RIETI Discussion Paper Series 12-E-016

Determinants of Transport Costs for Inter-regional Trade

KONISHI Yoko
RIETI

Se-il MUN
Kyoto University

NISHIYAMA Yoshihiko
Kyoto University

Ji Eun SUNG
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 12-E-016
March 2012

Determinants of Transport Costs for Inter-regional Trade

KONISHI Yoko
Research Institute of Economy, Trade and Industry

Se-il MUN
Kyoto University

NISHIYAMA Yoshihiko
Kyoto University

Ji Eun SUNG
Kyoto University／Research Institute of Economy, Trade and Industry

Abstract

This paper presents a microeconomic model of inter-regional freight transportation based on careful formulation of the cost structure in trucking firms and market equilibrium, which takes into account the feature of transport service as a bundle of multiple characteristics. We estimate the parameters of the model using the micro-data of inter-regional freight flows from the 2005 Net Freight Flow Census in Japan. Estimation results show that the determinants of transport cost incorporated in the model have significant effects in the ways that the model predicts. The degree of competition also has significant effect on freight charge. It is shown that there exist significant scale economies with respect to lot size and long-haul economies. The quantitative extent of these effects is also demonstrated .

Keywords: Freight charge, Cost structure, Transport service, Micro-shipments data
JEL classification: D24, L91, R41

The authors appreciate that various supports from the RIETI and are also grateful for helpful comments and suggestions by Professor Kyoji Fukao (Hitotsubashi University, RIETI-PD), Vice President Masayuki Morikawa (RIETI), President and CRO Masahisa Fujita (RIETI) and seminar participants at RIETI. This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Grant-in-Aid for Scientic Research (No. 17330052, 21330055, 30283378 and 22730183).

1. Introduction

Transport cost over distances is a major impediment of trade at any spatial scale, international or interregional. Reducing transport cost significantly benefits the economy in ways such as more firms selling their products in distant locations and consumers enjoying lower prices and greater variety. Understanding the structure of transport costs is essential for policy-making to design efficient transportation systems that contribute to reducing transport costs and thereby improve gains from trade.

There are several approaches to quantitative analysis of transport cost. The gravity model has been used to describe the pattern of trade flow in which volume of trade between countries decreases with distance, a proxy of transport cost. Anderson and Wincoop (2004) derived the gravity equation from the general equilibrium model of international trade, and proposed a method to measure the transport cost in terms of the ad valorem tax equivalent. Another approach is to use the data of the fob exporting price and cif importing price between the same trading partners, then the cif/fob ratio is taken as a measure of transport costs. Limao and Venables (2001) used the cif/fob ratio as the dependent variable of the regression to examine various determinants of transport cost, including infrastructure quality. These methods based on indirect information are developed mainly for international trade to cope with the data availability problem. At the interregional level (within the same country), Combes and Lafourcade (2005) developed a method to compute the generalized transport cost between regions. They combined geographical information system (GIS) data and various sources including traffic conditions, energy prices, technology, infrastructure, and the market structure of the transport industry. Based on a shift-share analysis of these components for road transport, they found that changes in the market structure (-21.8%) and technology (-10.9%) were the real engines of the decrease of transport costs for the 1978-1998 period in France. In contrast, infrastructure contributes 3.2% to the decrease of transport costs.

This paper empirically investigates the structure of transport costs for interregional trade by using microdata on freight charge. Note that the freight charges are determined through interaction in the transport market, where shippers demand and carriers (transport firms) supply transport services. Thus, freight charges paid by shippers should reflect the cost incurred by carriers. We focus on road transport, reflecting the fact that trucking has a dominant share in transporting goods between regions in Japan. In 2005, trucks transported

91.2% of overall domestic freight volume (sum of operating carriers and private trucks), while the second largest share was 7.8%, by coastal shipping. We develop a simple model of the trucking market and derive the freight charge equation. By estimating the parameters of this equation, we examine the effects of various factors on the level of freight charge. We use microdata from the 2005 Net Freight Flow Census (NFFC), in which information on freight charge and other variables for individual shipment are obtained. NFFC is drawn from stratified random samples of actual shipments, which are the best available data on interregional shipments. The data for other explanatory variables such as distance, toll payment, and wage are obtained from various sources. An advantage of our method is that our data represent the costs actually incurred by shippers or carriers, unlike those based on constructed data by Combes and Lafourcade (2005). We further examine the existence of economies of scale with respect to lot size (weight) and long-haul economies: transport cost per unit weight is decreasing with weight; transport cost per distance is decreasing with distance.

The rest of the paper is organized as follows. The next section presents the model of freight transportation. Section 3 specifies the equations for estimation, and section 4 describes the data for empirical analysis and presents the results of estimation. Section 5 concludes the paper.

2. Model

A trucking firm offers transport service between separated locations using capital (trucks), labor (drivers), and fuel as inputs. In practice, a single trucking firm takes orders for shipments with various sizes and origin/destination pairs (distance). The sum of these shipments for a given period of time becomes the output of the firm that is compatible with the standard definition in the model of production[1]. However, we consider the cost structure of each shipment[2]. More specifically, we formulate the cost function of the transport service by chartered truck, by which a transport firm uses a single truck exclusively to transport the

[1] In this context, there is a substantial body of literature on cost structure of motor carrier firms. Among them, Allen and Liu (1995) used firm-level data of motor carriers to examine the presence of scale economies in freight transportation. In contrast, we use the data for each shipment that provide useful information for the analysis of interregional transport cost structure.

[2] The relation between costs in firm level and each shipment is discussed in Appendix 1.

goods ordered by a single shipper[3].

The cost for each shipment is the sum of the expenditures for inputs and highway toll if it is used as follows

$$C_{ij} = r_i^L L_{ij} + r^K K_{ij} + r_i^X X_{ij} + r_{ij}^H H \tag{2.1}$$

where L_{ij}, K_{ij}, and X_{ij} are respectively the quantities of labor, capital, and fuel that are used to transport a good from region *i* to region *j*. *H* is the highway dummy taking *H*=1 when the truck uses highway, and *H*=0 otherwise. r_i^L, r^K, r_i^X, and r_{ij}^H are respectively the wage rate, capital rental rate, fuel price, and highway toll[4]. Labor input is measured in terms of time devoted by drivers, t_{ij}, which includes not only driving time but also time for loading and unloading, rest breaks, etc. The capital cost for each shipment is considered to be the opportunity cost of using a truck for the time required to complete the trip, so also measured in terms of time. Also note that the larger truck should be used to carry a larger lot size of cargo. We denote by q the lot size of shipment measured in weight, and then capital input is represented by $g(q)t_{ij}$, where $g(q)$ is an increasing function of q. It is observed that fuel consumption per distance depends on weight (lot size) q and speed s_{ij}, thus represented by the function $e(q, s_{ij})$ [5]. Highway toll depends on the distance and weight of the truck, and is written as $r_{ij}^H = r^H(q, d_{ij})$. Incorporating the assumptions above into (2.1), the cost function is written as follows,

$$C_{ij}(q, d_{ij}, t_{ij}) = r_i^L t_{ij} + r^K g(q) t_{ij} + r_i^X e(q, s_{ij}) d_{ij} + r^H(q, d_{ij}) H \tag{2.2}$$

In the above cost function, q, d_{ij}, t_{ij} are all considered as output variables. This implies that freight transportation is a bundle of multiple characteristics produced by the trucking firm. This is different from the conventional definition of output variables in transportation; i.e., the product of quantity and distance ($q_{ij} \cdot d_{ij}$ with our notations). Empirical analysis in the

---

[3] The other widely adopted type is the consolidated truck service in which a single truck carries cargo collected from several shippers.

[4] We assume that locations of the trucking firm and origin of trip are the same, so wage rate at the origin is applied. Firms may purchase fuel at any location along the route, so fuel prices should be given for the origin-destination pair. However, we assume that fuel price at the origin is applied, considering the difficulty of acquiring information concerning where trucks purchase fuel.

[5] $e(q, s)$ increases with weight q. On the other hand, the relation between fuel consumption and speed is U-shaped: $e(q, s)$ decreases (increases) with s at lower (higher) speed.

subsequent section examines whether the conventional definition is appropriate.

The price of a transport service, freight charge, is also defined for a bundle of characteristics as $P_{ij}(q, d_{ij}, t_{ij})$. We consider the market equilibrium in a similar manner to the hedonic theory developed by Rosen (1974), as follows. The market for freight transport is segmented by pairs of origin and destination. Suppose there are shippers in region *i* that demand the transport service, where the origin of transportation is the same as the shipper's location. Each shipper looks for the firm that undertakes the order of transportation every time it is required to transport a good of size q_{ij}, from *i* to *j*[6]. We assume there are a number of trucking firms willing to take the order as long as freight charge, $P_{ij}(q, d_{ij}, t_{ij})$ exceeds the cost, $C_{ij}(q, d_{ij}, t_{ij})$. The shipper solicits bids and awards the order to the lowest bidder. We assume that all trucking firms in market *ij* have the same production technology. The bid submitted by firm *n* is $C_{ij}(q, d_{ij}, t_{ij}) + \delta_{ij}^n + \upsilon_{ij}^n$, where δ_{ij}^n is the profit added over the cost and υ_{ij}^n is a random variable that reflects the attitude of the firm at the time of bidding. Each firm chooses δ_{ij}^n to maximize the expected value of profit, $R^n \delta_{ij}^n$, where R^n is the probability that firm *n* wins the bid. Note that R^n depends not only on the bid by firm *n* but also on those by its competitors, so the bidding competition is formulated as a game. In equilibrium, the following relation should hold.

$$P_{ij}(q, d_{ij}, t_{ij}) = C_{ij}(q, d_{ij}, t_{ij}) + \delta_{ij}^* \tag{2.3}$$

where $\delta_{ij}^* = \min_n \left\{ \delta_{ij}^n + \upsilon_{ij}^n \right\}$[7]. By using a similar but more general model, Holt (1979) showed that increasing the number of bidders decreases the equilibrium bid. Following this result, we expect that δ_{ij}^* is decreasing with the number of trucking firms in market *ij*. We allow a different degree of competition in the market for trucking transport since the number of trucking firms may vary by location[8]. In the empirical analysis, we use several proxy

---

[6] Distance d_{ij} is determined once origin *i* and destination *j* are given. On the other hand, t_{ij} may be variable for the same distance since trucks can deliver the cargo faster via highway, or increasing the number of drivers to save on break time, loading, and unloading. Shippers are also willing to pay a higher price for faster delivery. Thus, it is more appropriate to formulate the model in which t_{ij} is endogenously determined in market equilibrium. This issue is left for future research and discussed in section 5.

[7] With this formulation, perfect competition is a special case where $\delta_{ij} = 0$.

[8] Since the deregulation of entry and price-setting started in 1991, the number of trucking

variables to explain the variation of δ_{ij}^*.

3. Econometric Model and Methods

Based on the theory we developed in the previous section, we estimate the cost function of trucking firms using the Net Freight Flow Census data, detailed in the following section. We need to take into account that the data come from surveys of shippers, not trucking firms, which means that we must estimate cost function without input/output data of suppliers. In order to do this, we assume a certain relationship between the freight charge and its cost (2.3).

3.1 Regression specification

Remember that the cost of carrying cargo weighing *q* tons from region *i* to region *j* located at distance of d_{ij} km is decomposed into four components, drivers' wage, truck rent, fuel expenditure. and highway toll if it is used, as follows:

$$C_{ij}(q, d_{ij}, t_{ij}) = r_i^L t_{ij} + r^K g(q) t_{ij} + r_i^X e(q,s) d_{ij} + r^H (q, d_{ij}) H$$

Suppose that truck rent $g(q)$ depends linearly on the size of truck $w^T(q)$, or $g(q) = \alpha_1 + \alpha_2 w^T(q)$. Truck size (defined by category according to weight without cargo) is determined so that the truck accommodates the cargo of size *q*.[9] The fuel efficiency $e(q,s)$ of trucks is typically an increasing function of total truck weight $q + w^T(q)$, and a U-shaped function of speed *s*. We assume that one can drive at different but fixed speeds s^H on the highway and s^L on local roads, and thus

$$e(q,s) = \begin{cases} c^H \left(q + w^T(q)\right) & \text{highway} \\ c^L \left(q + w^T(q)\right) & \text{local road} \end{cases}$$

---

firms in Japan has increased consistently, with about 1.5 times more in 2004 than in 1990. The growth rate in the numbers of employees and truck drivers is relatively slower than that of trucking firms. This means that the scale of trucking firms is becoming smaller and the trucking industry is becoming more competitive. At the local level, however, sizes of markets vary widely depending on the level of economic activity in the regions of origin and destination and the distance between them.

[9] Details of the relation between lot size and truck size are given in section 4.

where c^H and c^L are the fuel consumption per weight for speeds at s^H and s^L, respectively, and $c^H < c^L$ is assumed.

Highway toll $r^H(q, d_{ij})$ depends on the truck size and the distance,

$$r^H(q, d_{ij}) = a + b\rho_1(w^T(q))\rho_2(d_{ij})d_{ij}$$

where $\rho_1(w^T(q))$ is the toll per distance applied for the truck category of $w^T(q)$ and $\rho_2(d_{ij})$ represents the discount factor for long-distance use of the highway.

We assume that the price is determined depending also on other factors $Z = (Z_1, \cdots, Z_6)$, as

$$P_{ij}(q, d_{ij}, t_{ij}) = C_{ij}(q, d_{ij}, t_{ij}) + \gamma' Z + \gamma_7 t_{ij}$$

$\gamma' Z$ includes the trucking firm's profit, represented by δ_{ij}^* in (2.3), other factors affecting the cost, and demand-side effects that come from shippers' preferences. These variables are described in Table 1. $Q_i\_sum/trucks(Z_3)$, num-truck-firms (Z_5) are proxy to the degree of competition, thereby the determinants of profit. intra-dummy (Z_1) is a dummy variable that takes the value one when it is the intraregional trade and zero otherwise. The variable border-dummy (Z_2), takes the value one when the two regions are contiguous and zero otherwise. These two dummy variables are included to capture some nonlinearity in terms of d_{ij}. The variable *imb* (Z_4) represents the trade imbalance calculated as $imb = Q_{ji} / Q_{ij}$, where Q_{ji} is the trade volume from region j to i and Q_{ij} is the trade volume from region i to j. If a truck carries goods on both directions of a return trip, then the firm is willing to accept a cheaper freight charge compared with the case in which the truck returns without cargo. iceberg (Z_6) is a proxy to the price of goods transported, which is included to examine if an iceberg-type cost applies in our data. As the demand-side factor, we include t_{ij} because it is generally more favorable for shippers if the goods (can) reach the destination earlier.

< insert Table 1. Variable Descriptions and Sources of Data here>

Allowing parameters η_i, $i = 1,2,3,4$, our empirical model turns out to be:

$$P_{ij}(q, d_{ij}, t_{ij}) = \eta_1 r_i^L t_{ij} + \eta_2 r^K \left\{\alpha_1 + \alpha_2 w^T(q)\right\} t_{ij} + \eta_3 r_i^X \left\{c^H H + c^L(1-H)\right\}(q + w^T(q))d_{ij}$$
$$+ \eta_4 r^H(q, d_{ij})H + \gamma_7 t_{ij} + \gamma' Z + \varepsilon$$

γ_7 is the parameter representing the preference of shippers and thus expected to be negative. $c^H H + c^L(1-H)$ in the term of fuel consumption is further rewritten as $c^L(1-\theta H)$, where $\theta = 1 - \frac{c^H}{c^L}$ is the ratio of saving fuel consumption from using the highway. We use empirical evidence concerning c_H / c_L. To this end, re-parameterizing the above equation, we have the final form of econometric model,

$$\begin{aligned} P_{ij}(q, d_{ij}, t_{ij}) = \beta_0 + \beta_1 r_i^L t_{ij} + \beta_2 t_{ij} + \beta_3 w^T(q) t_{ij} + \beta_4 r_i^X (1-\theta H)(q + w^T(q)) d_{ij} \\ + \beta_5 r^H(q, d_{ij}) H + \gamma' Z + \varepsilon \end{aligned} \tag{3.1}$$

and thus, the explanatory variables are

$$\{r_i^L t_{ij},\ t_{ij},\ w^T(q) t_{ij},\ r_i^X (1-\theta H)(q + w^T(q)) d_{ij},\ r^H(q, d_{ij}) H,\ Z\}.$$

We expect the following parameters sign,

$$\begin{aligned} &\beta_0 > 0 \\ &\beta_1 = \eta_1 > 0, \\ &\beta_2 = \eta_2 r^K \alpha_1 + \gamma_7 \\ &\beta_3 = \eta_2 \alpha_2 > 0, \\ &\beta_4 = \eta_3 c^L > 0, \\ &\beta_5 = \eta_4 > 0. \end{aligned}$$

On the sign of γ, we expect the following. When *imb* (Z_4) is large, the driver is likely to have freight on the way home and the price may be lower. The opportunity cost of an empty drive is also smaller for shorter trips. For this reason, γ_4 is expected to be negative. We include Q_i _*sum/trucks* (Z_3) and num-truck-firms (Z_5) in region *i* as proxies to competition in transportation market *ij*[10]. If Z_3 is large, there are not enough trucks in the region relative to the quantity of goods to be carried out of the region. Then, the competition should not be heavy and the price will be higher. Therefore γ_3 is expected to be positive. If Z_5 is large, we may regard that there are too many trucking firms, which results in heavy competition. Then, the price will be lower and γ_5 is expected to be negative. The iceberg hypothesis implies that the transport cost is positively correlated with value of the good, so the coefficient of iceberg (Z_6) should have a positive sign. Expected signs of coefficients discussed so far are summarized in Table 2.

---

[10] This is equivalent to assuming that competition takes place among trucking firms located in the same region as shippers.

< insert Table 2. Expected Signs of Coefficients here>

3.2 Endogeneity and 2SLS estimation

We can consider implementing OLS (ordinary least squared) estimation of eq.(3.1). There may, however, be endogeneity in some explanatory variables. We drop subscripts *i* or *ij* unless it is ambiguous. First, *t* can be endogenous because if there are no specific requests on the arrival time from the shipper, trucking firms can decide the efficient length of time spent for the freight. This is especially the case when the goods are consolidated. *H* can also be endogenous because the trucking firm can decide whether to use the highway depending on its own convenience. In such cases of endogenous regressors, OLS estimation does not provide consistent estimates.

A solution is to apply 2SLS (two-stage least squares) estimation using suitable instrumental variables. Valid instruments must have correlation with the endogenous regressors, but uncorrelated with the error terms. In the present context, we may pick *d* and the dummy variable of time-designated delivery D_T as its instruments. The shipper determines both the variables; thus, they are considered exogenous, but are correlated with *H*. We use *d* again as the instrument for *t*. It is likely that carriage time *t* depends on distance *d* between the home and destination, but *d* is exogenous for the trucking firm because it is determined by the order of the shippers. Thus, in the first stage, we run a probit estimation for dependent variable *H* regressing on d, D_T,

$$E(H \mid d, D_T) = P(\delta_0 + \delta_1 d + \delta_2 D_T \geq u \mid d, D_T) \tag{3.2}$$

where *u* is a standard normal variate. We implement OLS for t;

$$E(t \mid d) = \kappa_0 + \kappa_1 d \ . \tag{3.3}$$

Taking into account that *t* is likely to depend also on *H*, we may want to include *H* as an additional regressor to (3.3),

$$E(t \mid d, H) = \kappa_0 + \kappa_1 d + (\kappa_2 + \kappa_3 d) H \ .$$

However, as previously stated, *H* is also endogenous and thus it is not a suitable IV. Instead we can use predictor $\hat{H}$ from regression (3.2) as the regressor, or,

$$E(t \mid d, \hat{H}) = \kappa_0 + \kappa_1 d + (\kappa_2 + \kappa_3 d) \hat{H} \tag{3.4}$$

We obtain $\hat{H}$, the predicted values of H from (3.2), and $\hat{t}$, the predictor of t from either (3.3) or (3.4). Replace t and H in eq.(3.1) by $\hat{t}$ and $\hat{H}$ respectively, and we obtain second stage regression equation,

$$P_{ij}(q,d_{ij},t_{ij}) = \beta_0 + \beta_1 r_i^L \hat{t}_{ij} + \beta_2 \hat{t}_{ij} + \beta_3 w^T(q)\hat{t}_{ij} + \beta_4 r_i^X (1-\theta\hat{H})(q+w^T(q))d_{ij} + \beta_5 r^H(q,d_{ij})\hat{H} + \sum_{k=1}^{6}\gamma_k Z_k + \varepsilon. \quad (3.5)$$

Applying OLS estimation to (3.5), we obtain 2SLS estimates of β, γ that are consistent under endogeneity. (3.5) is slightly different from textbook 2SLS in the sense that some of the endogenous variables are multiplied by exogenous variables. We show that OLS of (3.5) works in Appendix 2.

4. Data and Empirical Results

We formulate an estimation model of the freight charge equation and explain the estimation strategies in the previous section. In this section, we first list the dependent variable and covariates from the 2005 Net Freight Flow Census (NFFC), National Integrated Transport Analysis System (NITAS), and other statistics. NFFC provides microdata on interregional shipments. NITAS is a system that the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) developed to compute transport distance, time, and cost between arbitrary locations. We adopt demand size and degree of competition of the transportation market to control regional heterogeneity by other statistics. Second, we show the data construction for our empirical study and then discuss the empirical results.

4.1 Data Description

In the previous section, we show the estimation model in eq. (3.5);

$$P_{ij}(q,d_{ij},t_{ij}) = \beta_0 + \beta_1 r_i^L \hat{t}_{ij} + \beta_2 \hat{t}_{ij} + \beta_3 w^T(q)\hat{t}_{ij} + \beta_4 r_i^X (1-\theta\hat{H})(q+w^T(q))d_{ij} + \beta_5 r^H(q,d_{ij})\hat{H} + \sum_{k=1}^{6}\gamma_k Z_k + \varepsilon.$$

The dependent variable is freight charges P_{ij} and the explanatory variables are

$$\{r_i^L t_{ij},\ t_{ij},\ w^T(q)t_{ij},\ r_i^X(1-\theta H)(q+w^T(q))d_{ij},\ r^H(q,d_{ij})H,\ Z\}$$

Z includes other explanatory variables, which can affect the price. Specifically, we use intra-dummy (Z_1), border-dummy (Z_2), $Q_i\_sum/trucks$ (Z_3) imb (Z_4), num-truck-firms (Z_5), and iceberg (Z_6). Table 1 provides the data sources to construct these variables.

We use the data from NFFC conducted by MLIT to obtain data on individual freight charge P_{ij}, lot size q, and transportation time t_{ij} that each shipment actually spent. We inform that t_{ij} might include times for loading and unloading of cargo, transshipment, driver's break, etc., which would vary widely with trucking firms and shipments.

The 2005 census uses 16,698 domestic establishment samples randomly selected from about 683,230 establishments engaged in the mining, manufacturing, wholesale trade, and warehousing industries. Each selected establishment reports shipments for a three-day period. This produces a total sample size of over 1,100,000 shipments, each of which has information on the origin and destination, P_{ij}, q, t_{ij}, the industrial code of the shipper and consignee, the code of commodity transported, main modes of transport, etc. We also collect data on transport distance d_{ij}, wage rate r_i^L, toll payments r^H, the number of trucking firms, number of trucks, etc. The data on transport distance d can be calculated by using NITAS from the information on the origin and destination for each shipment in NFFC. NITAS is a system that MLIT developed to compute the transport distance, time, and cost between arbitrary locations along the networks of transportation modes such as automobiles, railways, ships, and airlines. It searches for transportation routes according to various criteria, such as the shortest distance, shortest time, or least cost. We compute the transport distance between 2,052 municipalities as the distance between the jurisdictional offices along the road network with NITAS under the condition of minimizing the travel distance.

The driver's average wage per hour in the prefecture of origin r_i^L is calculated using the data on the monthly contractual cash earnings, scheduled hours worked, and overtime for drivers of small-middle-sized and large-sized trucks. These data are taken from the Basic Survey on Wage Structure by the Japan Institute for Labour Policy and Training. The general retail fuel r_i^X is the average diesel oil price as of October 2005 for prefecture of origin, which is published by the Oil Information Center. Truck size $w^T(q)$ is given by weight of a truck without cargo for categories according to lot size, as follows;

$$w^T(q) = \begin{cases} 2.356, & \text{if } q \le 2 \\ 2.652, & \text{if } 2 < q \le 3 \\ 2.979, & \text{if } 3 < q \le 4 \\ 3.543, & \text{if } 4 < q \le 5 \\ 5.533, & \text{if } 5 < q \le 12 \\ 7.59, & \text{if } 12 < q \le 14 \\ 8.765, & \text{if } 14 < q \end{cases}$$

We refer to Hino Motors' product specifications[11] to get $w^T(q)$. Highway toll $r^H(q,d)$ is from the East Nippon Express Company (E-NEXCO) and associated with each shipment's lot size and distance.

$$r^H(q,d) = \begin{cases} 0.84*(150+24.6*d)*1.05 \;\; if \;\; q < 2 \\ 0.84*(150+1.2*24.6*d)*1.05 \; if \;\; 2 \le q < 5, \\ 0.84*(150+1.65*24.6*d)*1.05 \;\; if \;\; q \ge 5 \end{cases}$$

0.84, 150 yen, and 1.05 are respectively the ETC or highway card discount, fixed cost, and consumer tax. Toll is 24.6 yen/km and there exists a vehicle type ratio (1.2, 1.65) that associates with the truck size $w^T(q)$ or q as below. While examining $r^H(q,d)$, we also reflect the tapering rate. If $100 < d \le 200$, we can get the discount rate 25% for distance exceeding 100 km, and if $d > 200$, a 25% discount for $100 < d \le 200$ and 30% discount for distance over 200 km are applied. There is a discount when the truck runs during the late night or early morning hours using ETC when there is a 30% or 50% discount. This is also considered in computing $r^H(q,d)$.

MLIT estimates the overall trade volume between prefectures based on shipment data from NFFC and publishes it via its website[12], and we use these data for Q_i, Q_{ji}, and Q_{ij} to construct the variables, $Q_i\_sum/trucks\,(Z_3)$ and imb (Z_4). We composed the num-truck-firms (Z_5) variable as 1,000 times the number of trucking firms per capita of prefecture of origin i. iceberg(Z_6) is defined by the monetary value (unit: yen) of annual

11 http://www.hino.co.jp/j/product/truck/index.html
12 http://www.mlit.go.jp/seisakutokatsu/census/census-top.html

shipments divided by its total volume (unit: tons) of annual shipments[13].

We would like to mention that definitions for region differ among the variables. t_{ij} and d_{ij} are municipality level data considering with both origin and destination regions, while r_i^L, r_i^X, r^H, $Q_i\_sum/trucks\,(Z_3)$, and num-truck-firms (Z_5) belong to prefectures of origin. $imb(\mathrm{Z}_4)$ is prefectural-level data made by origin and destination regions.

The descriptive statistics of these variables used in the estimation are summarized in Table 3[14].

< insert Table 3. Descriptive Statistics here>

In order to construct a target dataset for our analysis, first we abstract from the full dataset the data on the shipments that used trucks as the main mode of transport and then remove shipments with the following conditions: (1) Since this study focuses on the trucking industry, we exclude observations in regions inaccessible via a road network. Hokkaido, Okinawa, and other islands are excluded. (2) In order to observe the highway effects on P_{ij} clearly, we keep shipments that used only local roads or only highways. (3) We assume one truck and one driver are allocated for each shipment. We estimate that a large truck's maximum load capacity is less than 16 tons, which means if q is over 16 tons, carriers need multiple trucks. Thus, we removed the shipments for which q is over 16 tons. (4) We removed observations without freight charge P_{ij} data.

After abstracting our target dataset, 424,693 shipments and 8,155 shippers remain (full data set has 112,654 shipments and 16,698 shippers).

4.2 Estimation results

We estimated the econometric model eq. (3.5) using the data described in the previous section.

---

[13] These data are obtained from the NFFC annual survey of firms in manufacturing or wholesale industries. Thus, samples of shipments from the same firm should have the same value of $\mathrm{iceberg}(\mathrm{Z}_6)$

[14] Table 3 shows the descriptive statistics for chartered cargo. We also show a table comparing the descriptive statistics for chartered cargo with those for consolidated cargo in Appendix 3.

To implement estimation, we need to obtain a suitable value of θ to construct the explanatory variable $r_i^X(1-\theta\hat{H})(q+w^T(q))d_{ij}$. θ represents the fuel efficiency ratio of diesel trucks under two different speeds on highways and local roads. It is computed using the result by Oshiro, et al. (2001), who claim that

$$y(s) = 17.9/s - 9.6s + 0.073s^2 + 560.1$$

where $y(s)$ is fuel consumption efficiency (cc/km) and s is speed (km/hour). The weight is not controlled, but we can obtain an approximate ratio of $\theta = 1 - c_H / c_L$ assuming the efficiency ratio does not change with the weight of trucks. For example, supposing s^L =30(km/h) on local roads, the efficiency is $y(30)$ =338.4(cc/km). Similarly, when s^H =70 on highways, we have $y(70)$ =246.1. Combining the results, we obtain

$$\theta = 1 - \frac{c_H}{c_L} = 1 - \frac{(q+w^T(q))/e(q,s_H)}{(q+w^T(q))/e(q,s_L)} = 1 - \frac{e(q,s_L)}{e(q,s_H)} \approx 1 - \frac{246.1}{338.4} = 0.273$$

when the average speeds on highways and local roads are 70 km/h and 30 km/h, respectively. In Table 5, we report estimation results for $\theta = 0.2, 0.3, 0.4, 0.5$.

As suggested in section 3, we implemented both OLS and 2SLS estimation. Table 4 gives two kinds of estimates for all, chartered cargo and consolidated cargo observations with $\theta = 0.3$, which we think is the most reasonable value for θ. First we compare OLS and 2SLS regression shown in the table. Columns 2-7 give OLS estimation results, while columns 8-13 provide 2SLS estimates. In view of the estimation result of model 4, the coefficients of $r^L t$ and $r^H(w_T)H$ are not significant, which is obviously inappropriate. Those estimates for model 10 are all appropriate, including the signs of the parameters. We think that OLS estimation must be suffered from endogeneity bias. We believe that 2SLS is the suitable estimation method in the present model and data[15].

< insert Table 4. Estimation Results here>

Our main results are 2SLS estimation for chartered freight because there must be endogeneity

[15] We implemented 2SLS estimation for different sets of instruments based on the discussion in section 3, namely we take (3.3) and (3.4) in the first stage regression. The difference is that we use or do not use $\hat{H}$ in the first-stage estimation of t. In view of the estimates, we see the parameter estimates are not too different, and the significance of variables changes little. Therefore we report results only for (3.4). We also note that both regressors are significant in (3.4).

in some explanatory variables, as pointed out in section 3.2 and discussed above. We expect the sign of the estimates as stated in section 3, which is also tabulated in Table 2. The main estimation results are shown in Table 4, model 10. We obtain significant estimates with mostly right signs. The coefficient of labor input is significantly positive, as expected with $\beta_1 = 1.3696$. It is interesting that the level is between one and two. If only one driver carries goods all the time, the coefficient must be unity. But when they are carried for a long distance by, say, two drivers, one resting while the other drives, it will be two. If the data is a mixture of the two, it will take a value in [1,2]. We may also consider the case in which there is no cargo on the return trip. In this case, the trucking firm may like to charge the cost for two ways as well. β_2, the coefficient of time, is significantly negative. As discussed in section 3, the sign depends on two effects – one is related to the wage and truck rent, while the other is the shippers' preference; namely, they may be willing to pay more for faster delivery. There is a tradeoff between the two, with the former having a positive effect and the latter a negative effect on price P. We obtained the estimate of -3088.72 and, thus, we know that the latter dominates the former. β_3 is also the coefficient related to the truck rent. As the rent of larger trucks must be higher than for smaller ones, this coefficient is likely to be positive. β_4 is the coefficient of fuel consumption that is expected to be positive, and indeed it is. We cannot discuss its appropriate level since it depends on the mileage parameter of trucks. β_5 is the coefficient of highway toll, which is also significantly positive. As in the case of labor coefficient β_1, we expect this value to be in [1,2] because if the trucks do not have goods on their return trip, they may prefer to charge the shippers the highway toll for two ways. Indeed, the value is 1.2356, which lies in [1,2].

For additional variables of intra-dummy and border-dummy the coefficients are significantly negative. This may reflect that freight to very close places does not waste carriers' time for the return drive and thus the opportunity cost is lower. We also include the *imb* variable as the opportunity cost. *imb* is regarded as a proxy to the probability of obtaining a job on the return home. We expected that this has a negative impact on *P*, and this is right, but it turns out to be insignificant. We include $Q_i\_sum/trucks$ and num-truck-firms as proxies of freight industry competition. The coefficients are negative, as expected, but only the latter is significant. We can calculate the effect of an increase in the number of truck firms using this result. As shown in Table 3, the average number of trucking firms per 1,000 people is 0.420757. Because the standard deviation is 0.095, the change of 1 standard deviation from an average must be 0.095*5888=559(yen), noting that the coefficient is -5888. The area where the degree of competition is the highest is Ibaraki Prefecture, with the lowest in Nagano Prefecture. The difference of the degree of competition is 0.4082, which must be 0.4082*5888=2,404(yen) noting the maximum value of number of trucking firms per 1,000 people is 0.67458 and the minimum value is 0. 26638. Because the average freight charge is 26,737 yen, it is about 10% of the average of the freight charge. Though it is small, it is an effect that cannot be ignored. We include iceberg to examine whether the iceberg-type freight cost applies. The coefficient is positive as the iceberg hypothesis claims, but insignificant in our analysis. We conclude that this hypothesis does not hold in the Japanese truck freight industry.

We pick $\theta = 0.3$ as the default value based on the discussion at the beginning of this section. We examined the sensitivity by estimating the same model for different values of $\theta = 0.2, 0.3, 0.4, 0.5.$ Table 5 shows the results. The estimates are rather stable for all coefficients except those of $w^T t$ and $r^H(w_T)H$. The coefficient of $r^H(w_T)H$ becomes insignificant when $\theta = 0.2,$ while that of $w^T t$ remains significantly positive for all values of $\theta,$ but the level changes a great deal. One possible reason for this instability may be the means of construction of w^T. We construct w^T as stated in the previous section, but it should include noise that may not be ignorable. The present data does not in fact provide us with any information on what size of trucks are used for each service, and thus we cannot go further. A possible remedy is to use instruments for w^T in the estimation. We will pursue this direction in future research.

< insert Table 5. Estimation Results with Different θ here>

We estimated the model using the data of consolidated freight also, just for comparison. We do not believe our theoretical model suitably accommodates the case of consolidation because the cost structures must be different between the two services. We surmise that the trucking companies are likely to offer cheaper rates for consolidated service than chartered because the cost can be shared more efficiently among the shippers. However, we cannot confirm this conjecture straightforwardly comparing, say, estimates of models 10 and 12. We need to carefully construct the model of the freight price of consolidated freight service and estimate it.

NFFC classifies the shipments into nine groups by the variety of transported commodities; Agricultural and Fishery Products, Forest Products, Mineral Products, Metal and Machinery Products, Chemical Products, Light Industrial Products, Miscellaneous Manufacturing, Industrial Waste and Recycling Products, and Specialty Products. For example, high-valued and/or perishable commodities are expected to raise the cost of the trucking firm because they often require careful handling and/or faster transport service. We have already shown that the value of commodities does not affect the price of freight (see the coefficient of iceberg in model 10 of Table 4). In order to examine the commodity-specific effects on the freight charge, we also estimate the model for each commodity. Classification into groups and the detailed commodities in each group are described in Appendix 4. Table 6 provides the estimates for the eight categories. The levels and signs of the coefficients appear to be relatively appropriate for Metal and Machinery, Chemical Products, and Light Industrial Products, where sample sizes are significantly larger than for the others.

< insert Table 6. Commodity-wise Estimation Results here>

4.3 Scale economies and long-haul economies

Figures 1 and 2 plot elasticities of freight charge with respect to lot size *q* and distance *d*, which are calculated by the following formulas.

$$E_q(q,d) = \beta_4 \tilde{r}^X (1-\theta H) d \cdot q / \tilde{P}(q,d)$$

$$E_d(q,d) = \left[(\beta_1 \tilde{r}^L + \beta_2 + \beta_3 w^T(q))(\kappa_1 + H\kappa_3) + \beta_4 \tilde{r}^X (1-\theta H)(q + w^T(q)) + \beta_5 \frac{\partial r_H(w^T(q), d)}{\partial d} H \right] \frac{d}{\tilde{P}(q,d)}$$

where $\tilde{r}^X$ and $\tilde{r}^L$ are respectively the sample means of fuel price and wage rate shown in Table 3, $\tilde{P}(q,d)$ is obtained by substituting *q*, *d*, and sample means of other explanatory variables into (3.5).

Values of $E_q(q,d)$ and $E_d(q,d)$ provide the information on scale economies and long-haul economies: scale economies exist if $E_q(q,d) < 1$ and long-haul economies exist if $E_d(q,d) < 1$. The values shown in Figures 1 and 2 are significantly lower than 1, which indicates the existence of scale economies and long-haul economies in freight transportation.

$E_q(q,d)$ is increasing with q from 0.05 (at $q=1$ ton) to 0.45 (at $q=16$ ton), while $E_d(q,d)$ is increasing with *d* from 0.1 (at $d=50$ km) to 0.8 (at $d=800$ km). These results suggest that scale economies are stronger than long-haul economies.

As stated in footnote 1, the majority of existing studies on cost structure of motor carriers are based on firm-level data, and report that the motor carrier industry has a constant returns to scale technology. In contrast, our study shows the significant scale economies at the individual shipment level, which is important from the shippers' viewpoint. Note that freight charge per shipment is the real transport cost perceived by shippers, which they should take into account in making various decisions, such as choices of plant location and geographical extent of shipping destinations (i.e., market area). We find no literature on econometric estimation of long-haul economies in interregional transportation.

To obtain quantitative insights, we calculate the values of freight charge per ton-km for various combinations of *q* and *d*, as in Table 7. This calculation incorporates the effect of lot size through choice of truck size that is ignored in calculation of elasticities since marginal change in *q* does not affect $w^T(q)$. The table shows the results for two cases: using highways and local roads. Differences between the two cases contain the effects of several factors working in opposite directions, such as shippers' higher willingness to pay (+), trucking firms' cost savings from shorter transport time (-), and toll payment (+). In fact, the freight charges when using highways are higher if *q* and *d* are smaller, while the relations are reversed if *q* is larger. This may be attributed to the toll structure in which toll rate per weight decreases with truck size. In other words, highway use is advantageous for a larger cargo lot size. The table shows that variations in the unit freight charges for different combinations of *q* and *d* are quite large; e.g., from $\tilde{P}(1,50)=431.66$ (using highways) to $\tilde{P}(16,800)=19.14$ (using local roads). We also observe that the effects of changing lot size or distance vary depending on the level of *q* and *d*. Notwithstanding these results, it is somewhat surprising that the unit freight charges have similar values if the products of q and d, $q \cdot d$, are the same. For a fixed value of $q \cdot d = 800$, we have $\tilde{P}(2,400)=41.45$, $\tilde{P}(4,200)=40.04$, $\tilde{P}(8,100)=41.32$, $\tilde{P}(16,50)=39.02$. This suggests that the conventional definition of output, ton-km, turns out to be a good approximation.

5. Conclusion

This paper presents a microeconomic model of interregional freight transportation based on careful formulation of cost structure in trucking firms and market equilibrium, which takes into account the feature of transport services as a bundle of multiple characteristics. We estimate the parameters of the model using the microdata of interregional freight flows in Japan. Estimation results show that the determinants of transport cost incorporated in the model have significant effects in a manner consistent with theoretical predictions. The degree of competition also significantly affects the freight charge. Significant scale economies with respect to lot size and long-haul economies are shown to exist. Quantitative extents of these effects are also demonstrated.

We could extend the framework of empirical analysis in various directions in future research. First, time is a very important determinant of transport cost, as shown in the regression results. Shippers have an increasing willingness to pay for fast delivery, while trucking firms benefit from saving of opportunity costs of labor (drivers) and capital (trucks). It is widely recognized that transportation time savings account for the greatest part of the benefits from transport infrastructure improvement. Literature on estimating the value of transport time saving in freight transportation is relatively scarce compared with that on passenger transportation. It would be worth trying to develop a methodology to measure the value of time using microdata on freight charge. In this regard, we should note that transport time is an endogenous variable, which shippers and trucking firms choose for optimizing some objective. Second, this paper focuses on chartered truck service that has a relatively simple cost structure. We do not explicitly formulate the model of consolidated truck service, though it has a large share in interregional freight transportation. It is known that firms providing consolidated truck services adopt very complex production processes, such that they collect, consolidate, and distribute their shipments through networks consisting of terminals and breakbulk centers. Firms use advanced information and communication technologies, and construct their own infrastructure, such as terminals. Explicit modeling may be beyond the scope of our purpose, but a tractable framework that captures essential features of the service and is suitable for empirical analysis is needed. Third, there is an important research question regarding the widely observed fact that transport cost is decreasing over time. This may be explained by technological improvement and the increasing degree of competition due to deregulation. Which force is dominant? To address this question, we should develop methodology to define and measure productivity in transport sector, for which conventional methods such as total factor productivity (TFP) in the manufacturing sector are not applicable.

Finally, factor price changes or infrastructure improvement can significantly affect the behavior of agents as well as the equilibrium price of freight, which obviously affects social welfare. Structural estimation enables us to evaluate such effects, unlike simple regression estimation. We are planning to estimate the simultaneous equation system of freight price determination, time spent for delivery, and highway dummy, which have a complex relationship. Research in this direction is currently underway.

References

Allen, W. B. and Liu, D., 1995, "Service Quality and Motor Carrier Costs: An Empirical Analysis," *Review of Economics and Statistics*, 77, 499-510.

Anderson, J. and van Wincoop, E., 2004, "Trade Costs," *Journal of Economic Literature*, 42, 691-751.

Combes, P-P. and Lafourcade, M., "Transport Costs: Measures, Determinants, and Regional Policy Implications for France," *Journal of Economic Geography*, vol. 5, issue 3, 2005 pp. 319-349.

Holt, Jr., C., 1979, "Uncertainty and the Bidding for Incentive Contracts," *American Economic Review*, 69, 697-705.

Hummels, D., "Towards A Geography of Trade Costs," University of Chicago. mimeographed document, 2001.

Japan Trucking Association, *Truck yuso sanngyo no genjou to kadai - Heisei 19nen -*, Japan Trucking Association.

Limão, N. and Venables, A. J., "Infrastructure, Geographical Disadvantage, Transport Costs and Trade," *World Bank Economic Review*, 15, 2001, pp. 451-479.

Ministry of Land, Infrastructure, Transport and Tourism, *Land Transport Statistical Handbook*.

Ministry of Land, Infrastructure, Transport and Tourism (MLIT), Net Freight Flow Census. Statistics Bureau, Ministry of Internal Affairs and Communication. *Social Indicators By Prefecture*.

Oshiro, N., Matsushita, M., Namikawa, Y., and Ohnishi, H., 2001, "Fuel consumption and emission factors of carbon dioxide for motor vehicles," *Civil Engineering Journal* 43 (11), 50–55 (in Japanese).

Rosen, S., 1974, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure

Competition," *Journal of Political Economy*, 82, 34-55.

Table 1. Variable Descriptions and Sources of Data

Variable	Unit	Description	Source
P_{ij}	yen	Freight charge	Net Freight Flow Census (three-day survey)
r_i^L	yen/hour	Wage rate $r_i^L = \dfrac{\text{Monthly Contractual Cash Earnings}}{\text{Scheduled hours worked} + \text{over time}}$ * We use the data of Monthly Contractual Cash Earnings for small sized and medium sized truck driver if q<5, and those for large sized truck driver if q>5.	Basic Survey on Wage Structure, The Japan Institute for Labor Policy and on Training
t_{ij}	hours	Transportation time	Net Freight Flow Census (three-day survey)
w^T	tons	Vehicle weight $w^T(q) = \begin{cases} 2.356, & \text{if } q \le 2 \\ 2.652, & \text{if } 2 < q \le 3 \\ 2.979, & \text{if } 3 < q \le 4 \\ 3.543, & \text{if } 4 < q \le 5 \\ 5.533, & \text{if } 5 < q \le 12 \\ 7.59, & \text{if } 12 < q \le 14 \\ 8.765, & \text{if } 14 < q \end{cases}$	Hino Motors http://www.hino.co.jp/j/product/truck/index.html
r_i^X	yen	General retail fuel (diesel oil) price on October 2005	Monthly Survey, The Oil Information Center
q	tons	Lot size (disaggregated weight of individual) shipments	Net Freight Flow Census (three-day survey)
d_{ij}	Km	Transport distance between origin and destination	National Integrated Transport Analysis System (NITAS)
r^H		Highway toll $r_i^L = (\text{toll per 1km} \times \text{travel distance} \times \text{ratio for vehicle type} \times \text{tapering rate}+150) \times 1.05 \times \text{ETC discount}(=0.84)$ *toll per 1 km =24.6 yen/km *ratio for vehicle type $\Rightarrow$ 1.0 ($q \le 2$), 1.2 ($2 < q < 5$), 1.65 ($5 \le q$) *tapering rate $\Rightarrow$ 1 .0 if $d_{ij} \le 100$ $(100km \times 1.0 + (d_{ij} - 100km) \times (1-0.25)) / d_{ij}$ if $100 < d_{ij} \le 200$ $(100km \times 1.0 + 100km \times (1-0.25) + (d_{ij} - 200km) \times (1-0.30)) / d_{ij}$ if $200 < d_{ij}$	East Nippon Express Company (E-NEXCO)

Table 1. Variable Descriptions and Sources of Data

Variable	Unit	Description	Source
H		Dummy variable = 1 if highway is used; otherwise, 0	Net Freight Flow Census (three-day survey)
intra-dummy (Z_1)		Dummy variable = 1 if for intraregional trade; otherwise, 0	
border-dummy (Z_2)		Dummy variable = 1 if the trips between the two regions are contiguous; otherwise, 0	
$Q_i\_sum/trucks$ (Z_3)		$\frac{\text{Aggregated weight of Region i(origin)}}{\text{trucks}}$	Net Freight Flow Census (three-day survey) Policy Bureau, Ministry of Land, Infrastructure, Transport and Tourism
imb (Z_4)		Trade imbalances $imb=\frac{\text{Aggregated weight from Destination to Origin}}{\text{Aggregated weight from Origin to Destination}}$	Logistics Census, Ministry of Land, Infrastructure, Transport and Tourism http://www.mlit.go.jp/seisakutokatsu/census/8kai/syukei8.html
num-truck-firms (Z_5)	company per million people	Number of truck firms by prefecture Note: This is the number of general cargo vehicle operations if the main transport mode is charted and the number of special cargo vehicle operations if the main transport mode is consolidated service.	Policy Bureau, Ministry of Land, Infrastructure, Transport and Tourism
iceberg (Z_6)	million yen/ton	Proxy for properties of iceberg transport costs iceberg= $\frac{\text{The value of shipment of manufactruing industry \& wholesaler}}{\text{Estimated weight}}$	Net Freight Flow Census (annual survey)

Table 2. Expected Signs of Coefficients

Variable	Parameter	Expected Sign
$r_i^L t_{ij}$	β_1	+
t_{ij}	β_2	+/−
$w^T t_{ij}$	β_3	+
$r_i^X (1-\theta H)(q+w^T)d_{ij}$	β_4	+
$r^H (q, d_{ij})H$	β_5	+
intra-dummy (Z_1)	γ_1	−
border-dummy (Z_2)	γ_2	−
Q_i _$sum/trucks$ (Z_3) (Z_3)	γ_3	+
imb(Z_4)	γ_4	−
num-truck-firms (Z_5)	γ_5	−
iceberg (Z_6)	γ_6	0/+

Table 3. Descriptive Statistics

	Observation	Mean	Standard deviation	Minimum	Maximum
P_{ij}	83748	26737.21	38335.75	100	1974000
r_i^L	83807	1484.931	179.6792	1058.893	2102.116
t_{ij}	74381	5.155214	6.003164	0	240
w^T	83807	3.654444	1.675269	2.356	8.765
r_{ij}^X	83807	106.492	1.852045	103	115
q	83807	4.128685	4.034756	0.011	16
d_{ij}	83807	154.306	204.2325	0	1958.13
H	70096	0.3152962	0.464637	0	1
r^H	83807	2264.863	2781.883	79.38	29364.5
intra – dummy (Z_1)	83807	0.3864952	0.4869492	0	1
border-dummy(Z_2)	83807	0.2672211	0.4425114	0	1
$Q_i\_sum/trucks$ (Z_3)	83807	15.16853	4.451944	5.04197	64.7619
imb (Z_4)	83805	1.13686	3.099202	0.003106	274.077
num-truck_firms (Z_5)	83807	0.420757	0.095079	0.26638	0.67458
iceberg (Z_6)	67204	4.707244	135.2709	0.0000425	16000
$r_i^L t_{ij}$	74381	7718.265	9164.029	0	396732.5
$w^T t_{ij}$	74381	19.73064	26.98994	0	714.96
$r^H H$	70096	1107.907	2405.329	0	24343.7

Table 4. Estimation Results

Variables	OLS						2SLS					
	All		Chartered cargo		Consolidated cargo		All		Chartered cargo		Consolidated cargo	
	model1	model2	model3	model4	model5	model6	model7	model8	model9	model10	model11	model12
$r_i^L t_{ij}$	0.1569	0.2358	0.1125	0.2132	0.094	0.0809	0.3907	0.4942	1.4342	1.3696	0.1366	0.1869
	[5.49]***	[7.40]***	[0.68]	[1.22]	[9.65]***	[6.01]***	[9.97]***	[9.85]***	[7.13]***	[5.95]***	[10.34]***	[10.54]***
t_{ij}	-2536.0556	-2111.7866	-1148.8058	-1783.5209	-11301.7321	-11419.7068	-2840.7801	-2146.0242	-2010.2248	-3088.7248	4134.9585	4709.3272
	[-39.63]***	[-31.10]***	[-5.11]***	[-7.51]***	[-10.52]***	[-9.69]***	[-33.63]***	[-17.25]***	[-6.83]***	[-9.42]***	[6.57]***	[7.01]***
$w^T(q)t_{ij}$	676.7453	600.0916	359.7132	455.9242	4710.0306	4759.4979	-196.3241	-420.9516	314.6108	223.4514	-4920.8122	-5243.7421
	[28.89]***	[21.87]***	[14.11]***	[14.76]***	[10.34]***	[9.52]***	[-4.40]***	[-8.11]***	[5.60]***	[3.37]***	[-17.22]***	[-17.12]***
$r_i^x(1-\theta H)(q+w^T(q))d_{ij}$	0.7306	0.8296	2.9449	2.6334	-0.1038	-0.0548	0.1695	0.1689	0.0977	0.1002	0.5395	0.5526
	[17.26]***	[17.45]***	[27.06]***	[21.86]***	[-5.26]***	[-2.56]**	[43.75]***	[36.70]***	[18.39]***	[15.91]***	[81.03]***	[72.90]***
$r^H(q,d_{ij})H$	0.0765	0.0706	0.0757	0.0688	0.0343	0.04	2.8277	2.4949	-1.1421	1.2356	6.3681	6.3889
	[45.44]***	[34.50]***	[40.79]***	[29.70]***	[22.01]***	[17.46]***	[18.18]***	[10.27]***	[-2.76]***	[2.34]**	[69.37]***	[61.17]***
intra-dummy		-2665.6454		-7354.314		767.2074		-2683.2158		-6040.096		248.8293
		[-10.99]***		[-16.16]***		[4.89]***		[-10.43]***		[-12.06]***		[2.53]**
border-dummy		1262.1755		-2775.4975		1152.8691		-195.5927		-1928.8795		706.9904
		[6.20]***		[-6.89]***		[8.55]***		[-1.15]		[-5.19]***		[8.99]***
$Q_i\_sum/trucks$		86.8057		71.5525		31.5204		80.5111		-5.4344		56.0054
		[7.83]***		[2.86]***		[5.59]***		[12.56]***		[-0.23]		[17.07]***
imb		-40.8524		-104.4216		-6.6512		-12.2542		-41.4546		-1.8139
		[-4.26]***		[-1.01]		[-1.56]		[-2.42]**		[-0.80]		[-0.60]
num-truck-firms		26345.0031		-1739.0477		-112714.3285		24136.4124		-5892.8665		-196343.6211
		[44.44]***		[-1.47]		[-4.87]***		[60.05]***		[-4.66]***		[-17.31]***
iceberg		-0.1791		1.8205		-0.2154		-0.124		0.9196		0.0175
		[-1.26]		[0.95]		[-4.72]***		[-1.37]		[0.83]		[0.26]
Constant	12768.3942	2879.6988	12595.3986	18101.7652	2496.5544	1968.1952	13810.8297	9066.0948	8872.7262	19841.6146	19061.3234	18884.3862
	[100.30]***	[9.04]***	[81.48]***	[20.57]***	[24.31]***	[11.49]***	[61.03]***	[19.25]***	[19.54]***	[17.47]***	[93.07]***	[75.78]***
Adj-R	0.5239	0.5489	0.5015	0.5079	0.1233	0.1321	0.4882	0.5096	0.4503	0.4449	0.387	0.3925
Obs	136756	104471	64866	51602	71890	52869	267464	204138	83807	67204	183657	136934

Table 5. Estimation Results with Different θ

Variables	$\theta = 0.2$	$\theta = 0.3$	$\theta = 0.4$	$\theta = 0.5$
$r_i^L t_{ij}$	1.3976 [6.05]***	1.3696 [5.95]***	1.3302 [5.82]***	1.2765 [5.62]***
t_{ij}	-2947.0307 [-9.00]***	-3088.725 [-9.42]***	-3280.766 [-9.96]***	-3533.381 [-10.64]***
$w^T(q)t_{ij}$	301.2809 [4.72]***	223.4513 [3.37]***	143.1586 [2.07]**	65.5954 [0.90]
$r_i^X(1-\theta H)(q+w^T(q))d_{ij}$	0.0887 [14.93]***	0.1002 [15.91]***	0.113 [16.84]***	0.1267 [17.62]***
$r^H(q,d_{ij})H$	0.343 [0.63]	1.2356 [2.34]**	2.373 [4.61]***	3.7885 [7.41]***
intra-dummy	-6148.4718 [-12.17]***	-6040.096 [-12.06]***	-5942.274 [-11.99]***	-5868.238 [-12.00]***
border-dummy	-2004.2914 [-5.35]***	-1928.88 [-5.19]***	-1862.516 [-5.07]***	-1815.141 [-5.00]***
$\frac{Q_i\_sum}{trucks}$	-6.0736 [-0.25]	-5.4344 [-0.23]	-4.9084 [-0.20]	-4.5952 [-0.19]
imb	-41.3682 [-0.79]	-41.4546 [-0.80]	-41.3512 [-0.81]	-40.9696 [-0.81]
num-truck-firms	-5894.8293 [-4.66]***	-5892.867 [-4.66]***	-5884.539 [-4.67]***	-5867.273 [-4.66]***
iceberg	0.9121 [0.81]	0.9196 [0.83]	0.9267 [0.84]	0.9326 [0.85]
Constant	19021.4032 [16.79]***	19841.615 [17.47]***	20834.193 [18.27]***	22007.239 [19.16]***
Adj-R	0.4439	0.4449	0.4461	0.4473
Obs	67204	67204	67204	67204

* p<0.1, ** p<0.05, *** p<0.01

Table 6. Commodity-wise Estimation Results

Variables	Agricultural and Fisheries	Forest Products	Mineral Products	Metal and Machinery	Chemical Products	Light Industrial Products	Miscellaneous Manufacturing	Industrial Waste and Recycling
$r_i^L t_{ij}$	1.6496 [1.59]	0.162 [0.29]	-0.8648 [-0.36]	-0.4482 [-1.14]	1.6633 [3.20]***	1.5568 [7.11]***	1.3882 [3.73]***	6.3472 [1.93]*
t_{ij}	-2561.0204 [-1.79]*	243.0995 [0.23]	-1096.7526 [-0.31]	-419.4358 [-0.83]	-3751.921 [-4.35]***	-3779.0085 [-9.63]***	-2339.1046 [-5.07]***	-8126.1005 [-2.10]**
$w^T(q)t_{ij}$	-173.8142 [-0.51]	-237.7669 [-1.93]*	-771.304 [-1.98]**	659.1312 [4.56]***	-209.4111 [-1.14]	521.6024 [8.58]***	353.2259 [3.50]***	1701.7024 [2.17]**
$r_i^X(1-\theta H)(q+w^T(q))d_{ij}$	0.1705 [4.56]***	0.0712 [8.01]***	0.1324 [6.44]***	0.067 [4.98]***	0.1691 [10.53]***	0.0588 [10.61]***	0.0643 [7.75]***	-0.2029 [-2.24]**
$r^H(q,d_{ij})H$	-2.0151 [-1.61]	-1.593 [-0.13]	9.5851 [2.89]***	1.2463 [1.51]	1.7528 [1.25]	1.3704 [2.37]**	-1.5765 [-1.75]*	17.7173 [2.58]**
intra-dummy	-8232.0333 [-4.90]***	-5122.1026 [-2.05]**	-6125.6849 [-2.04]**	-6529.3893 [-7.22]***	-2119.4816 [-1.93]*	-3650.417 [-5.70]***	-13080.2737 [-11.85]***	7464.9411 [0.98]
border-dummy	-5029.6144 [-3.29]***	5140.8904 [2.07]**	3440.0146 [1.39]	-2279.0405 [-3.35]***	1254.581 [1.76]*	-302.3857 [-0.54]	-7503.931 [-7.62]***	17916.8437 [2.63]***
$\frac{Q_i\_sum}{trucks}$	1.657 [0.01]	-362.9136 [-4.81]***	-237.2564 [-1.05]	20.7556 [0.61]	-263.8902 [-4.00]***	64.8614 [2.10]**	770.605 [9.58]***	704.3764 [0.91]
imb	-474.0811 [-1.04]	-180.2589 [-2.30]**	-367.6703 [-2.88]***	-1.5696 [-0.03]	-117.7831 [-1.10]	-192.5421 [-1.88]*	-523.1289 [-1.54]	1667.9274 [1.43]
num-truck-firms	950.372 [0.25]	15661.2271 [1.61]	-772.8903 [-0.10]	-2831.8592 [-1.37]	-22356.248 [-6.92]***	7515.3657 [4.57]***	-1898.6012 [-0.63]	6509.1668 [0.44]
iceberg	343.6572 [3.03]***	31.7562 [1.76]*	7491.0812 [3.01]***	0.1851 [0.17]	-9.637 [-2.40]**	-7.2245 [-0.76]	-66.2656 [-1.42]	-2.4393 [-3.16]***
Constant	16114.02 [4.09]***	15485.3189 [3.64]***	25134.0726 [3.36]***	17777.1192 [8.80]***	28794.1274 [8.62]***	9611.4704 [5.68]***	10388.154 [3.99]***	-12010.2366 [-0.77]
Adj-R	0.6088	0.7666	0.6911	0.4672	0.3562	0.6636	0.5778	0.2832
Obs	1894	352	195	24444	17776	13524	6325	468

* p<0.1, ** p<0.05, *** p<0.01

Table 7. Values of Freight Charge per ton km

		50 km	100 km	200 km	400 km	800 km
1t	Local Road	284.2518	153.986	88.85312	56.28667	40.00344
	Highway	431.6629	230.0651	127.2556	75.44867	49.54521
2t	Local Road	147.4611	82.32825	49.76181	33.47858	25.33697
	Highway	219.5661	118.7672	67.36246	41.45901	28.50728
4t	Local Road	83.48051	49.93575	33.16337	24.77718	20.58408
	Highway	116.0308	65.77632	40.04586	27.05998	20.56705
8t	Local Road	56.12454	37.34685	27.95801	23.26359	20.91638
	Highway	66.15758	41.32738	28.49758	21.99973	18.75081
16t	Local Road	39.12315	28.46548	23.13664	20.47223	19.14002
	Highway	39.02973	26.80259	20.68901	17.27973	15.67876

Figure 1. Elasticity of freight charge with respect to lot size (q)

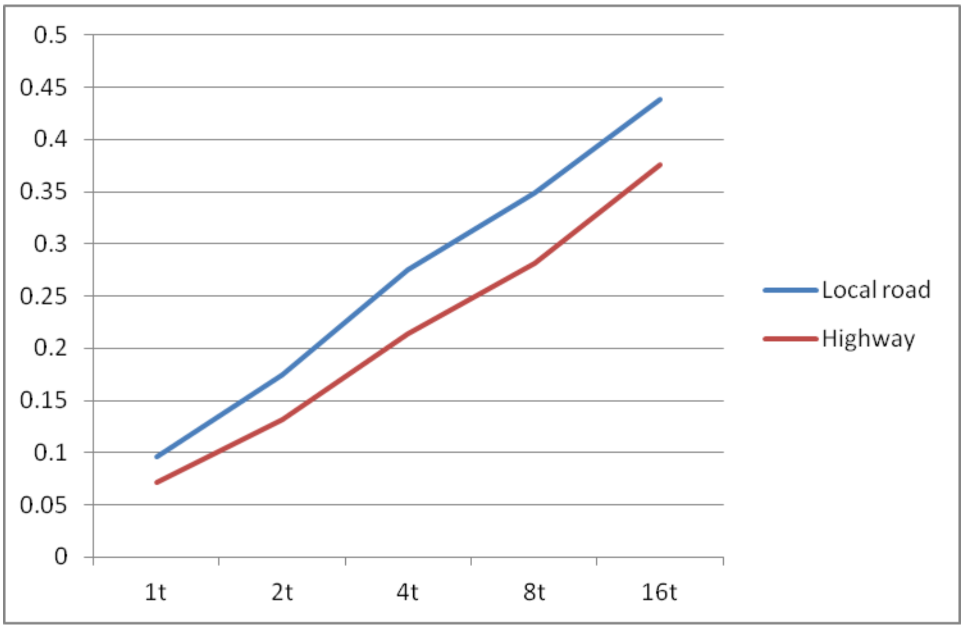
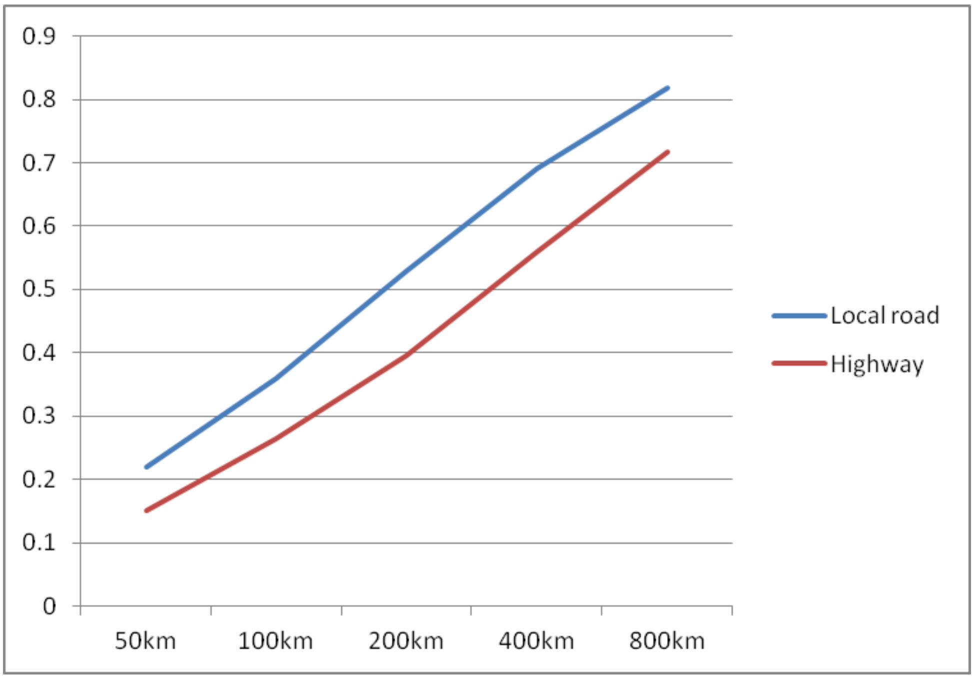


Figure 2. Elasticity of freight charge with respect to distance (d)



Appendix 1. Relation between costs in firm level and shipment level

The trucking firm takes orders of transporting cargo for various O-D pairs. We denote the number of shipments (orders) from i to j by m_{ij}. The total profit of the firm for a given period of time is written as follows.

$$\sum_{i,j} m_{ij}(P_{ij} - r_i^X X_{ij} - r^H(q,d_{ij})H) - r_i^L \overline{L} - r^K g(q)\overline{K} \tag{A1}$$

where $\overline{L}$ and $\overline{K}$ are labor (drivers) and capital (trucks) employed by the firm. In the short run, $\overline{L}$ and $\overline{K}$ are fixed, so the following constraints should hold

$$\sum_{i,j} m_{ij} t_{ij} \le \overline{L} \tag{A2}$$

$$\sum_{i,j} m_{ij} t_{ij} \le \overline{K} \tag{A3}$$

$\overline{L}$ and $\overline{K}$ are measured in terms of total time that the drivers and trucks in the firm can serve. If the firm takes an order, drivers and trucks are used during the trip, and their availability to serve other orders is restricted. In the short run, the trucking firm controls only m_{ij} to maximize (A1) subject to (A2)(A3). The optimality conditions are

$$P_{ij} - r_i^X X_{ij} - r^H(q,d_{ij})H - \lambda^L t_{ij} - \lambda^K t_{ij} = 0 \tag{A4}$$

where λ^L and λ^K are Lagrange multipliers associated with constraints (A2) and (A3), respectively. Thus $\lambda^L t_{ij}$ and $\lambda^K t_{ij}$ in (A4) are interpreted as opportunity costs of drivers and trucks. In the long run where $\overline{L}, \overline{K}$ are variables, optimal choices of the firm are described as follows

$$-r_i^L + \lambda^L = 0$$

$$-r^K g(q) + \lambda^K = 0$$

Putting the above equations into (A4) yields

$$P_{ij} = r_i^L t_{ij} + r^K g(q) t_{ij} + r_i^X X_{ij} + r^H(q,d_{ij})H .$$

The right-hand side of the above expression is equivalent to the cost of a shipment in Eq. (2.2).

Appendix 2

Consider the following endogenous regression model.

$$y_1 = \beta_0 + \alpha_1 x_1 y_2 + \beta' x + \varepsilon$$

where y_1, y_2 are endogenous and x_1, x are exogenous variables. OLS regression does not provide us with consistent estimates because $x_1 y_2$ is generally an endogenous variable. Supposing *z* is a valid instrument for y_2, or it satisfies

$$E(\varepsilon z) = 0,\ Cov(y_2, z) \neq 0,$$

then letting $\hat{y}_2 = \hat{\gamma}_0 + \hat{\gamma}_1 z$ be the OLS predictor of y_2 given *z*, $x_1 \hat{y}_2$ is a valid instrument for $x_1 y_2$.

This is the sketch of the proof. It suffices to show that

$$\frac{1}{n}\sum_{i=1}^{n} x_{1i}\hat{y}_{2i}\varepsilon_i \xrightarrow{p} 0.$$

Now

$$\frac{1}{n}\sum_{i=1}^{n} x_{1i}\hat{y}_{2i}\varepsilon_i = \frac{1}{n}\sum_{i=1}^{n} x_{1i}(\hat{\gamma}_0 + \hat{\gamma}_1 z_i)\varepsilon_i = \frac{\hat{\gamma}_0}{n}\sum_{i=1}^{n} x_{1i}\varepsilon_i + \frac{\hat{\gamma}_1}{n}\sum_{i=1}^{n} x_{1i} z_i \varepsilon_i .$$

Because $\hat{\gamma}_0 \xrightarrow{p} \gamma_0$, $\hat{\gamma}_1 \xrightarrow{p} \gamma_1$, and $\frac{1}{n}\sum_{i=1}^{n} x_{1i}\varepsilon_i \xrightarrow{p} 0$, $\frac{1}{n}\sum_{i=1}^{n} x_{1i} z_i \varepsilon_i \xrightarrow{p} 0$ by the exogeneity of (x_{1i}, z_i), we have the desired result.

Appendix 3. Descriptive Statistics

Variable	Observation		Mean		Standard deviation		Minimum		Maximum	
	Chartered cargo	Consolidated cargo	Chartered cargo	Consolidated cargo	Chartered cargo	Consolidated cargo	Chartered cargo	Consolidated cargo	Chartered cargo	Consolidated cargo
P_{ij}	83748	183650	26737.21	3961.465	38335.75	9248.058	100	100	1974000	460000
r_i^L	83807	183657	1484.931	1407.667	179.6792	141.4918	1058.893	1058.893	2102.116	1683.408
t_{ij}	74381	97040	5.155214	20.45417	6.003164	7.64649	0	10	240	90
w^T	83807	183657	3.654444	2.356641	1.675269	0.0159311	2.356	2.356	8.765	3.543
r_i^X	83807	183657	106.492	106.8483	1.852045	1.948149	103	103	115	115
q	83807	183657	4.128685	0.1288279	4.034756	0.2689434	0.011	0.001	16	4.32
d_{ij}	83807	183657	154.306	332.6557	204.2325	295.0032	0	0	1958.13	2074.33
H	70096	108220	0.3152962	0.5461467	0.464637	0.4978682	0	0	1	1
r^H	83807	183657	2264.863	3486.734	2781.883	2774.608	79.38	79.38	29364.5	20015.1
intra – dummy (Z_1)	83807	183657	0.3864952	0.1197776	0.4869492	0.3247022	0	0	1	1
border-dummy(Z_2)	83807	183657	0.2672211	0.202203	0.4425114	0.4016439	0	0	1	1
$Q_i\_sum/trucks$ (Z_3)	83807	183657	15.16853	14.27095	4.451944	5.368521	5.04197	5.04197	64.7619	64.7619
imb (Z_4)	83805	183576	1.13686	1.458757	3.099202	6.633833	0.003106	0.003106	274.077	322
num-truck_firms (Z_5)	83807	182248	0.420757	0.002232	0.095079	0.001567	0.2663796	0.0007036	0.67458	0.00997
iceberg (Z_6)	67204	136967	4.707244	35.74154	135.2709	407.0024	0.0000425	0.000019	16000	36475
$r_i^L t_{ij}$	74381	97040	7718.265	28736.62	9164.029	10581.01	0	10588.93	396732.5	131788.7
$w^T t_{ij}$	74381	97040	19.73064	48.20456	26.98994	18.03493	0	23.56	714.96	212.04
$r^H H$	70096	108220	1107.907	2167.824	2405.329	2811.127	0	0	24343.7	19097.9

Appendix 4. Classification and Commodity

Classification	Commodity
Agricultural and Fishery Products	Wheat
	Rice
	Miscellaneous grains, beans
	Fruits and vegetables
	Wool
	Other livestock products
	Fishery products
	Cotton
	Other agricultural products
Forest Products	Raw wood
	Lumber
	Firewood and charcoal
	Resin
	Other forest products
Mineral Products	Coal
	Iron ore
	Other metallic ore
	Gravel, sand, stone
	Limestone
	Crude petroleum and natural gas
	Rock phosphate
	Industrial salt
	Other non-metallic minerals
Metal and Machinery Products	Iron and steel
	Non-ferrous metals
	Fabricated metals products
	Industry machinery products
	Electrical machinery products
	Motor vehicles
	Motor vehicle parts
	Other transport equipment
	Precision instruments products
	Other machinery products
Light Industrial Products	Pulp
	Paper
	Spun yarn
	Woven fabrics
	Sugar
	Other food preparation
	Beverages

Appendix 4. Classification and Commodity

Classification	Commodity
Chemical Products	Cement
	Ready-mixed concrete
	Cement products
	Glass and glass products
	Ceramic wares
	Other ceramics products
	Fuel oil
	Gasoline
	Other petroleum
	Liquefied natural gas and liquefied petroleum gas
	Other petroleum products
	Coal coke
	Other coal products
	Chemicals
	Fertilizers
	Dyes, pigments, and paints
	Synthetic resins
	Animal and vegetables oil, fat
	Other chemical products
Miscellaneous Manufacturing	Books, printed matters and records
	Toys
	Apparel and apparel accessories
	Stationery, sporting goods and indoor games
	Furniture accessory
	Other daily necessities
	Wood products
	Rubber products
	Other miscellaneous articles
Industrial Waste and Recycling Products	Discarded automobiles
	Waste household electrical and electronic equipment
	Scrap metal
	Steel waste containers and packaging
	Used glass bottles
	Other waste containers and packaging
	Waste paper
	Waste plastics
	Cinders
	Sludge
	Slag
	Soot
	Other industrial waste