul spot price markets.