

# Transport costs, distance, and time : evidence from the Japanese Census of Logistics

著者	Tanaka Kiyoyasu
権利	Copyrights 日本貿易振興機構 (ジェトロ) アジア 経済研究所 / Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) <a href="http://www.ide.go.jp">http://www.ide.go.jp</a>
journal or publication title	IDE Discussion Paper
volume	241
year	2010-06-01
URL	<a href="http://hdl.handle.net/2344/903">http://hdl.handle.net/2344/903</a>

IDE Discussion Papers are preliminary materials circulated  
to stimulate discussions and critical comments

## IDE DISCUSSION PAPER No. 241

### **Transport Costs, Distance, and Time: Evidence from the Japanese Census of Logistics**

Kiyoyasu Tanaka\*

June 2010

#### **Abstract**

Geographic distance is a standard proxy for transport costs under the simple assumption that freight fees increase monotonically over space. Using the Japanese Census of Logistics, this paper examines the extent to which transport distance and time affect freight costs across shipping modes, commodity groups, and prefecture pairs. The results show substantial heterogeneity in transport costs and time across shipping modes. Consistent with an iceberg formulation of transport costs, distance has a significantly positive effect on freight costs by air transportation. However, I find the puzzling results that business enterprises are likely to pay more for short-distance shipments by truck, ship, and railroad transportation. As a plausible explanation, I discuss aggregation bias arising from freight-specific premiums for timely, frequent, and small-batch shipments.

**Keywords:** Transport cost, Transport time, Distance, Logistics, Selection bias

**JEL classification:** R30, R40, R41

---

\* Kiyoyasu Tanaka. Address: Economic Integration Studies Group, Inter-Disciplinary Studies Center, Institute of Developing Economies, 3-2-2 Wakaba, Mihama-ku, Chiba-shi, Chiba 261-8545 Japan. Phone: 81-43-299-9755; Fax: 81-43-299-9763.

E-mail: kiyoyasu\_tanaka@ide.go.jp

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

---

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by the Institute of Developing Economies of any of the views expressed within.

---

**INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO**  
**3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI**  
**CHIBA 261-8545, JAPAN**

©2010 by Institute of Developing Economies, JETRO

No part of this publication may be reproduced without the prior permission of the IDE-JETRO.

**Transport Costs, Distance, and Time:  
Evidence from the Japanese Census of Logistics**

**Kiyoyasu Tanaka**

Inter-disciplinary Studies Center  
Institute of Developing Economies  
Address: 3-2-2 Wakaba, Mihama-ku, Chiba-shi, Chiba, Japan, 261-8545  
Phone: 81-43-299-9755  
Fax: 81-43-299-9763  
Email: kiyoyasu\_tanaka@ide.go.jp

June 2010

**Abstract**

Geographic distance is a standard proxy for transport costs under the simple assumption that freight fees increase monotonically over space. Using the Japanese Census of Logistics, this paper examines the extent to which transport distance and time affect freight costs across shipping modes, commodity groups, and prefecture pairs. The results show substantial heterogeneity in transport costs and time across shipping modes, pointing to a trade-off between costs and time in transportation. Consistent with an iceberg formulation of transport costs, distance has a significantly positive effect on freight costs by air transportation after controlling for a wide range of unobserved fixed effects and sample selection bias. However, I find the puzzling results that business enterprises are likely to pay more for short-distance shipments by truck, ship, and railroad transportation. As a plausible explanation, I discuss aggregation bias arising from freight-specific premiums for timely, frequent, and small-batch shipments.

*Keywords:* Transport cost, Transport time, Distance, Logistics, Selection bias

*JEL classification:* R30, R40, R41

## 1. Introduction

Transport costs play a central role in shaping the pattern and location of economic activity. As the location of production is determined to minimize transport costs, the cost of moving products generates dispersion forces against industrial agglomeration (Fujita et al., 1999). In the international market, transportation costs create a barrier to an international transaction in goods and services, which in turn shapes the pattern of international trade (Anderson and van Wincoop, 2003). The costs of exporting motivate the firm to locate a production plant offshore for the savings of transport costs, which lead to the formation of multinational firms (Brainard, 1997; Markusen, 2002). Further, transport costs could even play an important role in economic growth (Gallup et al. 1999).

Despite the significance of transport costs, data limitations on direct measures of transport fees have posed a major challenge for empirically investigating the role of transportation in economic activities. Because it is difficult to observe an *ex ante* freight cost faced by agents in making economic decisions, empirical research has extensively relied on the geographic distance as a proxy for the size of transport costs. For example, the great-circle distance between capital cities in countries is used to measure transport costs between countries. This approach is justified by the simple assumption that the distance monotonically increases freight costs over space. A strong negative effect of the distance is commonly interpreted as suggesting that transport costs significantly discourage foreign trade (Disdier and Head, 2008).

However, the geographic distance captures not only transport costs over the geographic space but also other economic costs, including an acquisition of distant market information, communication with distant agents, and different preferences over goods (Anderson and van Wincoop, 2004). Fundamentally, a market rate for transport costs faced by producers would be determined by an interaction between demand for and supply of transport

services. In the market economy, the transport costs in equilibrium could further differ along many dimensions such as product characteristics, shipping mode, and the origin and destination markets. Because the geographic distance is a crude measure of actual transport fees that producers are willing to pay, the question remains as to what extent the geographic distance can explain heterogeneity in transportation costs.

The purpose of this paper is to investigate the role of distance in determining transport costs using extremely detailed data on domestic transport costs from the *Census of Logistics* (CL) in Japan. The CL data provides rich information on transportation costs paid by business enterprises in mining, manufacturing, wholesale, and warehouse industries.<sup>1</sup> I exploit data on freight costs, shipping time, and the volume of commodity flows, which are disaggregated by 8 major product categories, 11 transportation modes, and 47 prefecture pairs.

Exploiting the rich characteristics of transportation in Japan, I illustrate the pattern of transport costs in various categories. As expected, there is a wide dispersion of freight costs across regions, commodity groups, and transport modes. In particular, the average cost of air transportation is substantially higher than other transportation modes, pointing to the fact that freight costs differ most significantly by shipping mode. The average cost of freight fees appears to increase with respect to a dispersion of freight fees. By contrast, shipping time by air transportation is shorter than other transport modes, reflecting a trade-off between freight costs and time.

To explore the role of distance in transport costs, I estimate a transport cost function based on the iceberg formulation of transport costs in new economic geography. To isolate the distance effect from other factors, I control for transport time and a wide range of fixed effects in sending and arriving prefectures, commodity groups, transport modes, and year. Robust to a

---

<sup>1</sup> See Hummels (2007) for data sources on international transportation costs.

variety of alternative specifications, I find that transport fees are *negatively*, not positively, correlated with the distance. As the result may be driven by heterogeneity in transport mode and sample selection bias, I also employ a Heckman two-step estimation for samples by air, railroad, ship, and truck. Consistent with intuition, air transport fees increase with distance at a decreasing rate. However, distance continues to exhibit a significantly *negative* coefficient in other transport modes. To resolve this puzzling result, I discuss the plausible hypothesis that the negative distance effect may be driven by unobserved factors, including a premium for timely, frequent, and small-batch shipments in modern transportation systems.

The rest of this paper is organized as follows. Section 2 describes the Japanese Census of Logistics, followed in section 3 by a description of transport costs and time aggregated over regions, commodity groups, and shipping mode. Section 4 explains iceberg transport costs in new economic geography to derive a transport cost function for empirical specification. Section 5 presents the estimation results. Section 6 discusses a possible interpretation of the estimation results. Section 7 presents the conclusions.

## **2. Data on the Census of Logistics**

In this section, I describe the *Census of Logistics* (CL) used for analyzing the characteristics of transport costs in Japan. The Japanese Ministry of Land, Infrastructure, Transport and Tourism conducts the survey on business enterprises in 47 Japanese prefectures and 4 sectors: mining, manufacturing, wholesale, and warehouse industries. The primary objective of the logistics survey is to examine comprehensively the flow of domestic freight from the demand side of transportation in order to understand the origin and destination of freight as well as the relationship between logistics and industrial activity. Begun in 1970, the survey has been conducted every 5 years; however, the question concerning transport costs has

been included in the survey only since 1995. This paper exploits the survey results from the years 2000 and 2005.

The logistics survey defines freight as materials, manufactures, and commodities that are shipped in and out of the business enterprise for the purpose of production, purchase, and sale. However, the survey excludes freight that is not directly related to production/sale activities such as business documents, empty containers, and industrial waste. The destination of freight as defined above ranges from foreign markets, domestic industries, and individual persons. On the other hand, the origin of freight flows does not include industries such as agriculture, forestry and fishery, construction, retail, and services.

The sampling scheme of the logistics survey is carefully designed to estimate the actual characteristics of domestic transportation flows in the population as defined above. Specifically, the sample size is determined according to three strata: industry, employment, and prefecture. First, the survey defines the number of business enterprises in each industry of interest from other official statistics, and then decides the number of the enterprises to be sampled to meet the minimum sampling rates.<sup>2</sup> In 2005, 63,417 enterprises were surveyed by interview or through mailed questionnaire for shipments sent during three days in October. The survey questions included product, volume and quantity, transport route, and shipping time and cost. Responses were received from 21,026 of the surveyed enterprises. The rate of response was significantly higher for interviewed enterprises (78.1%) compared with those by mailed survey (31.8%). The response by mining and warehouse industries was over 40% while manufacturing and wholesale industries were below 40%.

From the census of logistics, I created a two-year panel data set on domestic transport costs disaggregated by pairs for 47 prefectures, 8 commodity categories, and 11 transportation

---

<sup>2</sup> For details, see <http://www.mlit.go.jp/seisakutokatsu/census/census.html>



modes. The types of goods range from agricultural and marine products, wood products, non-metallic minerals, metals and machinery, chemicals, light industrial products (paper, pulp, food, and beverages), various products (printing, leather, rubber, and plastics), to special goods (fertilizers, containers, and paper boxes). Transport modes range from railway by container, other railway, private truck, delivery-services truck, rental truck, commercial trailer truck, ferryboat, container ship, RORO ship, other marine shipping, air transport, and other.<sup>3</sup> From the census of logistics, I also use data on transport time disaggregated by transport modes and prefecture pairs as well as tonnage of transportation flows disaggregated by major goods, transport modes, and prefecture pairs.

In the prior literature, ad-valorem freight rate of trade and distance have been widely used to measure transport costs. Hummels (1999) estimates the relationship between freight cost and distance for imports of the U.S., New Zealand, and Latin American countries in 1994. He finds that the distance elasticity with respect to freight rates is, on average, 0.27. The estimates range from 0.46 for U.S. imports by air to 0.22 for those by ocean shipping. Combes and Lafourcade (2005) report that estimated transport costs for truck shipping increases with distance traveled across regions in France. They find that the distance is highly correlated with freight fees at a point in time, but not over time. Further, Limão and Venables (2001) estimate the determinants of transport costs for a standard container shipped from Baltimore in the U.S. Their findings indicate that shipping charges from Baltimore increase with respect to the distance to a destination market, with the pronounced positive impact of land distance.

There are, however, limitations on these proxy variables for understanding the relationship between the distance and willingness-to-pay for transport services. In particular, the

---

<sup>3</sup> RORO ship stands for roll-on, roll-off ship. The RORO ship can accommodate commercial vehicles and trailer trucks without lifting them by crane. Ferryboats and RORO ships are used for marine transportation at small-scale marine ports.

ad-valorem freight rate is measured between countries by customs officials, i.e., port-to-port shipping costs. The distance is measured between the points of each region. Consequently, these conventional measures may not well capture the precise size of transport costs between producers and consumers. In contrast, the CL data are distinctive in that transport cost is a direct measure of individual shipment fees paid by the business enterprise. This allows me to directly investigate the willingness-to-pay of the enterprises for moving commodities over space.

### **3. Patterns of Transportation**

This section describes characteristics of transport costs and time aggregated over a combination of three categories: goods, shipping mode, and region. To account for a distance effect on shipping fees measured in Japanese yen per ton, the transport costs per ton are divided by the distance in 100 kilometers between prefectural offices.<sup>4</sup> I then use data on freight fees per ton/100km while excluding shipments within prefectures. Prefectures are categorized into 8 regions according to their location from north to south: Hokkaido, Tohoku, Kanto, Chubu, Kinki, Chugoku, Shikoku, and Kyushu.<sup>5</sup> Transportation mode is aggregated over air, railroad, ship, and truck. Furthermore, I carefully examine the sample to mitigate possible reporting/aggregation errors. Specifically, I exclude the samples in which transportation costs and time fall in the top or bottom 1% tail of their distributions, respectively.

Table 1 presents descriptive statistics of transport costs in yen per ton/100km over the region of prefectures sending shipments and the freight mode. Transport cost is the highest for air transportation, followed by truck transportation. Railroad and ship transportation are less

---

<sup>4</sup> Alternative units of the measurement may include transport fees per ton-hours and ton-values, but there appears to be no consensus on the most appropriate unit of measurement.

<sup>5</sup> The correspondence between the prefecture and region is provided in the Appendix Table.

expensive than these modes. By contrast, the difference between sending regions appears to be relatively small for each transport mode. In particular, cross-regional differences in average transport cost are relatively small for railroad and ship transportation. In the case of air shipping, average freight fees from major economic areas such as Kanto, Chubu, and Kinki are relatively larger than other sending regions. These patterns indicate that unit transport cost differs much more by transport mode than by the origin of shipments. In addition, average transport costs are positively associated with a dispersion of transport cost measured by standard deviations. Transport costs are more variable in air and truck shipping than in railroad and ship freight.

**[Table 1]**

In Table 2, transport fees are aggregated over transportation mode and commodity group. Freight costs in air and truck transportation are large relative to railroad and ship freight. A cross-product variation in average freight fees appears to be relatively small for railroad, ship, and truck transportation. In the case of air shipping, a cross-product difference in shipping fees is significantly large, with shipments of non-metallic minerals being relatively expensive. Further, a dispersion of average freight fees measured by standard deviations seems to increase with respect to the average freight fees. These patterns are generally in line with the wide variation of international freight rates within commodities as found in Hummels (2001).

**[Table 2]**

The CL data also allow me to describe the pattern of transport time over prefecture pairs and transport modes. To isolate distance effects, shipping time is measured in hours per 100km. Table 3 shows a pattern of transport time aggregated over sending regions and transport modes. A distinctive feature is that average shipping time is relatively shorter for air and truck transportation than railroad and ship transportation. Variance in shipping time measured by standard deviations is also much smaller for air and truck transport. Thus, it appears that

average shipping time is positively associated with variance in shipping time. These patterns are in stark contrast with the observation on characteristics of transport costs, suggesting a trade-off in transport services between shipping costs and time. In addition, a cross-regional variation in shipping time is relatively small across transport modes.

[Table 3]

#### **4. Transport Cost Function**

Before proceeding with the empirical analysis, this section will explain the iceberg form of transport cost function to study the relationship between transport costs and geographic distance. As McCann (2005) points out in explaining the nature of transport costs in the new economic geography, Samuelson (1952) first introduced an iceberg formulation of transport costs that generated a price deviation of identical goods between home and foreign markets. Because the price difference was simply discontinuous between different markets, there was not specific role of the geographical distance in Samuelson's iceberg form of transportation costs. By contrast, Krugman (1991) introduced the explicit role of distance in the iceberg formulation of transport costs, which allows for investigating a spatial pattern of economic activities. Thus, I focus on Krugman's formulation of iceberg shipping costs to motivate the empirical analysis.

##### **4.1. Iceberg Transport Cost**

Suppose there is a producer in location  $i$  and a consumer in location  $j$ , with the geographic distance between the producer and consumer denoted by  $D_{ij}$  in kilometers. The price per ton of a good in locations  $i$  and  $j$  is  $P_i$  and  $P_j$ , respectively. The tonnage of the good in each location  $i$  and  $j$  is expressed by  $Q_i$  and  $Q_j$ . The logic of iceberg transport costs implies that a part of the good melts away in the process of transportation from  $i$  to  $j$ . As such, the speed of melting

is assumed to be a function of the distance (Krugman, 1991; Fujita et al., 1999). Specifically, the iceberg form of transport costs for the good delivered from  $i$  to  $j$  is:

$$Q_j = Q_i \cdot \exp(-\eta \cdot D_{ij}), \quad \eta > 0 \quad (1)$$

where  $\eta$  is an iceberg parameter of the proportion of the remaining tonnage of the delivered good. It is useful to express the cost of shipping in a relative price, rather than a relative quantity, of the delivered good. A consumer in location  $j$  must pay a price per ton ( $P_j$ ) of the good transported from location  $i$ , as expressed in a relative price:

$$P_j = P_i \cdot \frac{Q_i}{Q_j} \quad (2)$$

Using equation (1), the iceberg transport cost in an ad-valorem form is written as:

$$\tau_{ij}^{\text{ad}} \equiv \frac{P_j}{P_i} = \exp(\eta \cdot D_{ij}) \quad (3)$$

This expression indicates that the price per ton of the delivered good in  $j$  is increasing with the iceberg decay parameter and geographic distance between  $i$  and  $j$ .

Equation (3) describes the relationship between ad-valorem transport costs and distance, which is derived from the iceberg formulation of freight fees. Taking first and second derivatives of the equation with respect to  $D_{ij}$  gives the following properties of the transport cost function:

$$\frac{\partial \tau_{ij}^{\text{ad}}}{\partial D_{ij}} = \eta \cdot \exp(\eta \cdot D_{ij}) > 0 \quad (4)$$

$$\frac{\partial^2 \tau_{ij}^{\text{ad}}}{\partial D_{ij}^2} = \eta^2 \cdot \exp(\eta \cdot D_{ij}) > 0 \quad (5)$$

Equations (4) and (5) show that ad-valorem transport cost is increasing with the geographic distance between  $i$  and  $j$  at an increasing rate over the distance; the transport cost function is strictly convex with respect to distance.

The property of the iceberg transport cost function provides an underlying hypothesis for empirically analyzing the relationship between shipping costs and distance.

However, the transport cost function needs to be modified in order to meet the nature of the available data on real freight fees paid by producers while preserving the essential property of iceberg transport costs. In reality, producers are likely to pay directly for transport services by shipping companies to deliver their product to consumers/firms in a distant market. Then, the cost of moving the product is added to the price per ton of the good that consumers/firms in a destination market must pay. As a result, it is difficult to measure the precise transport costs from the relative price of identical goods between origin and destination markets.

To deal with the nature of real transportation, the cost of moving a commodity over space is redefined as the total amount of transport costs per ton of a good that a producer in location  $i$  pays to deliver to a consumer/firm in location  $j$  over distance  $D_{ij}$  in kilometers. This implies that transport cost can be expressed as a difference between price per ton in origin and destination markets  $i$  and  $j$ . In this paper, the price gap is specified as an additive function of transport distance, transport fixed cost, and other remaining determinants:

$$\tau_{ij} \equiv P_j - P_i = \exp(\eta \cdot D_{ij}) + f + e \quad (6)$$

The first term is the geographic distance of the form as implied by the iceberg transport cost function. It is evident that the iceberg property of transport cost with respect to distance is carried over in the modified equation; transport cost is strictly convex in distance. The second term,  $f$ , is the fixed cost of transporting a commodity between locations  $i$  and  $j$ . Finally, the last term,  $e$ , captures the remaining elements of the price gap, including producer's pricing to market, demand for different commodities, and product market regulation.

#### **4.2. Empirical Specification**

The purpose of this empirical analysis is to estimate the relationship between transport cost and distance as specified in equation (6). The relationship can be examined by estimating

the iceberg decay parameter of distance. However, it is difficult to directly estimate a nonlinear relationship between these variables while isolating other determinants of transport costs in the data. Instead, I estimate a linear empirical model that can maintain the hypothetical properties of iceberg transport cost with respect to distance. Specifically, a linear version of equation (6) is written as:

$$\tau_{ij} = \alpha_1 D_{ij} + \alpha_2 D_{ij}^2 + f + e \quad (7)$$

where  $\alpha_1$ ,  $\alpha_2$  are unknown parameters. A conventional assumption of iceberg transport cost suggests that shipping costs increase with respect to distance at an increasing rate (McCann, 2005). This suggests the following hypotheses for the unknown parameters in equation (7):

$$\frac{\partial \tau_{ij}}{\partial D_{ij}} = \alpha_1 + 2\alpha_2 D_{ij} > 0, \quad \frac{\partial^2 \tau_{ij}}{\partial D_{ij}^2} = 2\alpha_2 > 0$$

A deterministic version of equation (7) holds in theory, but there is no set of parameters for which the equation will hold exactly for an arbitrary set of observations. To address deviations from the theoretical assumptions, equation (7) can be transformed into a stochastic version of the transport cost function. Further, equation (7) implicitly assumes that transport costs vary solely by distance. A more realistic transport cost equation should reflect to some extent the heterogeneous freight fees. To this end, I allow for a variation in freight costs over delivered goods, transport mode, and time period.

Additionally, recent literature highlights the importance of delivery *time* in transporting goods across and within borders. (Hummels, 2001; Deardorff, 2003; Evans and Harrigan, 2005; Djankov et al. 2010). Producers should take into account the length of shipping time in choosing transport mode to minimize the level of transport costs per ton. The omission of transport time that is highly correlated with distance would cause an endogeneity bias. Thus, transport cost measured in yen per ton is further divided by average shipping time to obtain freight costs in yen per ton/hour.

Given the data on transportation, I estimate the following empirical model for sending prefecture  $i$ , arriving prefecture  $j$ , commodity category  $k$ , transport mode  $m$ , and year  $t$ :

$$\tau_{ijkmt} = \alpha_0 + \alpha_1 D_{ij} + \alpha_2 D_{ij}^2 + \theta_i d_i + \theta_j d_j + \theta_k d_k + \theta_m d_m + \theta_t d_t + \varepsilon_{ijkmt} \quad (8)$$

- $\tau$  is transport costs in yen per ton/hour
- $D$  is geographic distance in kilometers between prefectural offices<sup>6</sup>
- $d_i$  is a dummy variable for sending prefecture  $i$
- $d_j$  is a dummy variable for arriving prefecture  $j$
- $d_k$  is a dummy variable for commodity category  $k$
- $d_m$  is a dummy variable for transport mode  $m$
- $d_t$  is a dummy variable for time period  $t$
- $\varepsilon$  is an error term representing other unobservable influences on transport costs
- $\alpha_0, \alpha_1, \alpha_2, \theta_i, \theta_j, \theta_k, \theta_m, \theta_t$  are parameters to be estimated.

A consistent estimate of the parameters is obtained by ordinary least squares (OLS) under the assumption that the conditional expectation of the error is zero:  $E[\varepsilon | D, d_i, d_j, d_k, d_m, d_t] = 0$ .

Before discussing econometric issues, it is useful to contrast equation (8) with the trade cost function that is commonly assumed to estimate the gravity model of international trade. To account for trade friction between nations, unobservable trade cost is a loglinear function of distance between countries and other observable variables including national borders, common language, and colonial ties (Anderson and van Wincoop, 2003). By contrast, this paper focuses on transport costs within a country rather than across countries. This approach allows for isolating the influence of national barriers from the distance effect on trade costs. The estimate of the freight-distance relationship between domestic markets is distinctive from the estimate using ad-valorem freight rates (Hummels, 1999), and shipping company quotes (Limão and Venables, 2001).

---

<sup>6</sup> Distance data were obtained from the Japanese Geographical Survey Institute at <http://www.gsi.go.jp/KOKUJYOHO/kenchokan.html>.



### **4.3. Econometric Issues**

Consistency in OLS estimation depends on the assumption that the expected value of the error term is zero, conditional on observable explanatory variables. Because the geographic distance is fixed and can be measured precisely, there is little concern about endogeneity bias. Omitted variables bias can also be mitigated by controlling for a variety of fixed effects over several characteristics of transport costs. By contrast, sample selection bias is likely to be a serious concern. Specifically, this paper aims to estimate a transport cost function for all transport services within a country, but it is likely to obtain a random sample only for a subset of all the shipping transactions. The CL data collected in this paper allows for a maximum of 388,784 possible observations, but the actual sample size in the data is 40,465 observations before excluding possible outliers. Thus, the large number of missing samples of shipping transactions could be due to the selection mechanism in the market for freight services, resulting in nonrandom sampling of transport costs.

A plausible selection mechanism could emerge from an unobserved structure of underlying transport costs faced by producers for moving their goods. The key fact is that producers will choose a bundle of transport services to minimize the cost of transport services over distance. Excessively lengthy shipping should impose an upper boundry on transport fees for which producers are willing to pay, which in itself would cause transport transactions to cease. On the other hand, alternative modal choices over a specific transport route generate an incentive to concentrate on a specific mode of transportation for economies of scale. Because producers determine their payments for transport services according to the underlying schedule of freight costs, it is likely that only a subset of the samples of shipments will be observed.

To address a sample selection bias, I employ the sample selection model by Heckman (1979). As a first step, I specify a selection equation that determines whether the business

enterprises pay transport fees for shipments:

$$\Pr(d_{ijkmt}^r = 1 | V_{ijkmt}, X_{ijkmt}) = \Phi(\rho V_{ijkmt}, \pi X_{ijkmt}') \quad (9)$$

where  $V$  is the total tonnage of commodity flows and  $X$  is the vector of regressors. To distinguish the transport cost equation from the selection equation, the observed volume of commodity flows is assumed to influence the selection decision by producers over transportation. From a supply side, the volume of shipments over specific routes indicates the potential capacity of transport services, so that commodity flows are likely to represent a supply curve of transport services. Consequently, the commodity flows affect the probability that producers take advantage of transport services. In addition, it is assumed that commodity flows do not enter the transport cost function; the average cost of shipments over specific routes is not likely to be determined by the total weights of shipment flows. In the estimation, the estimated inverse Mills ratios from equation (9) are included in equation (8) estimated by OLS for the selected sample.

## 5. Estimation Results

This section presents the estimation results of equation (8) to examine the linkage between transport costs and distance. As discussed above in the empirical issues, there can be a sample selection mechanism on the decision by business enterprises over transport services. Because available samples generated by the selection mechanism differ from the full sample in which transport costs are either observed or unobserved, I make a clear distinction between the selected and full samples on transport costs.

### 5.1. Summary Statistics

Before discussing the estimation results, I will present the summary statistics and

correlation coefficients of key variables in the selected and full samples. The selected sample in Table 4 shows that transport costs in yen per ton/hour range from 0.19 to 286, with the mean of 6.67 and standard deviation of 12.2. On average, business enterprises are charged 6.67 yen per one hour for sending one tonnage of commodities. The distance of shipments delivered varies from 19 to 2,244 kilometers, with the mean of 438 and standard deviation of 299. In the data, domestic freight is delivered for 438 kilometers on average. Additionally, the full sample indicates that 10% of transport transactions are observed in the maximum possible sample size in the CL data.

**[Table 4]**

Table 5 lists the correlation coefficients between key variables. The selected sample shows that transport fees are negatively correlated with shipping distance, indicating that business enterprises are charged higher shipping fees for short-distance freight. In the full sample, the probability of observing transport costs is also negatively associated with distance. Short-distance shipments are more likely to be observed in the survey than long-distance freight. Further, the transport indicator is positively correlated with the volume of commodity flows in the full sample, which is consistent with the idea that a larger supply of transport services measured by commodity flows increases the probability that business enterprises will take advantage of transport services.

**[Table 5]**

## **5.2. OLS Estimation**

Table 6 shows the results of equation (8) estimated by Ordinary Least Squares (OLS) with robust standard errors. The model in column (1) includes a distance variable and year fixed-effects. The distance variable has a significantly *negative* coefficient, with the coefficient

size implying that an additional 100 kilometers of freight distance is associated with a reduction in transport costs in yen per ton/hour by 1.034. To allow for a quadratic relationship, column (2) includes a square term of the distance variable. As the square term is significantly negative, transport costs are concave with respect to shipping distance.

### **[Table 6]**

In column (3), I control for a variety of unobserved fixed effects over prefectures, transport modes, and commodity groups. As expected, the distance variables have the significant coefficients. An econometric interpretation is that transport costs are decreasing at a decreasing rate, but start to increase after the turning point of 910 kilometers. When measured at the long distance, the CL data support the iceberg assumption that transport costs are convex in relation to distance.

Since transport costs are often assumed to be a log-linear function of distance, I specify the estimating equation in logs. In column (4), the distance variable has a significantly negative coefficient, indicating that a 1% increase in distance reduces transport costs in yen per ton/hour by 0.95%. Column (5) shows that the negative elasticity depends positively on shipping distance. Even after controlling for a variety of fixed effects in column (6), the signs of the distance variables are unchanged.

### **5.3. Sample Selection Estimation**

Up to this point, the regression analysis has assumed away selection bias in the estimation, which may affect the estimated impact of distance in transport costs. Because producers take into account shipping distance and costs to determine freight, only a portion of the sample shipments is likely to be observed. This problem could be a possible reason for the negative distance effects in OLS estimation. To address a sample selection bias, I employ

Heckman's two-step estimation. In the selection model, commodity flows are defined as a log of the flows plus 0.001 to avoid a loss of observations on zero commodity flows.

Table 7 presents the results of the log-specification of equation (8) by the Heckman estimation with bootstrapped standard errors. In column (1), commodity flows have a significantly positive coefficient in the probit model, suggesting that business enterprises are more likely to take advantage of freight services when there are greater flows of goods delivered. The geographic distance also has a significantly positive impact on the transport indicator, suggesting that freight fees on long distant shipments are more likely to be observed. In column (2), the significant coefficient of the inverse Mills ratio indicates the presence of selection bias. Even after controlling for a sample selection bias, the distance variable has a significantly negative coefficient.

#### **[Table 7]**

To allow for nonlinearity in the log specification, I estimate the probit model with a quadratic term of the log of distance. In column (4), both distance variables are significantly negative, conditional on the selection bias corrected by the inverse Mills ratio. Finally, I include a variety of fixed effects in model 3. The results in column (6) indicate that the negative coefficients of the distance variables are robust to unobserved fixed effects over sending and arriving prefectures, transport modes, and commodity groups. Taken together, these results suggest that a robust negative link between transport costs and distance is not driven by sample selection bias.

#### **5.4. Sample Selection Estimation by Transport Mode**

The previous regressions have assumed that distance has identical influences on transport costs across shipping mode, which may mask the possibility that business enterprises

would choose specific transport modes based on the length of shipping distance. To address the issue of modal choice by firms, I separate the samples by rail, truck, ship, and air transportation. As shipping modes vary by the samples, I drop the fixed effects of transport modes from the estimating equation.

Table 8 presents the results of the Heckman estimation. The inverse Mills ratio is significantly positive across specifications, indicating the importance of controlling for selection bias in each freight mode. Column (1) shows that the log of distance is significantly positive, but the square term is significantly negative. This implies that an increase in distance is positively associated with transport costs for air shipping, but the increment declines over distance. While transport costs are convex in distance under the iceberg formulation of transport costs, air shipping fees observed in the CL data are concave in distance. A possible explanation for the concavity is that a large fixed cost in air transportation has a diminishing portion of freight costs over distance, meaning that additional costs of air shipments also decrease over distance.

**[Table 8]**

Estimates of the distance effects for railroad, ship, and truck transportation are presented in columns (2) through (4), respectively. The results imply that an increase in distance is associated with a decline in freight costs by railroad and truck at an increasing rate. On the other hand, an increase in distance is related negatively with freight fees by ship transportation, with the additional increase decreasing over distance. The estimates of distance effects in the previous regressions appear to be driven primarily by the samples for these transport modes. These results can be interpreted as suggesting that business enterprises are likely to pay higher freight fees for their shipments of short distance by railroad, ship, and truck transportation.

## **6. Discussions on Transport Costs and Distance**

The analysis up to this point has illustrated significant heterogeneity in transport costs and time across shipping modes, commodities, and regions. As faster shipping is more expensive than slower shipping, a systematic variation in freight costs and time supports the reliability of the Japanese Census of Logistics. Using the CL data, I find that air transport costs increase with respect to distance at a decreasing rate, after carefully accounting for a number of other determinants and selection bias. However, transport costs by other shipping modes tend to fall with respect to distance, suggesting that firms pay higher for short-distance shipments. These surprisingly puzzling results raise two related questions. What bias would remain in accounting for transport costs? If the data represent actual firm behaviors toward transport services, why are firms willing to pay more for shipments of shorter distance?

As discussed in the econometric issues, I address a wide range of influential bias in estimating distance effects. While this paper exploits extremely detailed information on freight costs, there is still aggregation bias that could play a large role in determining transport costs. Specifically, millions of individual freight transactions in the survey are aggregated to create data on transport costs that vary solely by prefectures, goods, and modes. As such, aggregate transport costs contain substantial variations in freight-specific costs, which could introduce an influential bias in an estimated distance effect.

Large differences in individual freight transactions are shown in the report on the Census of Logistics (2005). For instance, small-batch shipments have risen in importance for recent decades. The average volume per unit of shipment was 2.43 tons in 1990, which dropped to 2.13 in 1995, 1.73 in 2000, and 1.27 in 2005. Shipments of less than 0.1 ton account for almost 70% of the two millions total shipping transactions in the 2005 survey. This pattern also applies to such industries as manufacturing, wholesaling, and warehousing. The weight of unit

freight in the wholesale sector declined from 0.72 in 1990 to 0.36 in 2005. In the manufacturing sector, the corresponding figure decreased from 3.16 in 1990 to 2.06 in 2005.

Transportation in the modern economy is also characterized by timely shipping. The CL report (2005) provides information on the proportion of shipments whose arrival time is designated by hour, morning or afternoon, day, and none. In 2005, for instance, almost 80% of shipments in the manufacturing sector were subject to timely delivery in terms of aggregate freight weight; arrival time by hour was 35%, by morning or afternoon is 14.5%, and by day was 31.4%. At the aggregate industry level, these shares were 27.7%, 14.6%, and 31.4%, respectively. Taken together, it is evident that the majority of freight was sent in small batches by frequent and timely shipping. These elements should add substantially to freight fees at the aggregate level. Freight-specific costs are a plausible source of aggregation bias, which is worthy of further investigation for understanding transport costs.

Given that firms pay a substantial premium for timely, frequent, and small-batch shipments, the question remains as to why shipping premiums play a large role in short-distance freights. The starting point for a plausible hypothesis is the nature of modern manufacturing production. As Sakakibara et al. (1997) point out in their empirical analysis of Just-in-Time (JIT) manufacturing, the JIT production system has become prevalent in manufacturing production lines to reduce the cost of inventory holdings. To cut down on stocks, manufacturers receive only the necessary components and parts from suppliers only at the necessary time. For this reason, the delivery of the components has to be timely, frequent, and in small batches in the JIT system. As a result, the flexible logistics in the production system allow for a quick response to defective components and customer orders. In the JIT production network, a large opportunity cost of late delivery induces firms to place a high value on timely delivery.

Sophisticated logistics embedded in manufacturing production have implications on



the location of vertical production activity. Harrigan and Venables (2006) describe the theoretical mechanism whereby timeliness creates an incentive for the clustering of component producers and assembly plants. In the JIT production system, the opportunity cost of uncertain deliveries is potentially large because the delay in some stages of production processes could lead to the disruption of overall manufacturing operations. On the other hand, geographic proximity contributes not only to a reduction in transport costs over space, but also to an improvement in variability of freight arrival. As a result, parts and components suppliers in upstream sectors have a strong motivation to locate in close proximity to final-goods producers in downstream sectors. While industrial clusters may generate a spatial concentration of demand for costly transports, economic agglomeration can have a depressing effect on the demand for expensive fast transportation for long-distance freight. The link between industrial agglomeration and costly timely shipping may provide a plausible explanation for why firms are willing to pay higher for freights in transported shorter distances. Indeed, Evans and Harrigan (2005) have, in the case of U.S. retail sector, shown evidence for a model in which demand for timely delivery has shifted production to locations in proximity to the home market.

## **7. Conclusion**

Transport costs play a central role in accounting for a wide range of economic activities including industrial agglomeration, international trade, and foreign direct investment. In the empirical literature, geographic distance has been used extensively as a proxy for transportation costs under the simple assumption that transport costs increase monotonically over space. This paper provides an empirical analysis of the relationship between transport fees and distance from the demand side of transport services, using the extremely rich information on domestic transportation in Japan. As is common in the gravity model of international trade,

the iceberg formulation of transport costs is taken as a basis to specify the transport cost function in which transport cost is convex in distance. Then, I have modified the transport cost function to make it consistent with freight costs in the data, while preserving its convex property.

The descriptive analysis shows substantial heterogeneity in transport costs and time across transport modes, commodity groups, and regions. Consistent with causal empiricism, transport costs by air in yen per ton/km are the highest among alternative modes, followed by truck, ship, and railroad transportation. The dispersion of transport costs increases with respect to the average transport costs, indicating that expensive shipments tend to exhibit more variable freight fees. On the other hand, air shipping time in hours per kilometers is the shortest, followed by truck, ship, and railroad transportation. The dispersion of shipping time is smaller for transport modes with shorter average transport time. These patterns in transport costs and time indicate a trade-off between freight costs and time, which is consistent with intuition. From a theoretical point of view, however, the results are not consistent with the iceberg assumption that transport costs increase solely over distance; instead, unit freight costs vary substantially by shipping modes.

The estimation results of the transport cost function reveal that air transport costs increase significantly over distance, but at a decreasing rate. This is only partly consistent with the iceberg assumption that freight costs increase at an increasing rate over distance. Perhaps, fixed costs in air transportation might have a depressing effect on additional costs of long-distant shipments by air, leading to a concave relationship between freight costs and distance. In contrast, I find robust evidence that distance has a significantly *negative* correlation with transport costs by truck, ship, and railroad even after controlling for a number of other determinants and sample selection bias. Taken together, robust negative distance effects can be

interpreted as indicating that business enterprises in the survey are likely to pay higher for short-distance freight.

The CL data on transportation appear to be reliable, but generate a surprisingly puzzling pattern between transport costs and distance. A plausible reason is that the data on transport costs used in the analysis could be subject to aggregation bias arising from freight-specific premiums. Specifically, individual freight transactions are characterized by timely, frequent, and small-batch shipping. These expensive means of transportation in individual transactions should add substantially to the observed transport costs at the aggregate level. While the currently available data do not allow for analyzing transport costs at the freight-transaction level, it is worthy of further investigation to estimate freight-specific costs. Additionally, this paper discusses the linkage between industrial agglomeration and timely delivery in modern production networks, which could lead to a spatial concentration of demand for premium transport services. Given that firms tend to pay a high premium for short-distance freights, it may be possible to observe a negative correlation between distance and transport costs at the aggregate level. To investigate this hypothesis, it is necessary to identify a business relationship between business enterprises linked by transport services. These are questions that need to be examined in future empirical research.

## References

- Anderson, James E. and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review*, 93:1, pp. 170-92.
- Anderson, James E. and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature*, 42:3, pp. 691-751.
- Brainard, S. Lael. 1997. "An Empirical Assessment of the Proximity-Concentration Trade-off between Multinational Sales and Trade." *American Economic Review*, 87:4, pp. 520-44.
- Combes, Pierre-Philippe and Miren Lafourcade. 2005. "Transport Costs: Measures, Determinants, and Regional Policy Implications for France." *Journal of Economic Geography*, 5:3, pp. 319-49.
- Deardorff, Alan V. 2003. "Time and Trade: The Role of Time in Determining the Structure and Effects of International Trade, with an Application to Japan," in *Japan's economic recovery: Commercial policy, monetary policy, and corporate governance*. Robert M. Stern ed: Cheltenham, U.K. and Northampton, Mass., pp. 63-76.
- Disdier, Anne-Celia and Keith Head. 2008. "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *Review of Economics and Statistics*, 90:1, pp. 37-48.
- Djankov, Simeon, Caroline Freund, and Cong S. Pham. 2010. "Trading on Time" *Review of Economics and Statistics*, 92:1, pp. 166-173
- Evans, Carolyn L. and James Harrigan. 2005. "Distance, Time, and Specialization: Lean Retailing in General Equilibrium." *American Economic Review*, 95:1, pp. 292-313.
- Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. *The spatial economy: Cities, regions, and international trade*: Cambridge and London: MIT Press.
- Gallup, John Luke, Jeffrey D. Sachs, and Andrew D. Mellinger. 1999. "Geography and Economic Development." *International Regional Science Review*, 22:2, pp. 179-232.
- Harrigan, James and Anthony J. Venables. 2006. "Timeliness and Agglomeration." *Journal of Urban Economics*, 59:2, pp. 300-16.
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47:1, pp. 153-61.
- Hummels, David. 1999. "Toward a Geography of Trade Costs." Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University, GTAP Working Papers.
- Hummels, David. 2001. "Time as a Trade Barrier." Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University, GTAP Working Papers.
- Hummels, David. 2007. "Transportation Costs and International Trade in the Second Era of Globalization." *Journal of Economic Perspectives*, 21:3, pp. 131-54.
- Krugman, P. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy*, 99:3, pp. 483-499.
- Limão, Nuno and Anthony J. Venables. 2001. "Infrastructure, Geographical Disadvantage, Transport

- Costs, and Trade." *World Bank Economic Review*, 15:3, pp. 451-79.
- Markusen, James R. 2002. *Multinational firms and the theory of international trade*: Cambridge and London: MIT Press.
- McCann, Philip. 2005. "Transport Costs and New Economic Geography." *Journal of Economic Geography*, 5:3, pp. 305-18
- Ministry of Land, Infrastructure, Transport and Tourism. 2005. "*Report on the Census of Logistics*." Tokyo.
- Sakakibara, Sadao, Barbara B. Flynn, Roger G. Schroeder, William T. Morris. 1997. " The Impact of Just-in-Time Manufacturing and Its Infrastructure on Manufacturing Performance." *Management Science*, 43:9, pp. 1246-1257
- Samuelson, P. 1952. "The Transfer Problem and Transport Costs: the Terms of Trade When Impediments are Absent." *Economic Journal*, 62, pp. 278–304.

**Table 1. Transport Costs by Sending Region and Transport Mode**

Sending Region	Mode		<u>Air</u>		<u>Railroad</u>		<u>Ship</u>		<u>Truck</u>	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Hokkaido	46.7	64.1	2.74	4.87	1.41	1.31	15.5	55.5		
Tohoku	60.5	98.4	4.18	6.88	2.55	2.50	35.1	89.0		
Kanto	86.4	107.0	6.25	9.35	4.99	13.3	57.3	129.7		
Chubu	71.0	96.7	6.97	8.29	4.95	12.6	46.0	95.7		
Kinki	86.8	108.9	6.62	8.33	7.44	16.1	68.7	141.8		
Chugoku	55.5	125.9	5.31	10.4	6.92	25.5	43.6	101.1		
Shikoku	78.7	127.1	11.0	24.2	11.1	40.4	53.3	108.1		
Kyushu	60.1	103.5	3.86	7.95	4.02	8.47	39.0	94.8		
Total	69.0	103.8	5.78	9.81	5.93	19.4	49.7	112.3		

*Note:* transport cost is measured in yen per ton/100km.

**Table 2. Transport Costs by Commodity Group and Transport Mode**

Commodity Group	Mode		<u>Air</u>		<u>Railroad</u>		<u>Ship</u>		<u>Truck</u>	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Agricultural and marine products	41.0	67.3	3.1	3.7	3.9	7.1	72.8	150		
Wood products	n.a.	n.a.	1.6	1.1	5.5	9.0	60.8	149		
Non-metallic minerals	86.2	66.0	5.8	4.7	13.0	25.7	59.8	117		
Metals and machinery	70.7	108	6.1	9.4	6.1	19.6	45.5	107		
Chemicals	66.7	83.1	6.1	9.3	8.2	26.0	38.1	89.5		
Light industrial products	28.6	38.1	5.5	10.8	1.7	1.3	39.5	91.1		
Various (printing, leather, rubber)	87.6	124	6.6	11.3	4.7	15.6	58.8	124		
Other (fertilizers, containers, boxes)	50.8	63.4	4.9	4.3	3.1	2.7	73.3	135		
Total	69.0	104	5.8	9.8	5.9	19.4	49.7	112		

*Note:* transport cost is measured in yen per ton/100km.

**Table 3. Transport Time by Sending Region and Transport Mode**

Sending Region	Mode		Air		Railroad		Ship		Truck	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Hokkaido	3.21	1.39	7.27	4.64	6.66	2.08	5.77	2.31		
Tohoku	2.61	1.26	6.56	3.93	7.92	4.81	4.85	3.47		
Kanto	3.07	1.56	6.92	4.99	8.69	5.92	5.72	4.26		
Chubu	3.02	1.45	9.26	6.03	12.7	7.86	5.73	3.74		
Kinki	3.43	1.99	7.02	5.52	11.9	7.79	6.31	5.10		
Chugoku	2.93	1.36	8.36	5.21	7.90	3.32	5.41	3.66		
Shikoku	3.34	1.66	10.0	6.62	8.03	3.79	6.46	4.64		
Kyushu	2.99	1.84	6.13	3.86	7.21	4.75	5.68	3.99		
Total	3.06	1.65	7.53	5.24	8.96	5.91	5.75	4.18		

*Note:* transport time is measured in hours per 100km.

**Table 4: Summary Statistics**

Variable	Obs.	Mean	S.D.	Min	Max
<u>Selected Sample</u>					
Transport cost, yen per ton/hour	36848	6.67	12.2	0.02	286
Distance, 100km	36848	4.38	2.99	0.19	22.4
Distance squared, 100km	36848	28.2	36.9	0.04	503
Commodity flows, ton	36848	364	1788	0.001	116260
<u>Full Sample</u>					
Transport indicator: 1 if observed, zero otherwise	380512	0.10	0.30	0	1
Distance, 100km	380512	5.20	3.55	0.11	22.4
Distance squared, 100km	380512	39.6	53.8	0.01	503
Commodity flows, ton	380512	51.9	751	0	116260

**Table 5: Correlation Coefficients**

Variable	(1)	(2)	(3)	(4)
<u>Selected Sample : obs.=36848</u>				
Transport cost	1.00			
Distance	-0.25	1.00		
Distance squared	-0.18	0.95	1.00	
Commodity flows	0.05	-0.15	-0.11	1.00
<u>Full Sample: obs.=380512</u>				
Transport indicator	1.00			
Distance	-0.07	1.00		
Distance squared	-0.06	0.94	1.00	
Commodity flows	0.14	-0.06	-0.04	1.00

**Table 6. OLS Estimation of Transport Cost Function**

Dependent variable: Transport cost in yen per ton/hour

Variables	Specification	<u>Levels</u>			<u>Logs</u>		
		(1)	(2)	(3)	(4)	(5)	(6)
Distance		-1.034** (0.022)	-3.394** (0.083)	-3.732** (0.082)	-0.946** (0.007)	-1.069** (0.013)	-1.091** (0.011)
Distance squared			0.202** (0.006)	0.205** (0.007)		0.062** (0.007)	0.020** (0.006)
Sending prefecture FE		N	N	Y	N	N	Y
Arriving prefecture FE		N	N	Y	N	N	Y
Transport mode FE		N	N	Y	N	N	Y
Commodity FE		N	N	Y	N	N	Y
Year FE		Y	Y	Y	Y	Y	Y
Observations		36848	36848	36848	36848	36848	36848
Adj. R-squared		0.06	0.10	0.24	0.34	0.34	0.67

*Notes:* Figures in parentheses are robust standard errors; constant is not reported

\*\*: significant at 1%

\*: significant at 5%



**Table 7. Sample Selection Estimation of Transport Cost Function**

Dependent variable: Log of transport cost in yen per ton/hour

Variables	<u>Model 1</u>		<u>Model 2</u>		<u>Model 3</u>	
	Probit	Selection corrected	Probit	Selection corrected	Probit	Selection corrected
	(1)	(2)	(3)	(4)	(5)	(6)
Log of distance	0.375** (0.008)	-1.007** (0.007)	0.786** (0.015)	-0.860** (0.016)	0.222** (0.021)	-1.124** (0.013)
Log of distance squared			-0.192** (0.006)	-0.067** (0.007)	-0.075** (0.009)	-0.004 (0.006)
Log of commodity flows	0.334** (0.001)		0.335** (0.001)		0.363** (0.002)	
Inverse Mills ratio		1.203** (0.009)		1.213** (0.009)		0.891** (0.011)
Sending prefecture FE		N		N		Y
Arriving prefecture FE		N		N		Y
Transport mode FE		N		N		Y
Commodity FE		N		N		Y
Year FE		Y		Y		Y
Observations	378567	36879	378567	36879	378567	36879

*Notes:* Figures in parentheses are bootstrapped standard errors with 50 replications; constant is not reported.

\*\* : significant at 1%

\* : significant at 5%

**Table 8. Sample Selection Estimation by Transport Mode**

Dependent variable: Log of transport cost in yen per ton/hour

Variables	Mode	<u>Air</u>	<u>Railroad</u>	<u>Ship</u>	<u>Truck</u>
		(1)	(2)	(3)	(4)
Log of distance		0.747** (0.246)	-0.659** (0.109)	-1.474** (0.075)	-0.847** (0.014)
Log of distance squared		-0.391** (0.071)	-0.073* (0.041)	0.201** (0.035)	-0.082** (0.007)
Inverse Mills ratio		0.745** (0.033)	0.509** (0.041)	0.416** (0.059)	0.980** (0.011)
Sending prefecture FE		Y	Y	Y	Y
Arriving prefecture FE		Y	Y	Y	Y
Commodity FE		Y	Y	Y	Y
Year FE		Y	Y	Y	Y
Observations		34328	69000	103440	171799
Uncensored observations		1669	1857	1057	32296

*Notes:* Figures in parentheses are bootstrapped standard errors with 50 replications; constant is not reported; log of commodity flows is included in the probit estimation.

\*\* : significant at 1%

\* : significant at 5%

**Table A1: Prefecture and Region**

<b>Region</b>	<b>Prefecture</b>				
Hokkaido	Hokkaido				
Tohoku	Aomori	Iwate	Miyagi	Akita	Yamagata
	Fukushima				
Kanto	Ibaragi	Tochigi	Gunma	Saitama	Chiba
	Tokyo	Kanagawa			
Chubu	Niigata	Toyama	Ishikawa	Fukui	Yamanashi
	Nagano	Gifu	Shizuoka	Aichi	
Kinki	Mie	Shiga	Kyoto	Osaka	Hyogo
	Nara	Wakayama			
Chugoku	Tottori	Shimane	Okayama	Hiroshima	Yamaguchi
Shikoku	Tokushima	Kagawa	Ehime	Kouchi	
Kyushu	Fukuoka	Saga	Nagasaki	Kumamoto	Ohita
	Miyazaki	Kagoshima	Okinawa		