Title: Improving Freight Efficiency with Load Matching Technology

Project Number: 15-2.1e
Years: 2015-2018

Final Report

Date: October 2018

Principal Investigator

Dr. Seiji Steimetz

Researchers

Anete Brinke
Benjamin Olson

MetroFreight Center of Excellence
California State University Long Beach
Center for International Trade and Transportation
6300 State University Drive, Suite 255
Long Beach, California 90815

Table of Contents

List of Tables

Title: Improving Freight Efficiency with Load-Matching Technology

Abstract

Purpose: Load-matching technology for truckers and shippers helps an inefficient and often fragmented local trucking market by eliminating non-revenue-generating trips. The basic idea of the technology is to provide a real-time, GPS-based connection between shippers and carriers, somewhat similar to how Uber and Lyft connect drivers and passengers.

There is reason to believe that the market for this kind of service will only grow; however, expansion will depend upon a combination of economic and political factors. The roll-out of load-matching services in Los Angeles will provide useful lessons for their adoption in other locations.

Approach: Our research investigates the role that data-driven analytics can play in improving goods-movement efficiency through load-matching technology. We have secured access to one company's data for analyzing factors that influence the supply of load-matched carriers in short-haul trucking operations. At the time of acquiring the data, this company was using load-matching technology to connect over 400 businesses with more than 700 owner-operators in New York, Los Angeles, and the San Francisco Bay Area.

Research Goals:

1. Obtain and export data on acceptance rates and shipment characteristics (including prices).
2. Prepare data for statistical analysis.
3. Develop an econometric model relating acceptance rates and likelihood of touring to shipment characteristics.
4. Determine how those characteristics influence acceptance rates and likelihood of touring. For example, determine how a one-percent increase in price affects the probability of acceptance.

Findings: This project might generally be described as a "feasibility study". But its potential gains are substantial relative to its cost. For example, analyzing the tradeoffs carriers make when deciding whether to pick up a shipment might reveal pricing strategies that improve adoption. As a result, fewer truck trips would be generated, thereby reducing highway congestion and pollutant emissions. The case is the same in the analysis of the tradeoffs carriers make when deciding whether to link multiple shipments into a single trip.

While this project focuses on load-matching operations in Los Angeles, we anticipate that its findings will readily generalize. Moreover, they will provide a unique, analytical perspective on

how to implement load-matching technology in untapped markets, thereby expanding the potential for that technology to improve the efficiency of goods movement nationwide.

Research Impact: Load-matching is still a nascent technology, and there is still much to learn about how it can be improved to further reap its efficiency gains. Our approach allows us to examine the availability and condition of the data generated by load-matching operations. This includes information on pricing, transit times, service reliability, lead and turnaround times, and characteristics of shippers and operators.

Practical Impact: The general benefits of load-matching services are somewhat obvious. Full trucks mean fewer trucks, resulting in less highway congestion and reduced truck pollution.

Introduction

Load-matching technology, or digital freight matching (DFM), is used in freight movement by shippers to connect with carriers to move their goods from point A to point B. This research investigates the habits of shippers regarding what shipments they pick up: i.e., what factors influence the probability they will pick up a shipment posted to the virtual board or why they “link”, which is when multiple shipments are picked up in one area and are all delivered to one vicinity.

Digital freight matching could reduce truck traffic and pollution by reducing empty backhauls or “deadhead” trips and by consolidating shipments and coordinating “milk runs”, or linked shipments.

Project Objective

The research seeks to investigate DFM shipment data to develop an econometric model relating shipment characteristics to acceptance rates and the likelihood of touring. The results can then be used to determine how the characteristics influence driver behavior which can suggest better pricing and timing strategies to be implemented to increase efficiency and reduce truck emissions.

Project Description

This research develops two models analyzing the data from a DFM company. Both models employ a binary logit model. The first uses a dependent dummy variable with value of 1 indicating the shipment was picked up by an owner-operator carrier and 0 indicating the platform failed and an owner-operator did not pick up the shipment, and the DFM company had to contract a driver to recoup the loss. The dependent variable in Model 2 is 1 if a shipment was linked into a trip of two or more shipments, and 0 if it was a standalone shipment. Independent variables for these models include shipment weight and distance, pickup and delivery timing characteristics, and payment offered per mile and per ton. Furthermore, the study descriptively analyzes the current trends in the market for this DFM company, such as the median shipment weights and the number of trips picked up as well as linked either by origin location, destination location, or both.

Research Approach

Data Cleaning

Prior to analysis and modelling, several data assumptions prompted the need for cleaning of the raw dataset. The statistical programming software RStudio was used to perform this cleaning. The model assumptions are outlined here:

- Shipments must be short-haul
- Shipments must be using box trucks (rather than 53-footers)
- Shipments must all be within the Greater Los Angeles Area

In order to clean the data in alignment with these assumptions, shipments were removed from the raw data based on criteria. First, to exclude shipments that were not short-haul, shipments labelled “long-haul” or “dayrate” in the *shipmentType* column were removed as well as shipments with delivery distance greater than 100 miles. To include only box truck shipments, only shipments with weight between 2 and 12000 pounds were included. Furthermore, shipments that were posted by shippers that ended up with a carrier but were then canceled, for whatever reason, were removed from the data. Via communication with the company that provided the data, this was done by removing shipments with no revenue.

Next, data for only one market was included in the analysis by removing shipments with other geographical codes in the *market* column. Also, outliers were removed by excluding shipments with *revenue* or *driverpay* greater than or equal to $2000.

Next, certain dummy variables were created to test their significance in the model. These are outlined in Table 1 below

Table 1: Variables coded from raw data for modelling purposes

Variable name	Description
pickupFlexibility	1 if the *pickupAfter* and *pickupBy* times are greater than zero hours apart 0 otherwise
deliveryFlexibility	1 if the *deliverAfter* and *deliverBy* times are greater than zero hours apart 0 otherwise
pickupMorningPeak	1 if the *pickupBy* time is between 6am and 9am 0 otherwise
pickupEveningPeak	1 if the *pickupBy* time is between 3pm and 6pm 0 otherwise
deliverMorningPeak	1 if the *deliverBy* time is between 6am and 9am 0 otherwise

deliverEveningPeak	1 if the *deliverBy* time is between 3pm and 6pm 0 otherwise
oneDay	1 if the *pickupBy* date and *deliverBy* date are the same 0 otherwise
isEnterprise	1 if the *Enterprise.Customer* column indicates "yes" 0 otherwise
ooPickedUp (dependent variable)	1 if *revenue* is less than or equal to 4 and *driverpay* is greater than or equal to 3 0 otherwise

Other variables that were coded from the raw data include *payPerMile* (payment to driver divided by distance), *payPerTon* (payment to driver divided by weight), *tripCount* (number of shipments in trip), and *tripPickupCount* (number of picked up shipments in a trip).

Identification of linked trips

Certain criteria were set in order to identify trips as linked. A "linked trip" is defined as a driver picking up two or more shipments in the same truckload. There are three different ways this analysis classifies linked trips: by origin city, by destination city, or by both ("super" linked). That is, a driver perhaps picked up two or more shipments from the same origin city and took them out to different cities in the same area, and thus the shipments were linked by origin. Alternatively, a driver went around to a few different cities in the same area to pick up two or more shipments and then delivered them all in the same city, and the shipments were then linked by destination. In some cases, the driver picked up two or more shipments from the same city and then took them all to one other city, and in this case the shipments are linked by destination and by origin, which was dubbed a "super" link. Table 2 outlines the criteria used to match shipments in the raw data as part of the same linked trip. Additionally, shipments must have been identified as having been picked up (ooPickedUp = 1) in order to be considered part of a link. For the purposes of this study, *shipment* refers to a single shipment posted to the DFM platform and *trip* refers to a single trip that a driver takes to deliver one or more shipments. When two or more shipments are part of the same trip, they are referred to as *linked* and the trip is a *linked trip* (commonly referred to as a "tour").

Table 2: Criteria for identifying shipments in the raw data that are part of the same trip

Link type	**Definition criteria**				
	Driver ID	*Origin City*	*Destination City*	*"Pickup by" date*	*"Deliver by" date*
Origin	Same	Same	Different	Same	n/a
Destination	Same	Different	Same	n/a	Same
Super	Same	Same	Same	n/a	Same

Four dummy variables were created to indicate shipment linking. *isNotLinked* indicates whether a shipment was linked in any way. *isOriginLinked*, *isDestinationLinked*, and *isSuperLinked* all indicate whether a shipment was linked in each respective way. It must be noted that if a shipment is indicated as “super” linked, it is not indicated as both origin and destination linked as well.

Data

The data for this project was received from a digital freight matching company under a confidentiality agreement.

Analysis and Results

Descriptive Statistical Findings

Table 3 below summarizes the percentages of shipments and trips that are picked up by drivers and the various link types in the sample.

Table 3: Shipment and trip statistics

	Picked Up	Linked	Origin Linked	Destination Linked	"Super" Linked
Shipments	88%	57%	39%	12%	6%
Trips	80%	23%	13%	6%	4%

These percentages reveal that 88% of shipments were picked up by a subscribing driver ("driver"), which is a higher percentage than previously thought for DFM platforms.

Additionally, the weights of the shipments were analyzed. It was found that the median weight of trips with a single shipment is 451 pounds while the median weight of trips with at least two linked shipments is 1,563 pounds. The linked trips deliver shipments that are three times as heavy as individual shipments, indicating that linking shipments is associated with fuller truckloads. The mean weight of linked shipments is twice as large as that of the individual shipments. It should be acknowledged that these numbers portray an underutilization of the full capacity of box trucks. However, it is important to note that this analysis includes different sized vehicles and their full capacities vary. Therefore, we cannot distinguish between box trucks and smaller trucks in the data to make conclusions about the utilization of truck capacities. Nevertheless, this finding is noteworthy because it shows that linked trips deliver more heavy truckloads than trips with individual shipments.

Modelling

Two models were developed to examine the effects of different shipment characteristics on the probability of a shipment being picked up and the probability of a shipment being linked.

Binary logit estimation was used when modelling because it is cogent when the dependent variable is binary, i.e. takes the value 0 or 1. Stata was used to perform the modelling.

Model 1: Binary logit estimation of whether a shipment was picked up by a carrier

This model uses each individual shipment as a data point and *ooPickedUp* as the dependent variable. Table 4 shows the marginal effects of the model. The sample size is 113,486.

Table 4: Marginal effects of independent variables in Model 1

Independent variable	Percentage-point effect	Z-statistic
Pickup Flexibility	10.5	20.03
Delivery Flexibility	-4.5	-5.4
Pickup Morning Peak	1.4	54.19
Pickup Evening Peak	-1.4	-12.92
Delivery Morning Peak	0.12	45.65
Delivery Evening Peak	-0.8	-8.68
One Day	-10.8	-15.68
Enterprise Customer Indicator	-2.5	-14.2
Weight	2.3	55.18
Distance	2.4	16.69
Pay Per Ton	0.31	8.06
Pay Per Mile	0.25	6.22

The marginal effects of Model 1 are mostly consistent with intuition. Having pickup flexibility and higher payment per weight and distance increases the probability that a driver will pick up a shipment. Having a pickup time during the morning peak traffic hours also increases the probability of a pickup: this is likely due to most drivers looking to start their work day in the morning. Higher shipment weights and distances also increase the probability of a pickup. A lack of delivery flexibility, pickup and delivery times during the evening peak hours, and required same day delivery all decrease the probability of a shipment being picked up. In a background analysis, it was concluded that drivers show a stronger response to increases in payment during evening peak traffic hours for shipment pickups.

Model 2: Binary logit estimation of whether a shipment was part of a linked trip

This model analyzes the factors that influence whether or not shipments are part of a linked trip. Table 5 shows the marginal effects of the model. The sample size is 100,221.

Table 5: Marginal effects of independent variables in Model 3

Independent variable	Percentage-point effect	Z-statistic
Pickup Flexibility	-6.7	-6.47
Delivery Flexibility	-0.6	-0.54
Pickup Morning Peak	4.1	115.28
Pickup Evening Peak	-0.7	-6.24
Delivery Morning Peak	0.2	10.69
Delivery Evening Peak	-0.1	-1.22
One Day	15.6	26.98
Enterprise Customer Indicator	17.9	111.22

Weight	-3.7	-50.95
Distance	-5.2	-23.69
Pay Per Ton	-0.7	-15.28 9
Pay Per Mile	-1.2	-20.28

The negative signs for pickup flexibility, delivery flexibility, pickup and delivery during evening peak hours, weight, distance, and payment variables all indicate that these characteristics decrease the probability of drivers picking up more than one shipment. The negative coefficients for pickup and delivery flexibility and required same-day delivery indicate that drivers are looking for a schedule with less uncertainty when they are trying to handle more than one shipment at once. Heavier packages decrease the probability of shipment linking because heavy packages take more time to load and drivers might be constrained by the size of their trucks. Longer shipment distances are undesirable when linking because they potentially result in drivers having to go out of their way more.

Higher payment per shipment, on average, does not incentivize drivers to link shipments. This can be explained by drivers trying to capture higher profits by carrying multiple shipments at once rather than one high-paying individual shipment.

Analyzing average number of shipments per link type

Looking at the mean number of shipments per trip link type reveals within which link type drivers include more shipments. Table 6 below summarizes these statistics.

Table 6: Average number of shipments per link type

Trip Link Type	**Sample Mean**
Origin Linked	5.5
Destination Linked	3.2
Super Linked	2.6

This is insightful because linking shipments by origin is associated with a higher number of shipments per trip than destination or “super” linking. This is possibly because origin linking is logistically easier to manage for drivers. These numbers reveal that more can be done by DFM companies to make destination and “super” linking easier for drivers.

Conclusions

The analysis performed in this research presents several conclusions on the preferences of drivers. These represent the ways they can be incentivized to pick up and link together shipments posted to the DFM platform in a manner than reduces the number of trucks on the road and thereby reduces emissions.

First, the analysis reveals that most shipments posted to the platform are picked up and a majority is linked in a trip of at least two shipments. Precisely, 88% are picked up by truck owner-operators and 57% are linked.

The linked trips deliver shipments that are three times as heavy as individual shipments, meaning that linking shipments is associated with fuller truckloads.

Furthermore, the results from Models 1 and 2 show that the importance of timing characteristics of shipments to drivers is important depending whether the DFM company is looking to pick up shipments or link shipments. Timing flexibility (pickup and delivery windows and required same-day delivery) is only preferred as an incentive to have shipments picked up. Timing flexibility is undesirable when drivers seek to link shipments together. Interestingly, higher shipment payment per ton and per mile does not incentivize drivers to link. This is explained by drivers trying to capture higher profits by carrying multiple shipments at once rather than one high-paying individual shipment. Heavier shipments are viewed as an inconvenience to drivers when attempting to link shipments because they require more time to load and they may be constrained by the size of their trucks, which is why heavier packages decrease probability of shipment linking. From these results we can imply that drivers will generally opt for multiple shipments with less individual weight and shorter distances to capture higher pay.

Also, linking shipments by origin is associated with higher number of shipments per trip than destination and “super” linking. This is possibly because origin linking is logistically easier to manage for drivers. These numbers reveal that more can be done by DFM companies to make destination and “super” linking easier for drivers.

This analysis serves as a proof of concept that DFM platforms have the potential to alleviate congestion problems and deadhead miles. Inferences about driver behavior can be made to gain insights on how to improve DFM pricing strategies. However, more robust datasets would improve analyses and provide additional insights.

Acknowledgment

This research is supported by the Volvo Research and Educational Foundations through the MetroFreight Center of Excellence, at California State University Long Beach. The author acknowledges the data provider (confidential due to a non-disclosure agreement) and the researchers Anete Brinke and Benjamin Olson for their contributions to this research. All errors and omissions are the responsibility of the authors.