ON THE COSTS OF INTERNATIONAL TRADE AND INVESTMENT

TRAN THI THU TRANG

DOCTORAL DISSERTATION

GRADUATE SCHOOL OF ECONOMICS
HITOTSUBASHI UNIVERSITY

2023

ON THE COSTS OF INTERNATIONAL TRADE AND INVESTMENT

by

TRAN THI THU TRANG

Submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philoshophy

in

Economics

Graduate School of Economics

Hitotsubashi University

March 2023

Kunitachi, Tokyo

©2023 by Tran Thi Thu Trang

All Rights Reserved

Preface

In 1986, the Vietnamese government made a decision that changed the lives of its 62 million citizens. That decision was to transform the economy toward a market-oriented system and to gradually integrate its economy into the world's economy. Since then, the country's Gross Domestic Product (GDP) has doubled every decade and the average income per person as of 2020 quintupled its level from the 1980s. Since this development is attributed to export and investment growth, Vietnamese economic growth during these three decades could be described as the typical textbook case of international trade theory.

Indeed, in my youth, like other young Vietnamese, I was educated to prepare for the integration of the economy. International economics has become my focus during my undergraduate and graduate studies. Theories tell us that opening an economy requires a lot of effort to reduce all trade barriers, including tariffs, non-tariff barriers, transportation costs, information costs, etc. The globalization trend, where the case of Vietnam is an example, has dramatically reduced tariffs and non-tariff barriers. The improvement in transportation and information technology has significantly reduced other components of trade costs. And yet the recent comeback of protectionism and the skyrocketing rise in shipping costs around the globe have shown the relevance of studying the topic.

In particular, this dissertation tries to answer three questions. First, how much is the consumer's value of timeliness? Second, how often do sellers send their shipments to the buyer? Third, where do multinational firms locate their subsidiaries? The first two questions focus on the relationship between transportation costs and trade flows, while the third question focuses on the role of information costs on Foreign Direct Investment.

The first question is answered by utilizing the response of shipping companies to fuel shocks during the 2000s. The consumer's time value is modeled into the Armington-Anderson trade model, and shipping companies choose the optimal delivery time in response to fuel prices. Data on container ship movement is used to estimate the delivery time. By focusing on one mode of transport and creating a suitable instrument variable for the delivery time, the study can reliably estimate the time value. The results show that time costs are present even after controlling for the quality of goods. Consumers are indeed sensitive to delivery times.

The second question is explored by considering the storage and opportunity costs in the profit function of intermediate goods buyers. The optimal shipment frequency (and size) is achieved by balancing fixed costs per shipment, iceberg trade costs, and sellers' and buyers' productivity. Considering the buyer heterogeneity, the model provides some predictions on the number of buyers in relation to the decision of shipment frequency. These predictions are tested using the bill of lading dataset. The results show that the effects of iceberg trade costs and fixed shipment costs on trade volume are mainly from an increase in shipment frequency rather than shipment size. Interestingly, more than half of the increase in shipment frequency is from an increase in buyers.

The final question is examined through the decision on the location of manufacturing plants and wholesale subsidiaries of Japanese firms in European countries. The premise is that regional information costs can be reduced if more service firms connect these regions. The results show that Japanese firms tend to locate their wholesale subsidiaries in a region with many other Japanese service firms, even after controlling for industry and financial group agglomeration.

Manufacturing plant locations, however, do not always follow the same patterns. This shows that information is particularly important for wholesale subsidiaries, possibly to reach potential customers through other Japanese service firms.

The answers to these questions reinforce that international trade and investment costs are high. The world is not flat. The existence of trade costs motivates different agents to organize their operations to increase efficiency and reach more consumers. The studies have become possible partly thanks to the availability of micro data that offers a detailed view of the behavior of different economic agents. Hopefully, this micro data can be improved in accuracy and available to researchers with the improvement of technology and policies toward research activities.

Acknowledgements

I am indebted to Professor Jota Ishikawa, who has been supportive since I came to Japan twelve years ago. I owe Professor Yoichi Sugita for his insightful advice on the content of this dissertation and his support in finishing the final steps. I thank Professor Tomiura Eiichi, Professor Kensuke Teshima, and other members of the International Trade Seminar at Hitotsubashi University for their comments on various versions of the studies used in this dissertation. I thank Professor Takuma Matsuda and Mr. Seijiro Morimoto at Japan Maritime Center, whose discussions helped form the idea of the second chapter. I thank Professor Toshiyuki Matsuura at Keio University for his comments on the previous version of the fourth chapter. I thank Professor Taiji Furusawa for his advice on various research activities.

My gratitude goes to my dad, my sister, and Renee, who are always by my side and give me enormous encouragement. Finally, I thank my mom, who looks down on me from heaven.

Contents

List of Figures

List of Tables

Abbreviations

B/L	**B**ill of **L**ading
CES	**C**onstant **E**lasticity **S**ubstituition
CIF	**C**ost **I**nsurance **F**reight
EU	**E**uropean **U**nion
FDI	**F**oreign **D**irect **I**nvestment
FOB	**F**ree **O**n **B**oard
IMO	**I**nternational **M**aritime **O**rganization
IV	**I**nstrumental **V**ariable
knots	**n**autical **m**iles **p**er **h**our
MNE	**M**ulti-**N**ational **E**nterprises
OECD	**O**rganization for **E**conomics **C**o-operation and **D**evelopment
OLS	**O**dinary **L**east **S**quare
PPML	**P**suedo **P**oisson **M**aximum **L**ikelihood
TEU	**T**wenty-foot **E**quivalent **U**nit
UNCTAD	**U**nited **N**ations **C**onference on **T**rade and **D**evelopment
2SLS	**T**wo **S**tages **L**east **S**quare

Chapter 1

Introduction

1.1 Research Background

In some parts of the world today, consumers can shop from home by surfing the internet. After a click to buy, the goods will be delivered to their doors, sometimes within a day. This seamless movement of goods and information is thanks to the reduction in transportation costs. Figure 1.1 shows the trend of international transportation costs between 1995 and 2016. While there are some variations in types of products and country pairs, on average, the costs have declined by 17 percent during this period.

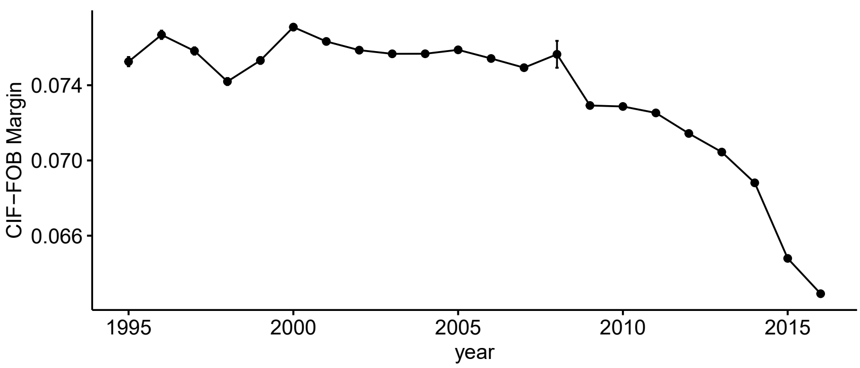


Figure 1.1: Transporation Costs

Source: Author's calculation based on OECD's International Transport and Insurance Costs of Merchandise Trade database

Political trade barriers are also decreasing, with more countries joining the global economy. The average tariff rates have decreased from about 15 percent at the beginning of the 1990s to about five percent in the 2000s.

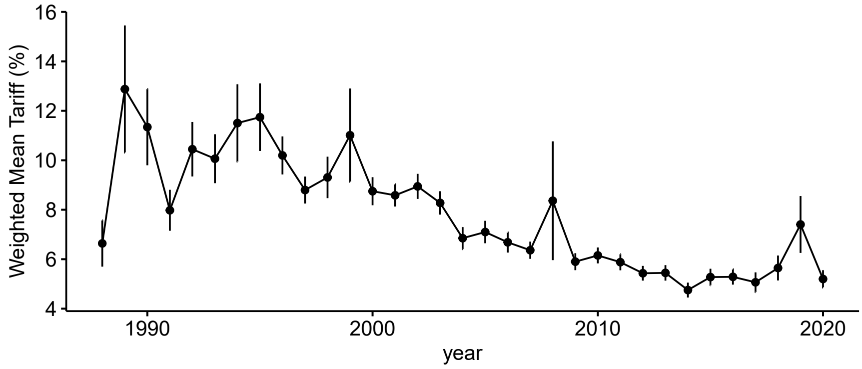


Figure 1.2: Weighted Mean Tariff Rate

Source: World Bank's estimates from data in the World Integrated Trade Solution System, based on data from the United Nations Conference on Trade and Development's Trade Analysis and Information System (TRAINS) database and the World Trade Organization's (WTO) Integrated Data Base (IDB) and Consolidated Tariff Schedules (CTS) database.

Despite the declining trend of transportation costs and tariffs, trade costs overall are still prevalent. Anderson and Van Wincoop (2004) estimates that the total trade costs can be as large as 170 percent tariff equivalent. Among these, transportation costs are about 21 percent tariff equivalent, including freight costs (12 percent) and time costs (9 percent). The border-related trade barriers are about 44 percent tariff equivalent, and local retail and wholesale distribution is approximately 55 percent tariff equivalent. Not to mention, some shocks always increase different portions of trade costs. For example, the rise of the US's protectionism has pushed the weighted average tariff of the US from around 1.6 percent to a record high of 13.8 percent in 2019. Another example is the rise of container freight during the second half of the COVID-19 pandemic. The price of sending one container from East Asia to the US's west coast rose eight times from 2000 USD/FEU at the beginning of 2020 to about 8000 USD/FEU in 2022.[1]

[1]FEU is short for a forty-feet container equivalent unit. It's often used in container shipping, in addition to TEU (twenty-feet container equivalent unit). The price mentioned here is from the container spot freight index for this particular route published monthly by Shanghai Shipping Exchange.

While trade costs have been studied extensively in the trade literature (Anderson and Van Wincoop, 2004), there are still many areas for more exploration. One area is the measurement of time costs.[2] The literature is started from the works of Hummels (2007) and Hummels and Schaur (2013). They utilized the variation of exporters to the US in choosing between air or ocean transport to estimate the time premium of air transport over ocean transport. Because goods transported by air and ocean are potentially different in quality and finding an appropriate instrument is not easy, these estimates can be improved by considering one mode of transportation. The second chapter of this dissertation estimates the value of time in international container trade.

Concerning the value of time for consumers, final producers also find themselves sensitive to time costs. For producers, time costs are related to the cost of storing intermediate inputs and the associated opportunity costs of idle capital. If the costs per shipment for these intermediate inputs are minimal, they can have them shipped continuously, eliminate storage costs, and reduce opportunity costs. While some multinational firms have succeeded by integrating the logistics system into their business models, most firms still have to deal with shipment costs and balance shipment frequency and size.

Studies on shipment frequency and shipment size have been explored by many authors, including Hornok and Koren (2015a) and Kropf and Sauré (2014). These models share a similar concept that consumers value the timeliness of shipment, but each shipment incurs a fixed cost. The optimal frequency ultimately depends on the trade-off between this fixed cost and the cost of holding inventory. On the other hand, trade literature also emphasizes the importance of firms' heterogeneity in shaping trade patterns (Melitz, 2003; Chaney, 2008; Bernard, Moxnes, and Ulltveit-Moe, 2018). The natural extension is to investigate how adding buyers'

[2] Perhaps, there is little dispute about whether customers value time. For example, Amazon Prime, an Amazon membership service offering free two-day shipping, launched in 2005. It quickly took off after 2011 and now has more than 200 million members across 19 countries. As of 2022, customers can enjoy one-day shipping services in some big cities.

heterogeneity affects sellers' decisions on shipment frequency and size in response to the change in trade barriers. This will be explored in the third chapter.

Finally, the fourth chapter departs from transportation costs to investigate information costs[3] in the location decision of multinational firms. Recent literature has pointed out the importance of multinational firms in shaping trade patterns through foreign affiliates.[4] It is crucial to study the impact of trade costs not only on trade but also on investments.

A vast body of literature has studied the role of trade costs in FDI's location decision (Alfaro and Chen, 2018). Trade costs generally decrease the flow of investment,[5] except for horizontal FDI.[6] The FDI literature distinguishes itself from the trade literature in the role of knowledge capital. The traditional assumption is that multinational firms' advantage in the free use of this joint input among subsidiaries is the motivation to invest instead of export. However, recent studies by Keller and Yeaple (2013) show that FDI's knowledge content does follow the law of gravity. Proximity is crucial for sharing knowledge.

While many papers have examined the benefit of proximity in the context of manufacturing plants[7] or headquarters' locations,[8] not many papers have examined the effects on wholesale subsidiaries.[9] Furthermore, Belderbos, Du, and Goerzen (2017) and Goerzen, Asmussen, and Nielsen (2013) point out the importance of city connectivity in reducing communication costs. The fourth chapter applies this idea to examine the relative locations between manufacturing plants

[3]Anderson and Van Wincoop (2004) This information cost barrier is categorized as part of the border-related trade barrier, which is about 6 percent tariff equivalent. Their calculation is based on the work of Rauch and Trindade (2002).

[4]See Bernard, Jensen, and Schott (2005) for example.

[5]These include vertical FDI (Helpman, 1984), complex FDI (Yeaple, 2003), export platform FDI (Ekholm, Forslid, and Markusen, 2007), merger and acquisitions (Head and Ries, 2008).

[6]Markusen (1984) is one of the first studies.

[7]See Head, Ries, and Swenson (1995), Head and Mayer (2004), and Defever (2012) for example. These papers mostly focus on agglomeration benefits.

[8]See Strauss-Kahn and Vives (2009), Belderbos, Du, and Goerzen (2017), and Goerzen, Asmussen, and Nielsen (2013).

[9]Anderson and Van Wincoop (2004) estimates the local distribution costs are about 55 percent tariff equivalent. His estimates are based on the works of Bradford (2003).

and wholesale subsidiaries.

1.2 Chapters Preview

Chapter 2. The Value of Time in International Container Trade

This chapter uses a spike in fuel prices during the first ten years of the 2000s to estimate the value of timeliness in international container trade. The theoretical framework shows fuel price affects trade through speed adjustment of the shipping sector (slow steaming). The ship can adjust delivery days and freight costs when fuel prices increase. The change in delivery days affects consumers as if there is a decrease in the quality of goods.

The evidence of slow steaming is confirmed using data on ship movement. Estimates of the elasticity of delivery days with respect to fuel prices were shown to depend on ship sizes. On average (for a ship with a size of $4,437$ TEU), a 10 percent increase in fuel prices increases delivery days by five percent. This translates to a delay of one day on the trade route between East Asia and North America (with an average delivery time of 20 days). The value of timeliness is estimated by an Instrumental Variables (IV) method. The IV considers the response of ships with respect to fuel prices and the composition of ship sizes for different trade routes. The trade elasticity with respect to delivery days is estimated to be about -0.0814. A one percent delay in delivery is equivalent to an additional one percent tariff. On the East Asia - North America trade route (20 days), one delayed day adds five percent tariff equivalent.

The study also shows geographical differences in ships' responses to fuel prices. Slow steaming does not happen in regional trade routes because it's easier to adjust schedules to respond to high fuel prices due to shorter distances. Among inter-regional trade routes, the slow steaming effects in the North-South group are twice as high as in the East-West group. It is because ship sizes on the East-West route are slightly bigger than that of the North-South, and the fuel consumption coefficient is higher for larger ships.

In general, the estimates of time value are less than the literature. This result by no means discredits the importance of time costs in international trade. On the contrary, it shows that time costs are present even after controlling for the quality of goods. Consumers are indeed sensitive to delivery time.

The study contributes to the literature in several ways. First, it is the first study to discuss a different channel for fuel prices to affect container trade: that is, through speed adjustment in the shipping sector. Second, it is the first study to provide a credible instrumental variable to control for the quality elements in delivery days. Third, it shows the potential of using new types of data (ship movement data) in exploring the patterns of trade costs in international trade.

Chapter 3. Trade Costs and Different Margins of Trade

This chapter investigates the relationship between trade costs and different margins of trade. In the theoretical model, iceberg trade costs affect trade directly through sales and indirectly through shipment frequency and the number of buyers. Shipment costs' effects on trade are only through adjusting shipment frequency. Sellers with a small number of shipments may reduce shipment size when iceberg trade costs decrease if the adjustment through shipment frequency dominates the sales effects. The theoretical models also provide testable hypotheses, which are examined using the Bill of Lading (B/L) data set.

The empirical analysis is carried out by estimating the gravity-like equation for the seller level and the seller-buyer level. The empirical results confirm that trade barriers reduce firms' trade volume, shipment frequency, and the number of buyers. The shipment costs may increase the average shipment size of a seller when the number of shipments is low, as predicted from the theoretical model.

The results also highlight the importance of buyer margins. While the effects of iceberg trade costs and fixed shipment costs on trade volume are mainly from an increase in shipment frequency rather than shipment size, more than half of an increase in shipment frequency is from an increase in the number of buyers. This new insight has not been explored in the literature on shipment frequency. It is mainly because the decision on shipment frequency comes after the buyers' decision.

The study contributes to the literature in several ways. First, it is the first study to examine the role of buyer heterogeneity in the decision of shipment frequency (and shipment size). More than half of an increase in shipment frequency is from an increase in the number of buyers. Second, it shows the potential of using the bill of lading data set in examining the relationship between exporters and importers at the firm level.

Chapter 4. Trade Costs and Multinational Firms' Location Decision

This chapter examines the decision of manufacturing and wholesale subsidiaries' location of Japanese manufacturing firms in the European Union (EU). It highlights the difference in geographical patterns between wholesale subsidiaries and manufacturing plants. The new insight is that wholesale affiliates tend to spread to a region with a strong presence of other Japanese service firms' affiliates. While more clustered services can create agglomeration-like effects for firms in certain regions, the connectedness potentially reduces information costs among regions. Hence, "information sharing" among multinational firms' subsidiaries does not need to happen in close proximity.

Regarding the relative location among firms' subsidiaries, manufacturing plants and wholesale subsidiaries of the same industry and the same firms are not located near each other. On the other hand, manufacturing firms locate their plants in regions with a strong presence in their financial group's wholesale subsidiaries, which stresses the importance of sharing information beyond industry agglomeration. The study also confirms the agglomeration force for manufacturing plants within the industry and financial groups. The industry agglomeration effects are less than half of the within-financial-group agglomeration effects.

The study contributes to the literature in several ways. First, it confirmed the importance of "information sharing" in the location decision of Japanese multinational firms (Blonigen, Ellis, and Fausten, 2005). Second, it highlights the difference in location patterns between manufacturing plants and wholesale subsidiaries.

To summarize, the remaining of the dissertation is as follows: chapter 2 discusses the value of time in international container trade, chapter 3 explores the role of buyers' heterogeneity in the sellers' decision of shipment frequency and shipment size, and chapter 4 discusses the role of information in firms' decisions of their foreign affiliates. Chapter 5 concludes.

Chapter 2

Estimating the Value of Time in International Container Trade

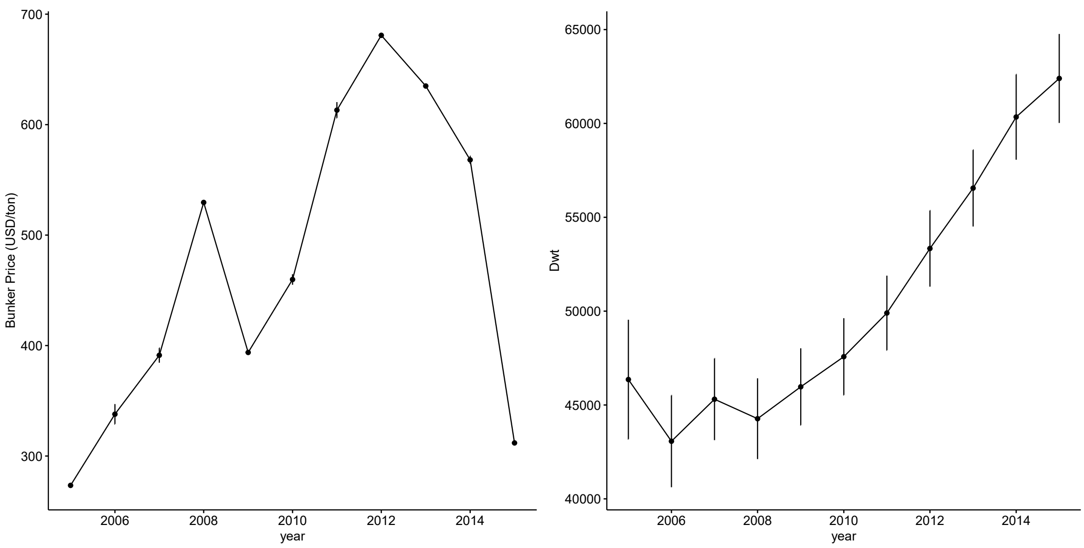
2.1 Introduction

With the improvement of transportation technology and logistics systems, it takes less time to transport cargo (Hummels, 2007). While air transport has grown more and more due to its dominantly shorter transit time, ocean shipping still makes up more than eighty percent of global trade volume. Container ships are the fastest among different types of cargo ships, and their speeds have improved thanks to more powerful engines in bigger ships (Stopford, 2009). While there has been a lot of improvement in ship designs, most ships still rely on the combustible engine to fuel their operations. Because fuel costs make up about half of the ships' operating costs (Stopford, 2009), saving fuel costs is crucial for shipping companies to maintain their profits.

However, fuel prices[1] increased dramatically in the first ten years of the twenty-first century. Figure 2.1a shows the trend of fuel prices for the world's

[1] Fuel prices are measured by the average prices of high-sulfur fuel oil (HSFO 3.5% sulfur) and very low-sulfur fuel oil (VLSFO), which is commonly used in ships' main engines. These fuels are usually purchased by ships in ports on their trips.

major areas. The average price almost doubled between 2006 and 2013. Many major shipping companies started to run their ships at lower speeds (i.e., slow steaming) to save fuel consumption.[2] Because ships with different sizes on different routes react differently in response to fuel price spikes, this phenomenon serves as a quasi-natural experiment to estimate the value of timeliness in international container trade. This is the first study to estimate the value of time in container trade using the slow steaming behavior of shipping companies.



(a) Fuel Price Trend   **(b)** Average Ship Size Trend

Note: Average ship size is measured by the average deadweight ton (DWT) of ships operating on 35 routes during this period. Data on ship movements and ship sizes are from IHS Markit. Fuel prices are calculated by taking the simple average fuel prices sold at major ports during this period. Data on fuel prices are from Clarkson Shipping Intelligence.

While reducing speed can save fuel consumption, it may make a round trip longer and affect the capacity of shipping (Stopford, 2009). However, this can be compensated by having more ships or bigger ships. In fact, more large ships were rolled out during this period. Figure 2.1b shows the trend of average ship size for all container ships running between 2005 and 2015. Average ship size[3]

[2] Ferrari, Parola, and Tei (2015) suggests that a reduction in 11 percent of the speed (from 48km/h to 40km/h) may halve the fuel cost of a mega container ship. While ships are often designed to run at certain speeds, shipping companies still have room to tweak these designed speeds by simply slowing down or, in some cases, degrading ships' engines (Psaraftis and Kontovas, 2013).

[3] Ship size is measured by the maximum weight (deadweight ton) a ship can carry. Specifi-

increased almost 30 percent in 2016 compared with 2006. This created a supply surplus situation and allowed shipping companies to reduce speed without worrying about lack of capacity (Stopford, 2009).

There have been a few papers studying the value of timeliness. Hummels and Schaur (2013) is a pioneer in estimating the value of time in international trade. They utilized the variation of exporters to the US in choosing between air or ocean transport to estimate the time premium of air transport over ocean transport. Because goods transported by air and ocean are potentially different, this paper aims to estimate the value of time using variations in different container routes responding to a spike in fuel prices. Considering the value of time in one mode of transportation reduces the possibility of contaminating delivery days with the choice of goods quality (i.e. "apples" vs. "apples" instead of "apples" vs. "oranges").

This study relates to the literature investigating the cost of shipment infrequency. A similar mechanism is considered: longer delivery time or infrequent shipment tends to reduce trade because consumers prefer on-time delivery (Hornok and Koren, 2015a). However, there is no explicit consideration of firms' decision on the shipping mode to save time or the number of shipments to save per-shipment fixed costs. Instead, speed adjustment in response to fuel prices is investigated in this study.

This study is also related to papers studying economic geography changes that affect trade. The gravity literature has used distance as a proxy for trade cost. However, it was pointed out the estimates of distance elasticity of trade cost were unrealistically high and have not dropped over time despite globalization.[4] Borchert and Yotov (2017) finds that the effect of distance on trade has

cally, it is the weight a ship can carry when loaded to its marks, including cargo, fuel, freshwater, stores, and crew. For containerships, TEU(twenty-feet equivalent unit) is often used. TEU is based on a measure of a standardized metal container of twenty feet in length. Thanks to standardization, it can also be used in railroad transportation (Stopford, 2009).

[4]See Anderson and Van Wincoop (2004) for a thorough discussion.

fallen by 9.34 percent over the period 1986–2006 by using country fixed-effects to control for the home bias effects. Hummels, Lugovskyy, and Skiba (2007) and Feyrer (2019) exert that the change in transportation technology (i.e., the rise of air freight) has shaped global trade and the world economy (changing distance elasticity of trade). Although previous papers have pointed out that the change in fuel prices may affect trade through transportation costs, very few papers have investigated the change in shipping companies' behaviors.

This study also relates to studies on the effects of oil prices on trade. Von Below and Vezina (2016) and Mirza and Zitouna (2009) investigated the effects of fuel prices on trade using a gravity equation framework. Their core assumption is that the interaction of oil prices and distance affects trade only via shipping costs.[5] Compared to these papers, this study shows a new channel for fuel prices to affect trade, that is, through speed reduction.

This study is close to the paper by Brancaccio, Kalouptsidi, and Papageorgiou (2020). They show that fuel price may affect the opportunity of a bulk carrier to get a match with shippers. Higher fuel prices make the ship less likely to move to other areas to look for a new contract and, therefore, is locked in locally. While this study also uses ship movement data to investigate ships' responses to high fuel prices, speed reduction in the container sector is the main focus.

In summary, this study contributes to the literature in two aspects. First, a new channel for fuel prices to affect trade is addressed. When fuel prices increase, ships reduce speeds to save fuel costs. As a result, it increases the delivery time at sea and potentially reduces trade. Second, the value of timeliness is estimated by constructing an instrumental variable (IV) containing information about fuel prices for delivery days. The evidence on slow steaming is confirmed using data

[5]Because oil prices may be endogenous to trade through countries' demand for foreign goods, the instrument of oil prices is the yearly number of conflicts in The Organization of the Petroleum Exporting Countries (OPEC) to control for the supply side (Von Below and Vezina, 2016). Mirza and Zitouna (2009) use oil prices, distance, and transport modes as instruments for freight rate.

on ship movement. Estimates of the elasticity of delivery days with respect to fuel prices were shown to depend on ship sizes. Using data on container trade value, the trade elasticity with respect to delivery days is estimated to be about -0.0814.[6] A one percent delay in a delivery day is equivalent to an additional one percent tariff. On the East Asia - North America trade route (20 days), one delayed day adds five percent tariff-equivalent.

The study is structured as follows: Section 2 illustrates the conceptual framework; section 3 describes the estimation strategy; section 4 describes data on ships' voyages and seaborne trade; section 5 presents estimation results and discussion; section 6 shows various robustness checks; and section 7 concludes.

2.2 Conceptual Framework

2.2.1 Consumer Preference of Timeliness

This study is based on Anderson Armington's model with time preference in consumer utility.[7] A representative agent in country j consumes goods from country i among N countries with delivery days D_{ij} with a constant elasticity (CES) utility function

$$U_j = \left(\sum_{i}^{N} D_{ij}^{-\tau} q_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)} \tag{2.1}$$

where τ is a parameter that shows the evaluation of consumers to delivery days (the higher value indicates a higher valuation of delivery days). The term $(\sigma - 1)/\sigma$ is the elasticity of substitution with the assumption that $\sigma > 1$. The variable D_{ij} is the number of days needed to deliver from country i to country

[6]However, it is worth noting that the estimates of the elasticities of trade with respect to delivery days are very sensitive to the functional forms of delivery days in the utility function. In the case of log-linear form as in Hummels and Schaur (2013), the trade elasticity with respect to days is $0.0087Day$. The tariff equivalent of one day is 0.26 percent.

[7]Hummels and Schaur (2013) assumes an exponential functional form $e^{-\tau D_{ij}}$. This will be shown in the robustness check section.

j. The variable q_{ij} is the quantity of goods consumed in the country j shipped from country i.

Assuming country j's income to be Y_j and the price of goods from country i in country j to be p_{ij}, and the budget constraints to be $\sum_i p_{ij}q_{ij} \leq Y_j$, the demand q_{ij} of country j for goods delivered from country i is the solution to the utility maximization

$$q_{ij} = \left(\frac{p_{ij}}{D_{ij}^{-\tau}}\right)^{-\sigma} \frac{Y_j}{P_j^{1-\sigma}} \tag{2.2}$$

where the price index is

$$P_j = \left(\sum_i D_{ij}^{-\tau\sigma} p_{ij}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}.$$

An increase of one percent on delivery day reduces demand by $\sigma\tau$.

Assuming iceberg trade costs T_{ij}, the price consumers in j pay for imports from country i can be stated as $p_{ij} = T_{ij}p_i$ where $T_{ii} = 1$ and $T_{ij} > 1$. The value of exports from i to j is

$$X_{ij} \equiv p_j q_{ij} T_{ij} = D_{ij}^{-\tau\sigma}\left(\frac{p_j T_{ij}}{P_j}\right)^{1-\sigma} Y_j. \tag{2.3}$$

Applying the goods market clearing condition for each commodity $Y_j = \sum_{i=1}^N X_{ij}$ and denote the world's income as $Y^W \equiv \sum_{j=1}^N Y_j$, the share of country j income in the world's income as $\eta_j \equiv Y_j/Y^W$, the bilateral trade is

$$X_{ij} = \frac{Y_i Y_j}{Y^W}\left(\frac{T_{ij}}{\Pi_i P_j}\right)^{1-\sigma} D_{ij}^{-\tau\sigma} \tag{2.4}$$

where the price indexes are

$$\Pi_i = \left(\sum_{j=1}^N \frac{T_{ij}}{\Pi_i P_j} D_{ij}^{-\tau\sigma}\eta_j\right)^{\frac{1}{1-\sigma}}$$

and

$$P_j = \left(\sum_{i=1}^{N} \frac{T_{ij}}{\Pi_i P_j} D_{ij}^{-\tau\sigma} \eta_i \right)^{\frac{1}{1-\sigma}}.$$

It is similar to Anderson and Van Wincoop (2003) except for the inclusion of the terms $D_{ij}^{-\tau\sigma}$. The term $D_{ij}^{-\tau\sigma}$ expresses the relationship between delivery days and bilateral trade. The longer the delivery days, the higher the perceived price for consumers, hence the lower demand and lower trade.[8] Without this term, the expression of bilateral trade is the same as in Anderson and Van Wincoop (2003).

Trade cost is assumed to have two portions:[9] unit freight cost FR_{ij} and other trade cost t_{ij}, so the trade cost is $T_{ij} = t_{ij} + FR_{ij}/p_j$. Denote $\Phi_i \equiv \frac{Y_i}{\Pi_i^{1-\sigma}}$ and $\Phi_j \equiv \frac{Y_j}{P_j^{1-\sigma}}$, the trade equation (2.4) becomes

$$X_{ij} = \Phi_i \Phi_j T_{ij}^{1-\sigma} = \Phi_i \Phi_j D_{ij}^{-\tau\sigma} \left(t_{ij} + \frac{F_{ij}}{p_j} \right)^{1-\sigma}. \tag{2.5}$$

The elasticity of trade with respect to fuel price f_{ij} is then

$$\frac{\partial \ln X_{ij}}{\partial \ln f_{ij}} = (1-\sigma) s_{ij}^{FR} \frac{\partial \ln FR_{ij}}{\partial \ln f_{ij}} - \tau\sigma \frac{\partial \ln D_{ij}}{\partial \ln f_{ij}} \tag{2.6}$$

where $s_{ij}^{FR} \equiv \frac{FR_{ij}}{p_j t_{ij} + FR_{ij}}$ is the share of freight costs in import price.

Fuel prices affect trade through two channels: the freight cost FR_{ij} and delivery days D_{ij}. The first channel has been studied in the literature, but the second channel has not. In other words, the second term has been treated as zero ($\partial \ln D_{ij} / \partial \ln f_{ij} = 0$) in previous studies on the effect of high fuel prices on trade.

The first channel corresponds to the first term of the right-hand side of equation (2.6). This term has three parts: the degree of substitution among exporters'

[8] This is not the same but relates to the idea of infrequent shipment as in Hornok and Koren (2015a), where an infrequent shipment is perceived as relatively more costly. Longer delivery days can cause infrequent shipment if the means of transport are limited.

[9] This is similar to Hummels, Lugovskyy, and Skiba (2007).

goods, the share of freight costs in import prices, and the fuel price elasticity of freight. Because $\sigma > 1$, the fuel price elasticity of freight costs $\epsilon_{F,f}$ is positive. A one percent increase in fuel prices decreases trade by $(1-\sigma)s_{ij}^{F}\epsilon_{F,f}$ percent. The magnitude of this impact depends on the value of σ and s_{ij}^{FR}. The literature has reported a wide range of estimates for σ, mostly between three to ten (Anderson and Van Wincoop, 2004; Hummels and Schaur, 2013). Trade literature also provides estimates on the share of freight costs in the imported price, which can be calculated using direct freight cost data or the CIF-FOB margin.[10] The trade-weighted average share of direct transport cost in FOB price for the US is about four percent, translated to about 3.8 percent compared to CIF price.

The estimates for $\epsilon_{F,f}$ usually depend on distance. However, previous studies provide very different numbers. For example, Von Below and Vezina (2016) estimated that the trade elasticity with respect to oil prices is between -0.187 and -0.271 multiplied by the natural logarithm of distance. Mirza and Zitouna (2009)'s estimate depends on $ln(distance)$ by a factor of -0.017.[11] Assuming a median distance of $4,675km$, the oil price elasticities can range from 0.1 to 2.3.[12] On the other hand, papers focusing on the container sector provide an estimate of about 0.2 (Hummels, Lugovskyy, and Skiba, 2007; UNCTAD, 2010). In summary, the second channel can be between -0.07 and -0.02 (when oil price elasticity of 0.2 and σ is between three and ten).

The second channel corresponds to the second term of the right-hand side of equation (2.6). In general, higher fuel prices induce ships to lower speed, which results in longer delivery days[13], that is $\partial lnD_{ij}/\partial f_{ij} > 0$. The more sensitive

[10] See Anderson and Van Wincoop (2004) for a thorough discussion.

[11] They calculate the fuel price elasticity of trade by multiplying the fuel prices elasticity of freight costs ($0.011ln(distance)$) by the freight costs elasticity of trade (-1.545).

[12] It is not quite straightforward to compare these two papers because they use different data from different periods. Von Below and Vezina (2016) uses all non-landlocked countries' aggregate trade data from 1962 to 2004. To control the endogeneity of freight costs to trade, Mirza and Zitouna (2009) uses the US's bilateral imports and freight charges data at the Standard International Trade Classification, Revision 4 (SITC4) product level (over 1000 products) from 1974-2001.

[13] Details are explained more in section 2.2.2.

consumers are to time (higher τ) or the more substitutable the exporters' goods (higher σ), the more negatively this channel affects trade.

2.2.2 Slow Steaming

Slow steaming refers to reducing speed to less than the designed speed to save fuel costs. Ferrari, Parola, and Tei (2015) suggests that the share of fuel cost in the operating cost of container ships is about 50 percent (the number depends on the size of the ship). During the high oil price period in the 2000s, speeds were reduced by about 30 percent, which may save 50 percent of fuel consumption.[14]

Speed reduction has become a trend during the 2000s. This can be done either at the design or operation steps. Speed reduction at the operation step is often called "slow steaming" which may cause some problems for engine maintenance if the speed is too low. To deal with this issue, the engine may be modified to drastically reduce speed ("super slow steaming"). The designed speeds have been reduced from 61 km/h in the 1960s to about 45 km/h (a drop of almost 33 percent).[15] This is partly due to the expansion of the fleet (number of ships). With a bigger fleet, the ship does not have to rush back to make another round of trips. It is also important to notice that the speed reduction during the 2000s is not paralleled with the increase in the number of ships but rather with the increase in total freight capacity (larger average ship size). This was reported to affect port-to-port transit times, and service quality.[16]

While slow steaming was observed for all ships, it is the most prevalent in the container sector. Figure 2.2 plots the ton-km-weighted average designed speeds and operated speeds for container ships, bulk carriers, and tankers. Container

[14]See Psaraftis and Kontovas (2013) and references therein.

[15]These two numbers are taken from Stopford (2009) and IMO (2015). The commonly used measure of ship speeds in ocean shipping is "knots" (nautical miles per hour). The original numbers are in knots but were converted to km/h. A knot is equal to 1.852 km/h. A ship traveling at 1 knot along a line of latitude travels approximately one minute of geographic latitude in one hour.

[16]See Ferrari, Parola, and Tei (2015) and references therein.

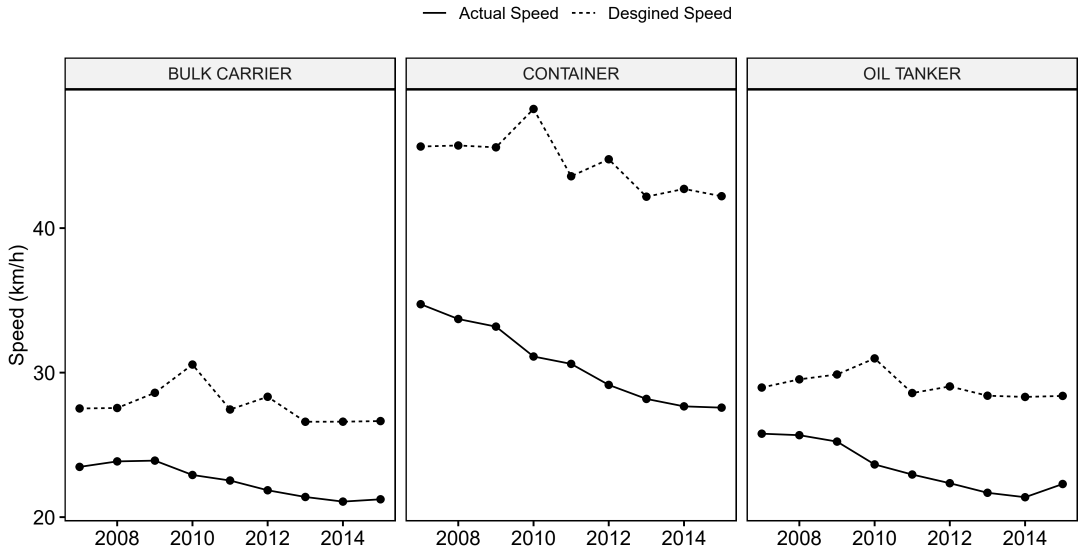


Figure 2.2: Average Speed of Different Ship Types

Note: Average speed is calculated as the ton-km-weighted average for operating speed at sea for different categories of ships within a sector reported in (IMO, 2015).

ships transport mainly manufactured products and intermediates, bulk carriers carry industrial materials, and tankers carry crude oil or gas. Due to the nature of their products, container ships tend to have higher designed and operating speeds. They also consume more fuel and have more incentive for slow steaming. The average operating speed of container ships was reduced by almost 21 percent during the 2007-2015 period, much higher than other sectors (12 percent for bulk carriers and 14 percent for oil tankers). Therefore, slow steaming in the container sector is the main focus of this paper.[17]

To formalize the analysis, assuming that the total cost function of a shipping service on route ij is

$$TC_{ij} = \left(f_{ij}\lambda v^{\alpha}\right)\frac{q_{ij}d_{ij}}{Cv} + FC_{ij} \tag{2.7}$$

[17]While shipping companies can freely adjust speed, they may need to operate at a certain speed under the charter party contract with the ship owner. In this case, fuel is paid by the ship owner, so there is less incentive for the shipping companies to reduce speed. This situation is not considered in this paper (Psaraftis and Kontovas, 2013).

where the variable v is the operating speed, the variable f_{ij} is the route-specific fuel price, the parameter $\lambda > 0$ is the fuel efficiency parameter, the parameter $\alpha > 1$ is the coefficient for fuel consumption, the variable d_{ij} is the distance, the variable q_{ij} is cargo volume, the variable C is the ship size, and the variable FC_{ij} is the fixed cost.[18]

The variable costs have two parts: the unit cost of shipping one unit of cargo for a day and the total transport work. The former consists of fuel costs per day per ton.[19] The latter includes the number of ships q_{ij}/C and delivery days.[20]

The fuel consumption (λv^{α}) is a non-linear function of speed. The value of α can be three for bulk carriers or tankers and four or five for containers (Psaraftis and Kontovas, 2013).

Because the distance on each route d_{ij} is fixed, and $D_{ij} = d_{ij}/v$, the cost function can be rewritten in terms of the delivery day D_{ij}. The marginal cost function is

$$MC_{ij} \equiv \frac{\partial TC_{ij}}{\partial q_{ij}} = f_{ij} D_{ij}^{-\beta} d_{ij}^{\alpha} z^{-1} \tag{2.8}$$

where $\beta \equiv \alpha - 1 > 0$ and $z \equiv \lambda C$. The derivative of marginal cost with respect to delivery day is

$$\frac{\partial MC_{ij}(z)}{\partial D_{ij}} = -\beta f_{ij} D_{ij}^{-\beta-1} d_{ij}^{\alpha} z^{-1}.$$

Slow steaming (lower v) on certain routes is equivalent to a longer delivery time (higher D_{ij}). Suppose that speed is not adjusted to save fuel costs; a one percent increase in fuel cost increases the marginal cost by one percent.

[18]See Stopford (2009), Psaraftis and Kontovas (2013), and Kosmas and Acciaro (2017) for more discussion on modeling shipping costs.

[19]There are also daily operating costs such as labor cost, ship maintenance cost, etc., which are often considered fixed for a ship of certain types and sizes. For ease of calculation, these costs are omitted in this paper.

[20]Delivery days include not only operating days d_{ij}/v but also the time at ports and the time on land. Because including these times does not bring any new insight into the analysis, and they can be controlled using a country-year dummy in the estimation, the model does not include these elements.

2.2.3 Shipping Market

From equation (2.2) and the assumption about trade cost, the demand for shipping on each route ij can be rewritten as

$$q_{ij} = Y_j P_j^{\sigma-1} e^{-\tau\sigma \ln D_{ij}} \left(p_j t_{ij} + FR_{ij} \right)^{-\sigma} \tag{2.9}$$

where $P_j^{1-\sigma} = \sum_j e^{\tau\sigma \ln D_{ij}} p_{ij}^{1-\sigma}$ and $p_{ij} = T_{ij} p_i$.

Hence,

$$\frac{\partial lnq_{ij}}{\partial FR_{ij}} = \frac{-\sigma}{t_{ij} p_j + FR_{ij}} \quad \text{and} \quad \frac{\partial lnq_{ij}}{\partial D_{ij}} = -\tau\sigma D_{ij}^{-1}. \tag{2.10}$$

Assuming that there is only one monopoly ship, which has the total shipping cost described in section 2.2.2; the profit of this ship is

$$\Pi_{ij} = q_{ij} FR_{ij} - TC_{ij}. \tag{2.11}$$

Solving the first order condition with respect to freight rate FR_{ij} yields

$$FR_{ij} - MC_{ij} = \frac{-1}{\frac{\partial \ln q_{ij}}{\partial FR_{ij}}} = \frac{p_j t_{ij} + FR_{ij}}{\sigma}, \tag{2.12}$$

so that

$$FR_{ij} = \frac{p_j t_{ij} + \sigma MC_{ij}}{\sigma - 1}. \tag{2.13}$$

Solving the first order condition with respect to day D_{ij} yields

$$FR_{ij} - MC_{ij} = \frac{\frac{\partial MC_{ij}}{\partial D_{ij}}}{\frac{\partial \ln q_{ij}}{\partial D_{ij}}} = \frac{\beta f_{ij} D_{ij}^{-\beta} d_{ij}^{\alpha}}{\tau\sigma z}. \tag{2.14}$$

Substituiting the expression of FR_{ij} in equation (2.13) and the expression of MC_{ij} in equation (2.8) into equation (2.14), the optimal delivery day is

$$D_{ij}^{*} = \left(A \frac{f_{ij} d_{ij}^{\alpha}}{p_j t_{ij} z} \right)^{1/\beta} \tag{2.15}$$

where $A \equiv \frac{\beta(\sigma-1)-\tau\sigma^2}{\tau\sigma(2\sigma-1)} > 0$ when $\frac{\beta}{\tau\sigma} > \frac{\sigma}{\sigma-1}$.

From equation (2.15), a one percent increase in fuel prices will delay the delivery day, but the magnitude depends on the fuel consumption coefficient $\beta \equiv \alpha - 1$.

Substituting the optimal day expression into equation (2.8), the marginal cost at optimal delivery day is

$$MC_{ij}(D_{ij}^{*}) = \frac{p_j t_{ij}}{A}. \tag{2.16}$$

Because of the specific functional forms of time valuation ($D_{ij}^{-\tau}$), fuel price f_{ij} is canceled out in the marginal cost function at optimal delivery day D_{ij}^{*}. The fuel price elasticity of marginal cost is zero (less than one in the case of no slow steaming). Slow steaming mitigates the negative impact of a fuel price increase.[21]

This could also be considered through a different lens. It has been shown in the work of Von Below and Vezina (2016) and Mirza and Zitouna (2009) that the freight costs does not increase to the magnitude of an increase in fuel prices. Hummels (2007) points out that the quality improvement in ocean transport services may not reflect well in price indices. If this applies during the high oil price period in the 2000s, it could mean the shipping company might as well adjust the quality of service instead of freight rates. A casual look at the freight rate trend during this period reveals that the freight rate did not increase much after 2010, even though fuel prices increased until 2011. Notteboom and Vernimmen (2009) also points out that bunker surcharge does not fully cover extra fuel costs due to a surge in fuel prices.

The markup is

$$\frac{FR_{ij} - MC_{ij}}{MC_{ij}} = \frac{\beta}{\tau\sigma}. \tag{2.17}$$

Ships that require more fuel consumption to maintain a certain speed charge

[21] In section 2.7.2, the different functional form shows that the fuel price elasticity of marginal cost is not zero.

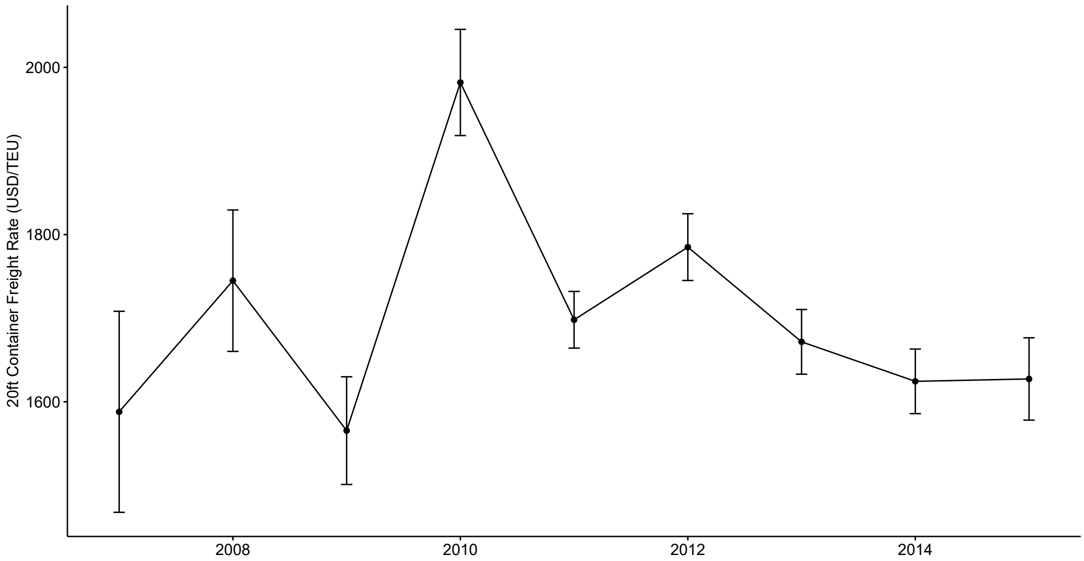


Figure 2.3: Yearly Average Freight Rate

Note: Average Freight Rate is a simple average of 20ft container freight rate between major ports. Data on fuel prices are from Drewry Container Freight Insight. This freight includes the Bunker Adjustment Factor (BAF), a surcharge that ship owners charge cargo owners for extra high fuel prices.

a higher rate. Consumers who are more sensitive to delivery days face higher markups.

Also, notice that the elasticity of fuel consumption with respect to fuel price can be calculated as

$$\frac{\partial \ln FuelConsumption(D_{ij}^*)}{\partial \ln f_{ij}} = \frac{-1}{\beta} = -\frac{\partial \ln D_{ij}^*}{\partial \ln f_{ij}}. \tag{2.18}$$

Substituting the expression of $MC_{ij}(D_{ij}^*)$ in equation (2.16) into equation (2.13) to get $FR_{ij}(D_{ij}^*)$ and the expression of trade costs

$$T_{ij}(D_{ij}^*) = \frac{2\sigma t_{ij}}{A(\sigma - 1)}. \tag{2.19}$$

In this setting, the trade costs do not depend on fuel cost or factory price p_i.

2.3 Estimation Strategy

2.3.1 Slow Steaming Evidence

To find the effect of fuel cost on delivery days, equation (2.15) can be estimated in log form using ship movement data. The interaction term between fuel prices and ship size is included in the regression to account for the heterogeneous effect of fuel prices on speeds for ships of different sizes. The variables γ_y and γ_q are year and quarter fixed effects. The ship fixed effects γ_s control for a ship's time-invariant characteristics, including fuel efficiency and ship size. The shipping operator fixed effect γ_o control for the shipping company's time-invariant characteristics. The fixed effects of importer γ_i, exporter γ_j, and country pair fixed effects γ_{ij} control for time-invariant market conditions. The main regression equation is

$$\begin{aligned}\ln D_{sijt} = \kappa_1 \ln f_{ijt} + \kappa_2 \ln Size_s \times \ln f_{ijt} + \kappa_3 \ln dist_{ijt} + \\ \kappa_4 \ln age_{st} + \gamma_s + \gamma_o + \gamma_{it} + \gamma_{jt} + \gamma_y + \gamma_q + \gamma_{ij} + \epsilon_{sijt}.\end{aligned} \tag{2.20}$$

The term $\kappa_1 + \kappa_2 \ln Size_s$ is the fuel price elasticity of delivery days. A positive κ_1 indicates that higher fuel prices induce slow steaming and longer delivery days. The magnitude of the effects depends on the size of the ship. A negative κ_2 indicates that bigger ships experience less slow steaming than smaller ships.

The elasticity of days with respect to fuel price is

$$\frac{\partial \ln D_{sij}}{\partial \ln f_{ijt}} = \kappa_1 + \kappa_2 \times \ln Size_s. \tag{2.21}$$

2.3.2 The Value of Timeliness

The value of timelines is expressed by the parameters τ in the consumer preference. From the gravity equation (2.5), one can obtain the estimates for $-\tau\sigma$

and $1-\sigma$ to calculate $-\tau$ from this estimate.[22]

Because the objective of this study is to estimate the value of timeliness, dividing both sides of the trade equation by trade cost $T_{ij}^{1-\sigma}$ gives the trade cost-adjusted trade equation

$$\frac{X_{ij}}{T_{ij}^{1-\sigma}} = \Phi_i \Phi_j e^{-\tau\sigma \ln D_{ij}} \tag{2.22}$$

where the left-hand side is freight cost-adjusted trade.

Rewriting equation (2.22) in regression form gives

$$\ln \frac{X_{ijt}}{T_{ijt}^{1-\sigma}} = \iota \ln D_{ijt} + \mathbf{W}_{ijt}\pi + \gamma_{it} + \gamma_{jt} + \gamma_{rij} + \varepsilon_{ijt}. \tag{2.23}$$

The coefficient ι is the estimate of $-\tau\sigma$. It captures the effects of delivery days purely from consumers' preference for timely delivery and not from the change in the freight costs. This set-up facilitates the usage of fuels cost in constructing the instrument for delivery days (details are explained in section 2.3.4). Coefficients γ_{it} and γ_{jt} are exporter-year and importer-year fixed effects. $\gamma_{r(ij)}$ are fixed effects for different sea routes. $\mathbf{W}_{ijt}$ contains gravity controls such as Common Language, Common Colony, Common Legal Systems, and Free Trade Agreement.

2.3.3 Delivery Day

The simplest way to aggregate delivery days from ship level to country pair level is to take a simple average. However, this approach treats ships of different sizes equally, even though they may react to fuel prices differently. A size-weighted average delivery day on each route is preferred.[23] The aggregate delivery day at

[22] This is the approach used in Hummels and Schaur (2013). However, what they measure on the left-hand side is not trade flow but the probability of switching from ocean shipping to air shipping.

[23] The motivation behind this is explained in Appendix 2.D, considering a monopolistic shipping market.

country pair level is

$$\ln D_{ijt} = \sum_{s} \theta_{sijt} \ln D_{sijt} \tag{2.24}$$

where

$$\theta_{sijt} \equiv \frac{ShipSize_s}{\sum_{s \in N_{ijt}^{ship}} ShipSize_s}$$

and the variable $ShipSize_s$ is the size of ship s measured in TEU, and the set N_{ijt}^{ship} is a set of ships operating on route ij at time t.

2.3.4 Endogeneity Issue

The challenge in estimating equation (2.23) is the endogeneity of delivery days. The left-hand side of this equation is freight cost-adjusted trade. Having the freight cost on the left-hand side allows us to separate the effects of delivery days on trade through freight cost and the direct effects of delivery days on trade demand.[24] The left-hand side contains the effects of fuel prices on trade through freight costs, whereas the right-hand side is left with the pure effects of fuel prices on trade through delivery days. However, it is still plausible that cost-adjusted trade is correlated with an unobservable "quality" (other than observable delivery days).[25]

To illustrate the problem, let us denote cost-adjusted trade as x and the unobservable "quality" as Ξ. The cost-adjusted trade can be expressed in matrix form as

$$x = D\iota + \Xi a + \varepsilon$$

where it is assumed that $a > 0$ so that higher quality increases cost-adjusted trade.

[24] Furthermore, it is worth noticing that this set-up depends on the multiplicative forms of trade costs, which include freight costs. This is commonly used in the trade literature using the gravity equation.

[25] Similar to delivery days, quality could affect trade directly through trade demand and indirectly through trade costs. Higher quality will likely increase trade directly through demand but also can decrease trade indirectly through trade costs. If the demand effects are larger than the trade costs effects, cost-adjusted trade positively correlates with quality.

The OLS estimator when Ξ is omitted is

$$\iota^{OLS} = (D'D)^{-1}D'x = \iota^{True} + (D'D)^{-1}D\Xi a + (D'D)^{-1}D'\varepsilon.$$

The second term contains the slopes in the least square regression of quality Ξ on delivery day D. If this slope is negative (higher quality means fewer delivery days), the OLS estimator is greater than the true value of ι. This calls for an instrumental variable (IV) to correct the bias.

2.3.5 Instrumental Variables

To address the possible endogeneity problem, candidate IV is the estimated delivery days from equation (2.20). However, the aggregation should not contain the time-variant factor from the change in demand for shipping. So the IV can be constructed as

$$\ln D_{ijt}^{IV} = \sum_{s \in N_{ijt}^{Ship}} \theta_{sij2010} \frac{\partial \ln D_{sij}}{\partial \ln f_{ijt}} \ln f_{ijt}. \tag{2.25}$$

By fixing the share of ship size in 2010, the time-variant portion of delivery days only contains information about the change in fuel prices. Because big ships react to fuel prices less than small ships, the trade route with a higher share of big ships in 2010 is likely to have less change in delivery days than those with smaller ships. Since the shares do not contain information about the difference in the composition of vessels after 2010, it does not include information about a possible increase in ship size due to an increase in market size after 2010.

Using the coefficients estimated from equation (2.20), the IV is

$$\widehat{\ln D_{ijt}^{IV}} = \sum_{s \in N_{ijt}^{Ship}} \theta_{sij2010}(\hat{\kappa_1} + \hat{\kappa_2} \times \ln Size_s) \ln f_{ijt}. \tag{2.26}$$

This IV captures the changes in delivery days in response to fuel price spikes.

The variations come from the change in fuel prices through time and across country pairs, as well as the cross-sectional difference in the composition of ship sizes. The way IV was constructed is similar to Feyrer (2019) and Sequeira, Nunn, and Qian (2020). The general idea is to run a stage “zero” regression to express the change in the potentially endogenous aggregated variable in terms of exogenous shocks and the heterogeneous effects of these shocks on the aggregated variable.[26] The zero-stage regression also backs out of all market conditions by including exporter-year and importer-year fixed effects.

The IV is likely to be exogenous (i.e. not correlate with the error terms) because their components are likely to be exogenous. First, the shares are predeterminant because they are from the beginning of the period. Second, the fuel prices themselves are not considered in the container trade and are not subject to simultaneous equations. Finally, the only channel through which fuel prices affect cost-adjusted trade is from delivery days, hence the IV does not suffer from omitted variable problems.

2.3.6 Relative Magnitude of Timeliness

There are two ways to examine the magnitude of timeliness. One is calculating the tariff equivalent amount of late delivery day as in Hummels and Schaur (2013) and as explained in Anderson and Van Wincoop (2004) and Head and Mayer (2014). The tariff equivalence of delayed delivery can be calculated as

$$Tequiv = exp\left(\frac{\hat{\iota}}{-\sigma}\right) - 1. \tag{2.27}$$

Another way is to compare the total effects of fuel prices. In the context of this paper, the effect of fuels on trade is shown in equation (2.6), which

[26]For example, in Feyrer (2019), the exogenous shocks are the change in the transportation distance among countries thanks to the adoption of air transport. This shock affects countries' trade volume differently because country pairs have different compositions between air and other modes of transportation.

is rewritten here for ease of reading. The first term on the right-hand side represents the change in the freight cost and the second term represents the effects of timeliness. Due to the specific assumptions, the first term is zero.[27] The second term depends on fuel prices and distances, as in equation (2.21).

$$\frac{\partial \ln X_{ij}}{\partial \ln f_{ij}} = \overbrace{\underbrace{(1-\sigma)s_{ij}^{F}\frac{\partial \ln F_{ij}}{\partial \ln f_{ij}}}_{FreightCost} \underbrace{-\tau\sigma\frac{\partial \ln D_{ij}}{\partial \ln f_{ij}}}_{Timeliness}}^{GrossEffects}.$$

To estimate the "gross" impact of fuel prices on trade (through both slow steaming and freight cost), the following regression is used

$$X_{ijt} = \exp(\psi_1 \ln f_{ijt} + \psi_2 \ln d_{ijt} + \gamma_t + \gamma_i + \gamma_j + \gamma_{ij} + u_{ijt}). \tag{2.28}$$

Coefficient γ_t are year-fixed effects. The fixed effects of importer γ_i, exporter γ_j, and country pair γ_{ij} control for time-invariant market conditions. Pseudo Poisson Maximum Likelihood (PPML) is used to run this regression.

The effect from "timeliness" can be calculated using

$$\begin{aligned} \left.\frac{\partial \ln X_{ij}}{\partial \ln f_{ij}}\right|_{Timeliness} &= -\tau\sigma\frac{\partial \ln D_{ij}}{\partial \ln f_{ij}} = \hat{\iota}\sum_{sij}\theta_{sij}\frac{\partial D_{sij}}{\partial \ln f_{ij}} \\ &= \underbrace{\hat{\iota}\hat{\kappa}_1}_{<0} + \underbrace{\hat{\iota}\hat{\kappa}_2\sum\theta_{sij}\ln Size_{sij}}_{>0}. \end{aligned} \tag{2.29}$$

While higher fuel prices affect trade negatively, the magnitude of this effect is smaller on a route with larger than average ship size.

[27] Section 2.7.2 check shows a case when the second term does not equal zero. The typical value for this channel is between -0.07 and -0.02 (with oil price elasticity of 0.2 and σ is between 3 and 10, and the freight costs share is about 0.38), as discussed in section 2.2.1.

2.4 Data Description

2.4.1 Ship Voyage

Ports call data from IHS Sea-Movement is processed to get the information on ships arriving in every port from 2010 to 2015. The number of travel days for a direct trip is calculated as the duration between two arrival dates minus the waiting time in the previous ports. An indirect trip between two ports may contain several port calls and is calculated by finding the shortest distance algorithm.[28] The number of travel days for indirect trips is the sum of all travel days on all legs. Information on ship registries (IMO numbers) is used to extract information about ship size and ship age. Trips at the port level are then aggregated to the country level.[29]

Table 2.1 shows the summary for ships in the final data set. There are about $289,000$ direct trips recorded between 2010 and 2015 (if one ship travels between two ports more than once each quarter, we take the average days for that quarter). The average speed is 34.6 km/h. The average number of days is 6.1. The average age is 8.4 years, and the average size is $4,437$ TEU.

2.4.2 Fuel Prices

Data for fuel prices are from Clarkson Shipping Intelligence. This data is reported at main major bunkering ports.[30] Fuel prices at each port are extrapolated using the fuel price of main bunkering ports and information about the ports' infrastructures. Fuel prices on a voyage are assumed to be the average prices of

[28] Maritime routes algorithm (https://github.com/eurostat/searoute) is used to calculate the maritime distance between two ports (using ports' longitude and latitude). The package cppRouting(Larmet, 2019) in R is used to find all possible connected ports, and only chronologically correct combinations are kept. The information on indirect trips is used to calculate the average delivery days between two countries when there are no direct trips.

[29] Some regions (including the US, Canada, Russia, and Australia) are divided into two regions facing two oceans. Details on data processing are discussed in Appendix 2.B.1.

[30] Ports that sell fuel.

Table 2.1: Summary Statistic: Ship Voyage and Container Trade

Statistic	N	Mean	St. Dev.	Min	Max
Ship Data					
Speed(km/h)	289,347	34.6	23.9	9.1	121.3
Day	289,347	6.1	5.9	1.0	95.0
Size(TEU)	289,347	4,436.9	3,416.9	88	19,870
Age	289,347	8.4	5.2	1	55
Distance(km)	289,347	4,675.0	4,642.3	218.2	24,784.7
Fuel Price(USD/ton)	289,347	544.1	135.2	256.3	706.4
Trade Data					
Cont.Trade(mil.USD)	20,502	1,038.2	6,718.9	0.0	295,411.2
CIF/FOB(w.average)	20,502	0.1	0.04	0.0	0.9
CIF/FOB(s.average)	20,502	0.1	0.02	0.0	0.3
Fuel price(USD/ton)	20,502	550.5	129.4	269.7	702.9
Distance(km)	20,502	12,085.6	9,031.2	226.2	96,451.0
ln(Day)	20,502	1.2	1.2	0.0	5.5
ln(Size)	20,502	3.9	3.2	0.02	9.8

Source: Ship data is from the IHS Sea web. Trade data is from IHS Markit. The fuel Price is from Clarkson. Sea distance is calculated using the SeaRoute algorithm created by Eurostat. CIF-FOB margin is from OECD-ITIC. Other gravity controls are from CEPII.

all ports along the voyage.

Figure 2.4 shows the extrapolated bunker prices for all ports.[31] The cross-sectional variation is mainly from the difference in maximum drafts and the geographical position of the ports in relation to the Middle East area. The serial variation reflects the general trend of fuel prices. Fuel prices increased drastically between 2011 and 2014, largely due to a constraint in supply, and fell in 2015 onward because of the decrease in Chinese demand and the expansion in the US's production (OECD, 2015).

Because the fuel prices are extrapolated for all ports using the fuel prices of main ports and ports' capacity (maximum draft and overall length of ships that can use the port), the variations in the fuel price across regions come from the variation of the ports in those regions.

[31] See Appendix 2.C for details on the extrapolation of bunker prices.

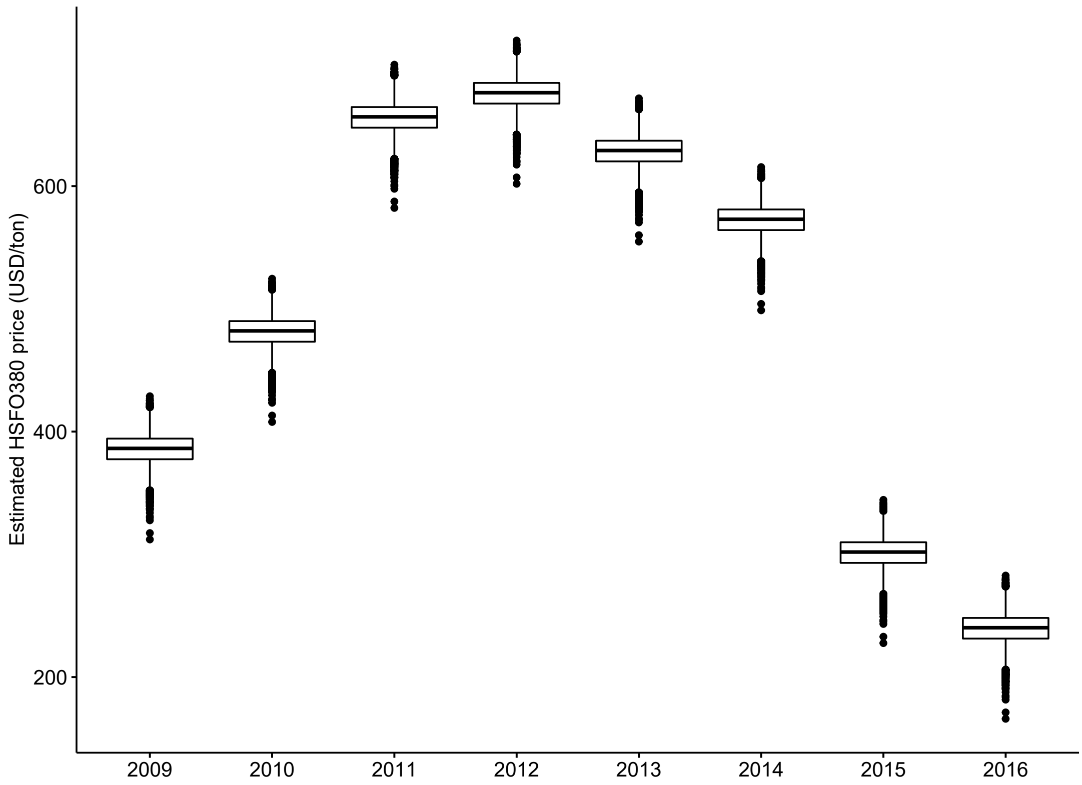


Figure 2.4: Extrapolated Fuel Prices

2.4.3 Country Level Data

The average delivery day between countries ij is the size-weighted average (in natural logarithm) delivery day of all direct voyages between ij for all ports.[32] If there is no direct service between ij, the above-calculated indirect voyages are used instead. Distances and fuel prices are calculated in the same manner.

Annual container trade data is from IHS Markit. To match the theoretical model, trade value is calculated by multiplying the quantity with a unit price for each commodity-pair-year observation. This unit price is calculated by dividing the total seaborne trade value by the total seaborne trade quantity for each commodity and pair of countries. All commodities are then summed up to get the total trade value for each country pair.

Transportation cost T_{ij} is approximated by CIF-FOB margins from OECD-

[32]The average of natural logarithm is used in the main regression. The average without natural logarithm is presented in the robustness check section in Section 2.7.2.

ITIC database (Miao and Fortainier, 2017). The original data are at HS 4-digit and are aggregated to the country pair level using trade value-weighted average.[33] However, because there are too many zeros in trade data, the main regression used the simple average CIF-FOB margins.

Other gravity controls, such as regional trade agreements, common borders, and common language, are from Head and Mayer (2014). Further details for all data is described in Appendix 2.A.

2.5 Findings

2.5.1 Slow Steaming Evidence

The estimation results for equation (2.20) are shown in Table 2.2. The coefficients for fuel prices are positively significant, while the coefficient for distance is significantly positive. This confirms the hypothesis that high fuel prices induce shipping companies to increase delivery time by reducing ship speeds. The magnitude of speed reduction depends on ship size. The bigger the ship, the less speed reduction. Younger ships tend to have shorter delivery days.

The elasticity of delivery days with respect to fuel prices as in equation (2.21) is

$$\frac{\partial \ln D_{sij}}{\partial \ln f_{ij}} = 0.6997 - 0.0232 \ln Size_s. \tag{2.30}$$

The average elasticity at the sample means 4,437 TEU is 0.5. On average, a 10 percent increase in fuel prices increases delivery days by five percent. This translates to a delay of one day on the trade route between East Asia and North America (with an average delivery time of 20 days).

Notice that $\hat{\kappa}_1$ corresponds to $1/(\alpha - 1)$, so $\hat{\alpha} = 2.4$, which is very close to the assumption often used in maritime economic literature (Stopford, 2009;

[33] Data on trade value is from The Observatory of Economic Complexity.

Table 2.2: Estimation Results: Delivery Days and Fuel Prices

Dependent Variable:	ln(Days)
Model:	(OLS)
ln(Fuel Price)	0.6997*** (0.1440)
ln(Fuel Price)*ln(Size)	-0.0232*** (0.0049)
ln(Distance)	0.5596*** (0.0076)
ln(Age)	-0.0228*** (0.0061)
Quarter = 2	-0.0623*** (0.0055)
Quarter = 3	-0.0389*** (0.0050)
Quarter = 4	-0.0073** (0.0028)
Observations	289,347
R^2	0.76371
Within R^2	0.09056

Clustered (Pair-Company) standard-errors in parentheses
Fixed effects: exporter-year, importer-year, country pair, company, ship
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Psaraftis and Kontovas, 2013; Kosmas and Acciaro, 2017). As in equation (2.18), the elasticity of fuel consumption with respect to fuel price is minus the elasticity of delivery days with respect to fuel price. An average ship (4,437 TEU) can save five percent fuel consumption by speed reduction when fuel prices increase by 10 percent.

2.5.2 The Value of Timeliness

The results of equation (2.23) are shown in Table 2.3. To calculate the left-hand side of this equation, it was assumed that $\sigma = 8$, following Anderson and Van Wincoop (2004).[34] The variable $\ln(Day)$ is the weighted average of delivery day as in equation (2.24) and the IV is constructed using equation (2.25).

Column one reports the OLS regression, column two reports the second stage of the IV regression, and column three reports the regression on the IV variable. While both OLS and IV specification shows significant results for delivery days,

[34] Results with different values of $\sigma \in \{3, 10\}$ are shown in Section 2.7.1.

Table 2.3: Estimation Results: Adjusted Trade and Delivery Days ($\sigma = 8$)

Dependent Variable:		ln(Adj.Trade)($\sigma = 8$)	
Model:	(OLS)	(2SLS)	(Reduced Form)
ln(Day)	-0.1571***	-0.0814**	
	(0.0164)	(0.0325)	
Day_IV			-0.0568**
			(0.0228)
Language	0.4988***	0.5063***	0.5122***
	(0.0830)	(0.0837)	(0.0841)
C.Colony	0.3677***	0.3835***	0.4003***
	(0.1383)	(0.1390)	(0.1402)
Legal	0.2236***	0.2202***	0.2179***
	(0.0633)	(0.0637)	(0.0641)
FTA	0.6344***	0.6685***	0.7020***
	(0.0769)	(0.0793)	(0.0775)
Observations	20,196	20,196	20,196
R^2	0.81028	0.80974	0.80801
Within R^2	0.06842	0.06574	0.05728

Clustered (Pair) standard-errors in parentheses
Fixed effects: Exporter-Year, Importer-Year, Sea Route
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

the magnitude of the coefficients in the IV regression is only half of the OLS results and very close to the reduced form results. It suggests that if the endogeneity is not addressed, OLS results may overestimate the timeliness value. It is because the unobservable "quality" positively correlates with cost-adjusted trade but negatively correlates with delivery days (as discussed in Section 2.3.4).

Table 2.4 shows the results of the first and second stages of IV regression side by side. The result of the F-test shows that the constructed IV is not a weak instrument. The Wu-Hausman statistics cannot reject the endogeneity of ln(Day). This confirms the validity of the endogeneity issue.

The value of timeliness can be calculated as

$$\hat{\tau} = -\hat{\iota}/\sigma = 0.0814/8 = 0.0102$$

Table 2.4: Estimation Results: Adjusted Trade and Delivery Days (IV)

Dependent Variables:	ln(Days)	ln(Adj.Trade)
IV stages	First	Second
Day_IV	0.6973***	
	(0.0160)	
ln(Day)		-0.0814**
		(0.0325)
Language	-0.0727**	0.5063***
	(0.0367)	(0.0837)
C.Colony	-0.2057***	0.3835***
	(0.0570)	(0.1390)
Legal	0.0289	0.2202***
	(0.0280)	(0.0637)
FTA	-0.4114***	0.6685***
	(0.0324)	(0.0793)
Observations	20,196	20,196
R^2	0.40109	0.80974
Within R^2	0.14438	0.06574
F-test (IV only)	2,624.1	7.5545
Wu-Hausman, p-value		0.00608

Clustered (Pair) standard-errors in parentheses
Fixed effects: Exporter-Year, Importer-Year, Sea Route
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

which is slightly smaller than Hummels and Schaur (2013), who estimated that $\tau = 0.069/3.301 = 0.0209$. This is intuitive because their estimation represents the premium between air and ocean shipping. Whereas this paper's results reflect the value of timeliness within container shipping.[35]

Using equation (2.31), the tariff equivalent amount of a one percent decrease in delivery day is

$$Tequiv = exp\left(\frac{\hat{\iota}}{-\sigma}\right) - 1 = 1.023\%. \tag{2.31}$$

For the route between East Asia and North America (average 20 days), an increase in one day is equivalent to a five percent additional tariff.

[35] To better compare with theirs, a robustness check with the same functional form of delivery day and value of σ is performed in Section 2.7.2.

2.5.3 Relative Magnitude of Timeliness

The estimation results of equation (2.28) are reported in the first column of Table 2.5. A one percent increase in fuel price results in a decrease of about one percent in trade. The results for adjusted trade in the second column also show similar results. This implies that most adjustments in the shipping market do not fully materialize in the shipping cost. However, there are two caveats. First, the significant level is quite low. Second, the transportation cost was approximated by CIF-FOB margins for all modes of transportation.

Table 2.5: Estimation Results: Trade and Fuel Prices

Dependent Variables: Model:	Trade (PPML)	Adjusted Trade($\sigma = 8$) (PPML)
ln(Fuel Price)	-1.091*	-1.026*
	(0.6390)	(0.6106)
ln(Distance)	-0.0476**	-0.0487**
	(0.0223)	(0.0228)
Observations	20,502	20,502
Squared Correlation	0.99374	0.99468
Pseudo R^2	0.99579	0.99578
BIC	4.38×10^{11}	7×10^{11}

Clustered (Pair) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Fixed effects: exporter, importer, year, country pair

Assuming a Δ percent change in fuel prices, the change in trade is

$$\%Trade\Big|_{Gross} = \hat{\psi}\Delta = -1.091\Delta. \tag{2.32}$$

The change from slow steaming adjustment via timeliness is

$$\begin{aligned} \%Trade\Big|_{Timeliness} &= \hat{\iota}(\hat{\kappa_1} + \hat{\kappa_2} \sum_{s \in N_{ijt}^{Ship}} \theta_{sijt} \ln Size_s)\Delta \\ &= -0.057\Delta + 0.002 \sum_{s \in N_{ijt}^{Ship}} \theta_{sijt} \ln Size_s \Delta. \end{aligned} \tag{2.33}$$

The second term is the weighted average of ship size in the natural logarithm. Because the maximum ship size is less than $20,000$TEU, all routes experience negative impacts. However, routes with large ship sizes have smaller impacts. On average, a one percent increase in fuel prices causes a decrease of 0.05 percent through timeliness.[36]

Table 2.6 shows the calculation for the fuel price increase between 2010 and 2011. During this time, fuel prices increased by 36 percent on average. It caused a decrease in delivery time of 22 percent on average. Trade decreased by 39.6 percent, of which about 4.5 percent is from timeliness.

Table 2.6: Effects of Fuel Price Increases in 2011

Statistic	N	Mean	St. Dev.	Min	Max
Day	3,417	15.6	12.9	1.0	111.5
κ_1	3,417	0.7	0.0	0.7	0.7
κ_2	3,417	−0.1	0.1	−0.2	−0.001
Slow Steaming Coef.	3,417	0.6	0.1	0.5	0.7
%Fuel Change	3,417	36.3	1.8	30.2	44.4
%Day Change	3,417	22.0	2.8	15.0	30.2
%Trade Change (Timeliness)	3,417	−1.8	0.2	−2.5	−1.2
%Trade Change (Total)	3,417	−39.6	2.0	−48.5	−32.9
%Share Timeliness	3,417	4.5	0.5	3.5	5.2

2.6 Slow Steaming on Different Trade Routes

Table 2.8 shows the results of the equation (2.21) for different sub-samples: voyages within regions, voyages between eastern and western hemisphere, voyages between northern and southern hemispheres.[37] The coefficients of fuel prices for regional routes are insignificant, while those of inter-regional routes are signifi-

[36] Brancaccio, Kalouptsidi, and Papageorgiou (2020) shows that an increase of one percent in fuel price reduces trade by about 0.35 percent. Their estimate is ten times higher than this study. It is possible because they study the bulk carrier market, where a larger portion of the shipping cost is fuel cost. This sector also has fewer long-term service contracts, so ships respond more aggressively to fuel prices by repositioning, further increasing shipping costs.

[37] There are 36 routes: 8 regional routes, 13 East-West routes, and 15 North-South routes.

Table 2.7: Ship Size (TEU) by Route

	Group	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1	Regional	88	1,296	2,664	3,900	5,552	19,870
2	East-West	95	2,702	4,616	5,206	6,763	19,870
3	North-South	95	2,506	4,253	5,024	6,690	19,870

8 regional routes: North America, South America, Europe, Africa, Middle East, South Asia, East Asia, Oceania, 13 East-West routes: North America-Europe, North America-Middle East, North America-South Asia, North America-East Asia, South America-Africa, South America-Oceania, Europe-Middle East, Europe-South Asia, Europe-East Asia, Middle East-South Asia, Middle East-East Asia, South Aisa-East Asia
15 North-South routes: North America-South America, North America-Africa, North America-Oceania, South America-Europe, South America-Middle East, South America-South Asia, South America-East Asia, Europe-Africa, Europe-Oceania, Africa-Middle East, Africa-South Asia, Africa-East Asia, Middle East-Oceania, South Asia-Oceania, East Asia-Oceania

cant. The slow steaming effects in the North-South group are twice as high as in the East-West group.

There are a few reasons why the coefficients are different for different trade routes. For regional routes, it's easier to adjust schedules to respond to high fuel prices because distances are shorter than inter-regional routes. The reason for the difference between East-West and North-South, however, is not very straightforward. One possible reason is that ship sizes on the East-West route are slightly bigger than that of the North-South route (Table 2.7). And fuel consumption coefficient (α) is higher for larger ships, as shown in the main regression.

Table 2.8: Estimation Results: Slow Steaming by Sea Route

Dependent Variable:	ln(Day)			
Sea Route	Full sample	Regional	EastWest	NorthSouth
Model:	(1)	(2)	(3)	(4)
ln(Fuel Price)	0.6997***	0.1412	0.8163**	1.532***
	(0.2343)	(0.2480)	(0.3891)	(0.4937)
ln(Fuel Price) × log(Size)	-0.0232***	-0.0116	-0.0689***	-0.0063
	(0.0071)	(0.0089)	(0.0157)	(0.0143)
ln(Distance)	0.5596***	0.5584***	0.4717***	0.5822***
	(0.0249)	(0.0279)	(0.0306)	(0.0286)
ln(Age)	-0.0228***	-0.0002	-0.0588***	-0.0380***
	(0.0067)	(0.0086)	(0.0170)	(0.0126)
Observations	286,566	155,406	55,258	75,902
R^2	0.76397	0.63367	0.69361	0.76195
Within R^2	0.09088	0.13305	0.01945	0.03602

Clustered (Pair) standard-errors in parentheses
Fixed effects: exporter-year, importer-year, quarter, country pair, company, ship
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Using the coefficients for different trade routes to construct the IV to estimate equation (2.23), the fuel coefficient becomes -0.1256, slightly higher when all routes have the same coefficients.[38]

2.7 Robustness Check

2.7.1 Different Values of σ

Table 2.9 shows a summary of results for different values of σ. The estimates for $\hat{\iota}$ are almost identical. However, because the magnitude of $\hat{\iota}$ is very small, a change in each unit of σ causes a drastic change in the tariff equivalent value.

[38] See table 2.E.1 in Appendix 2.E for detailed estimation results.

Table 2.9: Results for Different Values of σ

σ	$\hat{\iota}$	$\hat{\tau}$	Tequiv
3	-0.08300	0.02767	2.8052
4	-0.08269	0.02067	2.0887
5	-0.08237	0.01647	1.6611
6	-0.08206	0.01368	1.3771
7	-0.08175	0.01168	1.1747
8	-0.08144	0.01018	1.0232
9	-0.08113	0.00901	0.9055
10	-0.08082	0.00808	0.8114

All estimation follows the same specifications in equation (2.22). *All coefficients are significant at less than* 5% *significant level. See Table 2.E.2 in Appendix 2.E for details on estimation results.*

2.7.2 Different Functional Forms of Day

Assuming the functional forms as in Hummels and Schaur (2013), the utility of a representative consumer is

$$U_j = \left(\sum_{i}^{N} e^{-\tau D_{ij}} q_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}. \tag{2.34}$$

The derivative of log demand with respect to delivery day is

$$\frac{\partial lnq_{ij}(z)}{\partial D_{ij}(z)} = -\tau\sigma. \tag{2.35}$$

The optimal delivery days for ship z is a solution of the following equation

$$\frac{1}{D_{ij}(z)^{\beta}} \left(\frac{\beta}{\tau\sigma D_{ij}(z)} - \frac{\sigma}{\sigma+1} \right) = \frac{(2\sigma-1)p_j t_{ij} z}{(\sigma-1) f_{ij} d_{ij}^{\alpha}}. \tag{2.36}$$

The left-hand side decreases in D_{ij} and the right-hand side is constant with respect to D_{ij}. Hence, there is a unique solution for delivery day, denoted as D_{ij}^* (higher f_{ij}, longer D_{ij}^*).

Substituting D_{ij}^* into equation (2.13) gives an expression of FR_{ij} that depends on f_{ij}.

The trade equation is

$$\ln \frac{X_{ijt}}{T_{ijt}^{1-\sigma}} = \iota D_{ijt} + \mathbf{W}_{ijt}\pi + \gamma_{it} + \gamma_{jt} + \gamma_{r(ij)} + \epsilon_{ijt}. \tag{2.37}$$

The IV is

$$\begin{aligned} \widehat{D_{ijt}^{IV}} &= \sum_{s \in N_{ijt}^{Ship}} \theta_{sij2010} \exp[(\hat{\kappa_1} + \hat{\kappa_2} \times \ln Size_s) \ln f_{ijt}] \\ &= \sum_{s \in N_{ijt}^{Ship}} \theta_{sij2010} f_{ijt}^{\kappa_1 + \kappa_2 \ln Size_s}. \end{aligned} \tag{2.38}$$

Estimation results of equation (2.37) are shown in Table 2.10. The coefficient $\hat{\iota} = -0.0087$, ten times lower than the log functional form (0.0814) and the Hummels and Schaur (2013) estimation (0.069). The elasticity of trade with respect to delivery day is $-0.0087Day$. On East Asia - North America trade route (20 days), the elasticity equals -0.175. The tariff equivalent amount of one day is 0.26 percent, less than the Hummels and Schaur (2013) estimation (2.1 percent), which resonates with the qualitative results in the main regression.

It is worth noticing that the estimates of the trade elasticities with respect to delivery day are very sensitive to the functional forms of delivery day in the utility function. In the main specification, the conventional log-log form produces a constant elasticity of trade with respect to days of about -0.083. In the case of log-linear form as in Hummels and Schaur (2013), the elasticity of trade with respect to days is $-0.0087Day$. To further examine this issue, equation (2.23) is estimated using a quadratic functional form.[39] The elasticity, in this case, is $0.365 - 0.011Day$. It seems that the log-linear form produces similar estimates to the case of constant elasticity when delivery days are 10 days, while the quadratic form does so at 25 days.

Comparing these estimates with the elasticity of adjusted trade with respect

[39] The IV is constructed as in the main specification. The results are shown in Table 2.11.

Table 2.10: Estimation Results: Adjusted Trade and Delivery Days ($\sigma = 3.301$)

Dependent Variable:	ln(Adj.Trade)($\sigma = 3.301$)		
Model:	(1)	(2)	(3)
Day	-0.0136***	-0.0087**	
	(0.0016)	(0.0040)	
Day_IV			-0.0065**
			(0.0030)
Language	0.5108***	0.5155***	0.5215***
	(0.0832)	(0.0837)	(0.0840)
C.Colony	0.3795***	0.3884***	0.4042***
	(0.1382)	(0.1385)	(0.1397)
Legal	0.2128***	0.2111***	0.2093***
	(0.0634)	(0.0636)	(0.0640)
FTA	0.6800***	0.6971***	0.7241***
	(0.0765)	(0.0786)	(0.0772)
Observations	20,196	20,196	20,196
R^2	0.81108	0.81086	0.80948
Within R^2	0.06701	0.06592	0.05912

Clustered (Pair) standard-errors in parentheses
Fixed effects: Exporter-Year, Importer-Year, Sea Route
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

to distances (of about -0.475 for when $\sigma = 8$),[40] these estimates are quite small. Utilizing the relationship between distance and delivery days ($Day = Distance/Speed$), the elasticity of days is the elasticity of distance minus the elasticity of speed. This suggests that the elasticities of speed can be almost as high as distance elasticities.

[40]Table 2.E.3 in Appendix 2.E provides estimation results for different value of σ.

Table 2.11: Adjusted Trade and Delivery Days ($\sigma = 8$)

Dependent Variables:	Day	Day Square	ln(Adj.Trade)
IV stages	First		Second
Model:	(1)	(2)	(3)
Day_IV	1.802	31.60	
	(1.235)	(48.07)	
Day_IV square	1.256***	53.05***	
	(0.3930)	(15.57)	
Day			0.3645**
			(0.1653)
Day square			-0.0107**
			(0.0047)
Language	-0.6731**	-25.71**	0.4806***
	(0.3055)	(12.69)	(0.1130)
C.Colony	-1.702***	-41.35*	0.5732**
	(0.5164)	(25.06)	(0.2386)
Legal	0.2594	0.7882	0.1265
	(0.2558)	(16.81)	(0.1632)
FTA	-3.010***	-83.66***	0.8991***
	(0.2689)	(10.99)	(0.1479)
Fit statistics			
Observations	20,196	20,196	20,196
R^2	0.36148	0.17979	-1.4113
Within R^2	0.12221	0.03475	-10.840
F-test (IV only)	1,130.9	304.61	22.374
Wu-Hausman, p-value			4.51×10^{-7}

Clustered (Pair) standard-errors in parentheses
Fixed effects: Exporter-Year, Importer-Year, Sea Route
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

2.8 Conclusion

This study used a semi-natural experiment to estimate the value of timeliness in international container trade. A spike in fuel prices during the first ten years of the 2000s induced shipping companies to practice slow steaming. While this helps shipping companies save fuel costs, it caused delivery delays, which in turn negatively affected trade.

The theoretical framework shows fuel price affects trade through speed adjustment of the shipping sector. When fuel prices increase, the ship can adjust delivery days and freight costs. The change in delivery days affects consumers as if there is a decrease in the quality of goods. Slow steaming also softens the effect of fuel prices on freight costs.

In the empirical exercise, the evidence of slow steaming is confirmed using data on ship movement. Estimates of the elasticity of delivery days with respect to fuel prices were shown to depend on ship sizes. The average elasticity at the sample means $4,437$ TEU is 0.5. On average, a 10 percent increase in fuel prices increases delivery days by five percent. This translates to a delay of one day on the trade route between East Asia and North America (with an average delivery time of 20 days). This will also save fuel consumption by five percent. Larger ships have higher fuel consumption coefficients, hence saving more fuel.

The value of timeliness is estimated by constructing an IV for delivery days. The IV considers the response of ships with respect to fuel prices and the composition of ship sizes for different trade routes. The elasticity of trade with respect to delivery days is estimated to be about -0.0814. A one percent delay in delivery is equivalent to an additional one percent tariff. On the East Asia - North America trade route (20 days), one delayed day adds five percent tariff equivalent. Between 2010 and 2011, fuel prices increased by 36 percent. This caused delays of 22 percent on average (for the East Asia - North America trade route, an equivalent of four days). This spike in fuel price decreased trade by

39.6 percent, of which about 4.5 percent is from timeliness.

There are some geographical differences in ships' responses to fuel price increases. The coefficients of fuel prices for regional routes are insignificant, while those of inter-regional routes are significant. It's easier for regional routes to adjust schedules in response to high fuel prices thanks to their shorter distances than inter-regional routes. Among inter-regional routes, the slow steaming effects in the North-South group are twice as high as in the East-West group. It is because ship sizes on the East-West route are slightly bigger than that of the North-South, and the fuel consumption coefficient is higher for larger ships.

Robustness checks for different values of elasticity of substitution (σ) show similar results for the elasticity of trade with respect to delivery day. Changing functional forms of delivery days from constant to variable elasticity results in the elasticity of trade with respect to delivery day to be $-0.0087Day$. On the East Asia - North America trade route (20 days), the elasticity equals -0.175. In general, the estimates of the value of timeliness are smaller than in the literature. This is because previous literature measures the premium between air and ocean transportation, whereas this study measures the variation within the container shipping sector.

For future analysis, data on freight costs can improve the accuracy of the estimation. In this study, container transportation cost is proxied by the CIF-FOB margins that include all modes of transportation. While an increase in fuel cost is likely to increase all modes of transportation costs, the magnitudes are different due to the different types of fuels and different shares of air or ocean transportation for different country pairs. The model can also be extended in several ways, including changes in ship allocation and shipment frequency.

Appendices

2.A Data Source

The data used in this chapter is a combination of port arrival data, container trade data, fuel prices, and CIF-FOB margin.

2.A.1 Port Arrival Data

The data contain information about port arrivals of container ships. It has the following information: ship identification number (IMO number), operating company, arrival and departure date and time, port name, country name, and draft.

This information is processed by IHS Sea-web movement (data provider) mainly from AIS (automatic identification system).[41] AIS data contain messages sent from a ship to a satellite to record their position. The initial use is to avoid a collision. From these messages, the data provider can extrapolate information about ship movements, including speed and positions. Because it was sent from the ship automatically, it is considered to be more accurate than data about the arrival and departure draft, which is manually input by ship operators.

There is a restriction on downloading the data for a day. To facilitate the download, data shown on the screen was saved as a snapshot and then converted back into a text file.

[41] https://maritime.ihs.com/EntitlementPortal/Home/Information/Seaweb_Movements

The same service also has information about the ship's characteristics: size, age, owner, engines, etc.

This provider also has data about port information. Some ports are added by checking the ports' home pages. It contains information about ports, including maximum depth, draft, and geographical locations.

2.A.2 Container Trade Data

The data has information about the volume of the container, dry bulk, liquid bulk trade, the value of seaborne trade, and the volume of seaborne trade for about 270 commodities.[42] These commodities are grouped by the data provider but can be grouped back to HS 2-digit level. The value of container trade is calculated by multiplying the volume by the unit price (calculated by dividing the value by the volume of seaborne trade).

The data is calculated by various sources, but mainly from custom data. Because only a few countries report trade data by mode of transport. The data provider extrapolated container trade data using custom data and some conversion factors. I have asked the company about the details of their calculation, but they only provide the above explanation.

2.A.3 Fuel Prices

Fuel prices are the quarterly average of spot fuel prices at the main fuel ports. The fuel type is HSFO 380 CST. Data is from Clarkson's Shipping Intelligence Network.[43] However, only fuel prices of main ports are reported. Missing data is extrapolated as in Section 2.C.

[42] https://www.spglobal.com/marketintelligence/en/mi/products/gta-forecasting-data-lake.html

[43] https://sin.clarksons.net

2.A.4 CIF-FOB Margin

This was estimated by OECD[44] using FOB and CIF trade data. The data has the ratio between CIF and FOB at the HS-6 digit level. It was aggregated for each country using trade value data.[45] Miao and Fortainier (2017) describes the detailed methodology and original data.

2.B Data Processing

2.B.1 Ship Data

The original data of arrival ports have more than 5.4 million records. Because data before 2009 is not credible, they are excluded from the dataset. Some records do not have information about ports in the port arrival data, which is filled with the most frequently visited ports of each region. Each country is a region except for the USA, Canada, Russia, and Australia. These four countries are divided into two regions because they face two oceans. These involve steps 1 to 7 in Table 2.B.1.

Direct voyage data is created from arrival data by two consecutive dates for each ship. Delivery days are the gap between two dates minus waiting time in the previous port. Voyages with the same port pair and negative delivery times are removed (step 8). Distances between ports in different regions are calculated using maritime routes (see figure 2.B.1). In contrast, distances between ports in the same region are lengths of a straight line between two points.

Indirect trips are added by using the shortest distance algorithm, considering chronologically correct order (steps 9, 10, 11). Voyages within one country are removed (step 12). Voyages with outlier speed are removed (steps 13, 14). Voyages between ports are aggregated to voyages between regions by taking the

[44] https://stats.oecd.org/Index.aspx?DataSetCode=CIF_FOB_ITIC
[45] https://oec.world

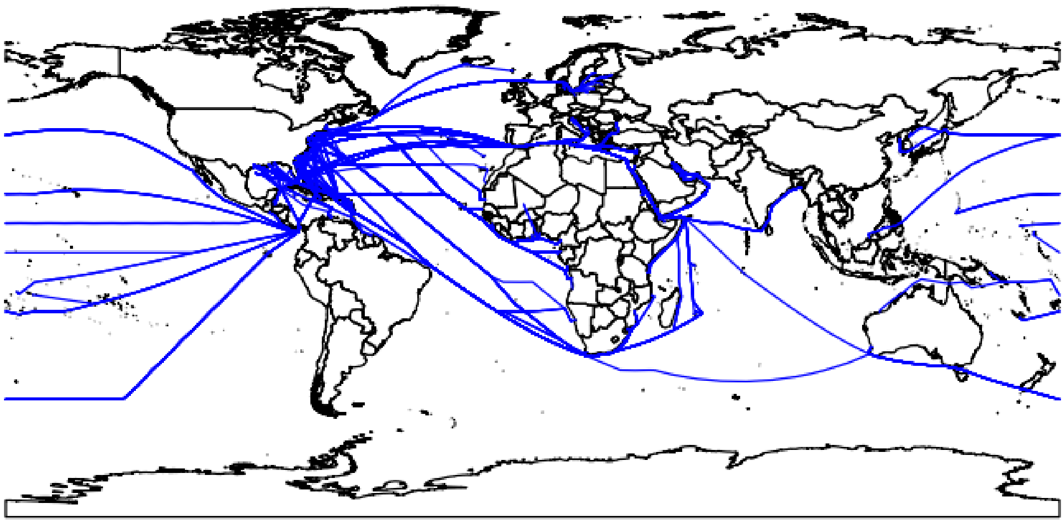


Figure 2.B.1: Maritime Route from the US's East Coast

Source: SeaRoute algorithm (https://github.com/eurostat/searoute) and package rworldmap (South, 2011)

average delivery days and speed for each ship in a region pair for each quarter (step 15). Region pairs with outlier speed are removed (step 16).

Table 2.B.1: Ship Data Processing

	Step	N.obs	N.direct	N.Oper	N.Ship	N.Ex	N.Port.Ex	N.Im	N.Port.Im
1	Raw	5,401,421		489	4,341			186	2,517
2	Keep between 2009 and 2016	4,493,711		486	4,324			186	2,509
3	Removed NA country call	4,469,886		486	4,324			185	2,401
4	Fill NA port with locf, nocb	727,800		473	4,279			157	927
5	Fill NA port most obs by Operator, Ship, Country call	524,369		468	4,265			151	751
6	Fill NA port most obs by Country call	163,049		380	3,606			118	115
7	Removed obs with one NA country call for whole period	4,469,799		486	4,324			185	2,401
8	Removed negative duration, same port pair	2,928,973		482	4,308	185	2,394	185	2,385
9	Indirect routes added	10,499,493	2,663,633	482	4,308	189	2,394	189	2,385
10	Non-choronical removed	5,067,502	874,991	482	4,308	189	2,394	189	2,385
11	Zero duration removed	4,323,050	690,672	482	4,307	188	2,364	189	2,361
12	Cabotage removed	3,325,565	517,770	471	4,274	188	2,315	189	2,308
13	Outlier speed removed >6 and <200	2,727,953	387,460	468	4,260	188	2,237	189	2,226
14	Outlier speed removed >3 and <97 percentile	2,591,638	365,936	467	4,259	188	2,210	189	2,190
15	Aggregated port to region	1,417,301	314,546	458	4,245	180		181	
16	Remove outlier <3 and > 97 percentile	1,318,142	289,347	457	4,242	180		181	

2.B.2 Trade Data

Trade data are merged with the trade cost database, gravity control, and fuel data. Unbalanced data are removed. Details are shown in Table 2.B.2.

Table 2.B.2: Trade Data Processing

	Note	Nobs	Trade (mil.USD)	Origin	Dest
1	Raw	511,938	38,140	239	239
2	Tradecost	-190,305	-1,911	-217	-217
3	Gravity control	-90,090	-181	-239	-239
5	Removed no fuel	40,103	22,227	130	142
6	Removed unbalanced	20,502	21,286	113	125

2.C Fuel Prices

Fuel prices at each port are extrapolated using the following equation

$$f_{pt} = a_1 MaxDraft_p + a_2 MaxLOA_p + a_3 TimeZone(ME = 0)_p + \gamma_{year} + \gamma_{quarter} + \epsilon_{pt}. \tag{2.C.1}$$

$MaxDraft_p$ and $MaxLOA_p$ are the maximum drafts and overall length that port p can accommodate. The bigger these values, the larger the port. $TimeZone(ME = 0)$ indicates the time zone of the ports, adjusted so that the time zone of the Middle East is equal to 0.

The remaining missing data are filled with the average region's prices. Estimation results are reported in Table 2.C.2. Fuel prices on a voyage are the average of fuel prices of all ports along the voyage. Fuel prices between regions/countries are the average fuel prices on voyages between those regions/countries.

Table 2.C.1: Summary Statistics for Bunkering Ports

Statistic	N	Mean	St. Dev.	Min	Max
Max Draft	2,621	9.858	4.681	1.500	32.300
Max LOA	1,878	213.621	89.271	5.000	510.000
TimeZone	6,495	−2.037	5.106	−15.000	9.000

Source: IHS Markit

Table 2.C.2: Estimation Results: Fuel Prices at Port Level

Dependent Variable:	Fuel Price			
Model:	(1)	(2)	(3)	(4)
Maximum Draft	-2.530***	-2.530***	-2.530***	-2.530***
	(0.8089)	(0.7964)	(0.7699)	(0.6571)
Maxium LOA	0.0587	0.0587	0.0587	0.0587
	(0.0395)	(0.0389)	(0.0377)	(0.0386)
Time Zone (ME=0)	2.160***	2.160***	2.160***	2.160***
	(0.5190)	(0.5110)	(0.5108)	(0.7775)
SE type	NW(1)	NW(1)	NW(2)	Conley(10^3)
Observations	424	424	424	424
R^2	0.91199	0.91199	0.91199	0.91199
Within R^2	0.06116	0.06116	0.06116	0.06116

All regressions have year and quarter-fixed effects.
All SEs are adjusted for a small sample except for column 2. NW(k) is Newey-West with k lag. Conley(10^3) accounts for spatial correlation within 1000km.

2.D Monopolistic Shipping Market

This section shows the underlying theoretical motivation for aggregation delivery day from ship level to country level as in equation (2.24). Assuming that a service aggregator consolidates shipping services on a specific route ij from M_{ij} ships that have size z following a distribution $G(z)$ on $(0, \infty)$, the aggregated shipping demand is

$$Q_{ij} = \int_0^\infty q_{ij}(z) M_{ij} dG(z) \tag{2.D.1}$$

where

$$q_{ij}(z) = Y_j P_j^{\sigma-1} e^{-\tau\sigma \ln D_{ij}(z)} \Big(p_j t_{ij} + FR_{ij}(z) \Big)^{-\sigma}. \tag{2.D.2}$$

Hence, the partial derivatives of log demand with respect to freight rate and delivery days are

$$\frac{\partial ln q_{ij}(z)}{\partial FR_{ij}(z)} = \frac{-\sigma}{t_{ij} p_j + FR_{ij}(z)} \quad \text{and} \quad \frac{\partial ln q_{ij}(z)}{\partial D_{ij}(z)} = -\tau\sigma D_{ij}(z)^{-1}. \tag{2.D.3}$$

The marginal cost function for each ship z is

$$MC_{ij}(z) \equiv \frac{\partial TC_{ij}(z)}{\partial q_{ij}(z)} = f_{ij} D_{ij}(z)^{-\beta} d_{ij}^{\alpha} z_s^{-1}. \tag{2.D.4}$$

Optimal days (speed) and freight solve the profit maximization problem.

$$\Pi_{ij}(z) = q_{ij}(z)[FR_{ij}(z) - MC_{ij}(z)] - FC_{ij}. \tag{2.D.5}$$

Solving the first order condition with respect to freight yields

$$FR_{ij}(z) - MC_{ij}(z) = \frac{-1}{\frac{\partial \ln q_{ij}}{\partial FR_{ij}}} = \frac{p_j t_{ij} + FR_{ij}}{\sigma} \tag{2.D.6}$$

so that

$$FR_{ij}(z) = \frac{p_j t_{ij} + \sigma MC_{ij}(z)}{\sigma - 1}. \tag{2.D.7}$$

The first order condition with respect to day yields

$$FR_{ij}(z) - MC_{ij}(z) = \frac{\frac{\partial MC_{ij}(z)}{\partial D_{ij}}}{\frac{\partial \ln q_{ij}}{\partial D_{ij}}} = \frac{\beta f_{ij} D_{ij}^{-\beta} d_{ij}^{\alpha}}{\tau \sigma z}. \tag{2.D.8}$$

Combining two first-order conditions results in the optimal delivery days D_{ij}^*

$$D_{ij}^* = \left(A \frac{f_{ij} d_{ij}^{\alpha}}{p_j t_{ij} z} \right)^{1/\beta} \tag{2.D.9}$$

where $A \equiv \frac{\beta(\sigma-1) - \tau\sigma^2}{\tau\sigma(2\sigma-1)} > 0$ when $\frac{\beta}{\tau\sigma} > \frac{\sigma}{\sigma-1}$.

This shows the same qualitative results as in the main text. The higher f_{ij} the longer D_{ij}^*.

Substituting D_{ij}^* into the second The first order condition, FR_{ij} is then

$$FR_{ij}(z) = \frac{p_j t_{ij} + \sigma f_{ij} D_{ij}^{-\beta} d_{ij}^{\alpha} z^{-1}}{\sigma - 1} = B p_j t_{ij}. \tag{2.D.10}$$

where $B \equiv \frac{\beta - 2\tau\sigma^2}{\beta(\sigma-1)-\tau\sigma^2} > 0$ when $\frac{\beta}{\tau\sigma} > 2\sigma$.

The markup is

$$\frac{FR_{ij}(z) - MC(z)}{MC(z)} = \frac{\beta}{\tau\sigma}. \tag{2.D.11}$$

With a mass of firm M_{ij}, the price index is

$$\begin{aligned} P_i &= \left(\sum_i \int_0^\infty M_{ij} e^{-\tau\sigma \ln D_{ij}} p_{ij}^{1-\sigma} dG(z) \right)^{\frac{1}{1-\sigma}} \\ &= \left[\sum_j M_{ij}^{1-\sigma} p_{ij} \tilde{D}_{ij}^{-\tau\sigma} \right]^{\frac{1}{1-\sigma}} \end{aligned} \tag{2.D.12}$$

where

$$\tilde{D}_{ij} \equiv \left[\int_0^\infty D_{ij}(z)^{-\tau\sigma} dG(z) \right]^{\frac{1}{-\tau\sigma}}. \tag{2.D.13}$$

The operating profit is

$$r_{ij}(z) = (B+1)^{-\sigma} Y_i P_i^{\sigma-1} e^{-\tau\sigma \ln D_{ij}^*(z)} p_j^{-\sigma} t_{ij}^{-\sigma}. \tag{2.D.14}$$

Because $\frac{\partial r_{ij}}{\partial z} > 0$, $\exists z^*$ so that $r(z^*) \geq FC_{ij}$ for $z \geq z^*$.

Also, the operating profit ratio is the ratio of optimal delivery day, which is increasing in ship capacity z.

$$\frac{r_{ij}(z_1)}{r_{ij}(z_2)} = \left[\frac{D_{ij}^*(z_1)}{D_{ij}^*(z_2)} \right]^{-\tau\sigma}. \tag{2.D.15}$$

The average operating profit is the revenue evaluated at average delivery day $\tilde{D}_{ij}$, which can be shown as

$$\bar{r}_{ij} = r_{ij}(\widetilde{D}_{ij}) = \left[\frac{\widetilde{D}_{ij}^*(z^*)}{D_{ij}^*(z^*)} \right]^{-\tau\sigma} r_{ij}(D_{ij}^*(z^*)). \tag{2.D.16}$$

The zero cut-off profit pin downs z^* with given p_j

$$r_{ij}(D_{ij}^*(z^*)) = FC_{ij}. \tag{2.D.17}$$

The aggregate delivery day is

$$\widetilde{D}_{ij}(z^*) \equiv \left[\frac{1}{1-G(z^*)}\int_{z^*}^{\infty} D_{ij}(z)^{-\tau\sigma}dG(z)\right]^{\frac{1}{-\tau\sigma}}. \tag{2.D.18}$$

This is the capacity z weighted average of delivery days. In the empirical analysis, it is replaced by the size-weighted average.

Furthermore, the aggregate revenue is

$$R_{ij} = M_{ij}\int_{z^*}^{\infty} r_{ij}(z^*)dG(z) = M_{ij}\bar{r}. \tag{2.D.19}$$

The equation $R_{ij} = FC_e$ pins down M_{ij}, where H_{ij} is fixed entry cost for ij.

2.E Estimation Results for Robustness Checks

Table 2.E.1 shows results of equation (2.23) when using different coefficients for different routes to construct the IV as in equation (2.25) ($\sigma = 8$). Details were discussed in Section 2.6.

Table 2.E.2 shows results of equation (2.23) for different values of σ. Details were discussed in Section 2.7.1.

Table 2.E.1: Estimation Results: Adjusted Trade and Delivery Days ($\sigma = 8$) (Different Slow Steaming Coefficients by Route)

Dependent Variables:	ln(Day)	ln(Adj.Trade)
IV stages	First	Second
Model:	(1)	(2)
Day_IV	0.2443***	
	(0.0062)	
ln(Day)		-0.1256***
		(0.0362)
Language	-0.1336***	0.5680***
	(0.0431)	(0.0929)
C.Colony	-0.0404	0.1933
	(0.0671)	(0.1596)
Legal	-0.0012	0.2765***
	(0.0302)	(0.0645)
FTA	-0.3727***	0.3419***
	(0.0387)	(0.0817)
Observations	16,020	16,020
R^2	0.41214	0.83158
Within R^2	0.10026	0.04344
F-test (IV only)	1,395.7	12.718
Wu-Hausman, p-value		0.43165

Clustered (Pair) standard-errors in parentheses
Fixed effects: Exporter-Year, Importer-Year, Sea Route
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2.E.2: Estimation Results: Adjusted Trade and Days (IV)

σ	3	4	5	6	7	8	9	10
1st Stage :	Dependent Variable: ln(Day)							
Day_IV	0.6973***	0.6973***	0.6973***	0.6973***	0.6973***	0.6973***	0.6973***	0.6973***
	(0.0160)	(0.0160)	(0.0160)	(0.0160)	(0.0160)	(0.0160)	(0.0160)	(0.0160)
Language	-0.0727**	-0.0727**	-0.0727**	-0.0727**	-0.0727**	-0.0727**	-0.0727**	-0.0727**
	(0.0367)	(0.0367)	(0.0367)	(0.0367)	(0.0367)	(0.0367)	(0.0367)	(0.0367)
C.Colony	-0.2057***	-0.2057***	-0.2057***	-0.2057***	-0.2057***	-0.2057***	-0.2057***	-0.2057***
	(0.0570)	(0.0570)	(0.0570)	(0.0570)	(0.0570)	(0.0570)	(0.0570)	(0.0570)
Legal	0.0289	0.0289	0.0289	0.0289	0.0289	0.0289	0.0289	0.0289
	(0.0280)	(0.0280)	(0.0280)	(0.0280)	(0.0280)	(0.0280)	(0.0280)	(0.0280)
FTA	-0.4114***	-0.4114***	-0.4114***	-0.4114***	-0.4114***	-0.4114***	-0.4114***	-0.4114***
	(0.0324)	(0.0324)	(0.0324)	(0.0324)	(0.0324)	(0.0324)	(0.0324)	(0.0324)
R^2	0.40109	0.40109	0.40109	0.40109	0.40109	0.40109	0.40109	0.40109
F-test (IV only)	2,624.1	2,624.1	2,624.1	2,624.1	2,624.1	2,624.1	2,624.1	2,624.1
2nd Stage:	Dependent Variable: ln(Adj.Trade)							
ln(Day)	-0.0830**	-0.0827**	-0.0824**	-0.0821**	-0.0818**	-0.0814**	-0.0811**	-0.0808**
	(0.0324)	(0.0324)	(0.0324)	(0.0325)	(0.0325)	(0.0325)	(0.0326)	(0.0326)
Language	0.5160***	0.5140***	0.5121***	0.5101***	0.5082***	0.5063***	0.5043***	0.5024***
	(0.0834)	(0.0835)	(0.0835)	(0.0836)	(0.0836)	(0.0837)	(0.0837)	(0.0838)
C.Colony	0.3869***	0.3863***	0.3856***	0.3849***	0.3842***	0.3835***	0.3829***	0.3822***
	(0.1384)	(0.1385)	(0.1386)	(0.1387)	(0.1389)	(0.1390)	(0.1391)	(0.1393)
Legal	0.2113***	0.2131***	0.2149***	0.2167***	0.2185***	0.2202***	0.2220***	0.2238***
	(0.0635)	(0.0635)	(0.0636)	(0.0636)	(0.0636)	(0.0637)	(0.0637)	(0.0638)
FTA	0.6909***	0.6864***	0.6819***	0.6774***	0.6730***	0.6685***	0.6640***	0.6595***
	(0.0788)	(0.0789)	(0.0790)	(0.0791)	(0.0791)	(0.0793)	(0.0793)	(0.0794)
Observations	20,196	20,196	20,196	20,196	20,196	20,196	20,196	20,196
R^2	0.81138	0.81106	0.81074	0.81041	0.81008	0.80974	0.80939	0.80904
Within R^2	0.06818	0.06770	0.06722	0.06673	0.06624	0.06574	0.06525	0.06475
Wu-Hausman, p-value	0.00464	0.00490	0.00517	0.00546	0.00576	0.00608	0.00641	0.00677

Clustered (Pair) standard-errors in parentheses
Fixed effects: Exporter-Year, Importer-Year, Sea Route

Table 2.E.3: Adjusted Trade and Distance

Dependent Variable:	ln(Adj.Trade)							
	σ =3.301	σ =4	σ =5	σ =6	σ =7	σ =8	σ =9	σ =10
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Distance)	-0.4952***	-0.4922***	-0.4879***	-0.4836***	-0.4793***	-0.4750***	-0.4707***	-0.4664***
	(0.0394)	(0.0395)	(0.0396)	(0.0396)	(0.0397)	(0.0398)	(0.0399)	(0.0400)
Language	0.4888***	0.4876***	0.4860***	0.4843***	0.4826***	0.4810***	0.4793***	0.4776***
	(0.0815)	(0.0816)	(0.0816)	(0.0817)	(0.0818)	(0.0819)	(0.0820)	(0.0821)
C.Colony	0.3173**	0.3173**	0.3173**	0.3173**	0.3173**	0.3173**	0.3173**	0.3173**
	(0.1366)	(0.1367)	(0.1369)	(0.1371)	(0.1373)	(0.1375)	(0.1377)	(0.1379)
Legal	0.1927***	0.1941***	0.1960***	0.1979***	0.1998***	0.2018***	0.2037***	0.2056***
	(0.0617)	(0.0617)	(0.0618)	(0.0619)	(0.0619)	(0.0620)	(0.0621)	(0.0621)
FTA	0.4909***	0.4891***	0.4866***	0.4840***	0.4814***	0.4788***	0.4763***	0.4737***
	(0.0771)	(0.0772)	(0.0773)	(0.0774)	(0.0775)	(0.0777)	(0.0778)	(0.0780)
Observations	20,196	20,196	20,196	20,196	20,196	20,196	20,196	20,196
R^2	0.81721	0.81691	0.81648	0.81604	0.81559	0.81514	0.81469	0.81423
Within R^2	0.09730	0.09655	0.09549	0.09442	0.09335	0.09229	0.09122	0.09016

Clustered (Pair) standard-errors in parentheses
Fixed effects: Exporter-Year, Importer-Year, Sea Route

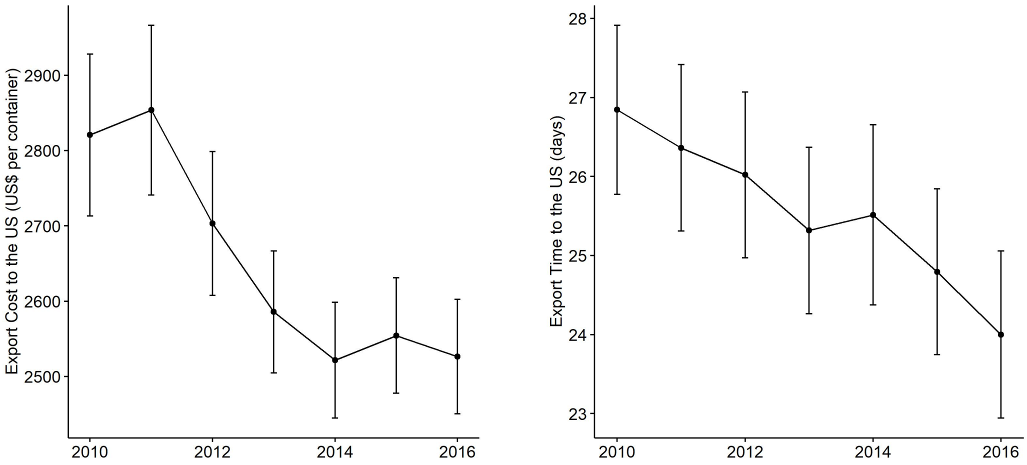
Chapter 3

Trade Costs and Different Margins of Trade

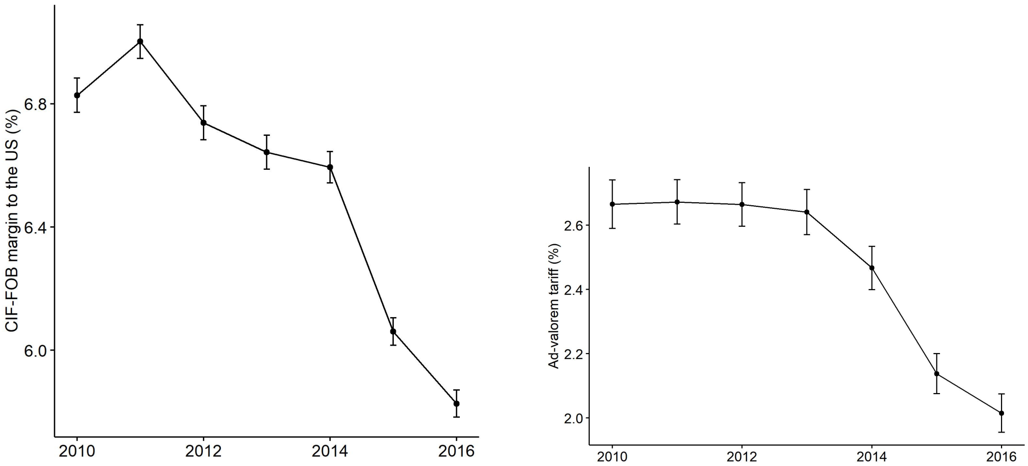
3.1 Introduction

The reduction in trade costs between developed and developing countries has contributed greatly to the "second unbundling" and the development of the global value chain (Baldwin and Lopez-Gonzalez, 2015). Figure 3.1 depicts the decreasing trend of export costs to the US in recent years. The administrative cost to export one container to the US has dropped by 16 percent on average from 2010 to 2016. The transportation cost (measured by the CIF-FOB margin) also dropped by 10 percent. Export time cost also decreased by 13 percent during this period. Similarly, the trade-weighted tariffs are also on a declining trend. Thanks to the reduction of these costs, sellers can send their goods at cheaper prices and in a more reliable and faster manner (Hummels, 2007; Hornok and Koren, 2015b). This study investigates how sellers adjust their decision on shipment frequency, size, and the number of buyers in response to the change in trade costs.

There have been a few papers investigating the effects of trade costs on ship-



(a) Administrative Cost **(b)** Export Time



(c) CIF-FOB Margin **(d)** Weighted Average Import Tariff

Figure 3.1: The Decline of Export Costs to the US

Source: Author's calculation using World Bank's Doing Business database, OECD's International Transport and Insurance Costs of Merchandise Trade database, and World Integrated Trade Solution

ment frequency and size. Békés and Muraközy (2012) shows that trade liberalization leads to an increase of extensive margin (more trade relationships) for both permanent and temporary exporters, where an adjustment in the intensive margin (average size) is more pronounced for permanent exporters. Hornok and Koren (2015b) and Kropf and Sauré (2014) investigate the frequency and size of shipment in response to a change in trade barriers. Hornok and Koren (2015b) shows that an increase in shipment costs reduces shipment frequency but increases shipment sizes. They formalize this idea in Hornok and Koren (2015a) by introducing a fixed cost per shipment. Still, their model cannot predict the relationship between shipment size and fixed cost per shipment because they assume all shipments have equal sizes. By adding storage costs on top of fixed costs per shipment, Kropf and Sauré (2014) solves for the optimal shipment frequency and size. They conclude that both variables are larger with higher market size and demand elasticity, but they are smaller with higher iceberg trade cost and production cost. These models share a similar concept that consumers value the timeliness of shipment, but each shipment incurs a fixed cost. The optimal frequency ultimately depends on the trade-off between this fixed cost and the cost of holding inventory. When each shipment is considered a transaction, this reflects the economic scale of transaction (Alessandria, Kaboski, and Midrigan, 2010).

There is extensive literature on how a change in trade barriers affects different trade margins. The extensive margins stemming from firms' heterogeneity are first introduced by Melitz (2003). Chaney (2008) introduces an extensive margin from seller heterogeneity, which refers to the change in the number of sellers in addition to the change in the size of each seller's export (the intensive margin). His main finding is that the elasticity of substitution has opposite effects on each margin. A decrease in trade barriers makes exports increase in an intensive margin higher than an extensive margin under high elasticity of substitution. Arkolakis (2010) incorporates convex marginal marketing costs into a trade model with seller heterogeneity to capture the higher growth rate of small existing sellers. This allows him to introduce the new consumer margin in addition to the intensive margin and new seller margin. Bernard, Moxnes, and Ulltveit-Moe (2018) also introduces buyer margin by assuming buyer heterogeneity ex-post, while ex-ante seller's unit cost of export is higher for smaller buyers. One of the most important messages from these papers is that the extensive margin of trade stems from seller and buyer heterogeneity.

This study contributes to the literature in two ways. First, the study presents a theoretical framework to show how iceberg trade costs and shipment costs affect not only shipment frequency and shipment size but also the number of buyers for each seller. The model is an extension of Bernard, Moxnes, and Ulltveit-Moe (2018) where sellers and buyers are heterogenous, and they need to make decisions about shipment frequency as in Kropf and Sauré (2014). In the model, iceberg trade costs affect trade directly through sales and indirectly through shipment frequency and the number of buyers, while shipment costs' effects are only indirectly through shipment frequency. Sellers with a small number of shipments may reduce shipment size when iceberg trade costs decrease if the adjustment through shipment frequency dominates the sales effects.

Second, the above effects are quantified using Bill of Lading (B/L) data of the US container import. The empirical results resonate Hornok and Koren (2015b), Hornok and Koren (2015a), and Kropf and Sauré (2014) in that trade barriers reduce shipment frequency. Furthermore, the empirical results show that shipment costs may increase average shipment size when the numbers of shipments are small, as predicted from the theoretical model.[1] In terms of magnitude, the effects of both trade barriers (iceberg trade costs and shipment costs) on trade volume are mostly from an increase in shipment frequency rather than shipment size; more than half of an increase in shipment frequency is from an increase in the number of buyers.

The next section describes the definition of shipment size and shipment frequency in the context of the B/L dataset and compares them with previous literature. Section 3.3 introduces a theoretical model based on Bernard, Moxnes, and Ulltveit-Moe (2018) and Kropf and Sauré (2014) to explore the effects of trade barriers on shipment size, shipment frequency, and the number of buyers. Section 3.4 introduces some testable hypotheses, estimation results of the grav-

[1] Hornok and Koren (2015b) also shows that shipment costs increase average shipment sizes using country-level trade data. However, they did not provide a theoretical explanation.

ity equation at the seller level and the seller-buyer level, and some discussions. Section 3.5 provides some robustness checks. Section 3.6 concludes.

3.2 Shipment Size and Frequency Definition

3.2.1 Data on Transactions

Even before data on detailed transactions became available, there had been numerous studies that provided evidence of infrequent shipments in international trade. Alessandria, Kaboski, and Midrigan (2010) uses the US monthly export to several countries to construct a Herfindahl-Hirschman index and finds that exports to most countries (except for Mexico) concentrate in certain months. Hornok and Koren (2015b) uses the US and Spanish export data that reports the number of shipments per month for each product and finds that a product is typically shipped only once or twice each month to a given destination. Kropf and Sauré (2014) uses Swiss Customs data and defines one transaction recorded in a custom form as a shipment. This allows them to examine shipment frequency at the seller level and find that a seller sends 3.5 shipments per year on average.

This study utilizes the US's container export at the Bill of Lading (B/L) level. B/L is proof that the seller has delivered the goods to the shipping companies and is crucial for a transaction to be successful. Each B/L provides information about the buyers, sellers, product classifications, quantities, arrival dates, and origin countries.[2]

Because the data contain all container inflow into the US, including commercial and non-commercial transactions, some data processing is needed to remove

[2]Each B/L records a transaction between a foreign seller and a domestic buyer, including information such as the buyer's name and address, the seller's name and address, product descriptions and quantities, departure and arrival ports, countries, and dates, information about shipping conditions such as ship number, shipping company, etc.

the latter. For buyers (consignee) and sellers (the shipper), only sellers and buyers indicated as companies are kept, and banks, transportation, military, government, embassy, and other non-commercial entities are excluded. All affiliates of the same companies in different states are treated as one buyer. Even though there may be different affiliates of the same companies in different countries that sell goods to the US, each shipper in a country is treated as a seller (for example, Honda Japan and Honda Thailand are treated as two different sellers).

For product classification, information about HS 2-digit was available.[3] Some transactions include more than two products but only transactions with one single product are kept.[4] The final dataset consists of about 27.7 million observations with the total quantity of 1.7 million tons of US container imports from 2010 to 2015. Each B/L is considered a shipment, which is also a transaction. In this sense, this study's approach is similar to that of Kropf and Sauré (2014), but different because the buyer of each transaction is identified.

3.2.2 Different Margins of Trade

Equipped with this data, a seller-level export to the US is decomposed into three dimensions: number of shipments, number of buyers, and average shipment size per shipment per buyer. Let s denote a single shipment, the yearly export volume X_{fig} of a seller f in country i for a product g to the US can be rewritten as

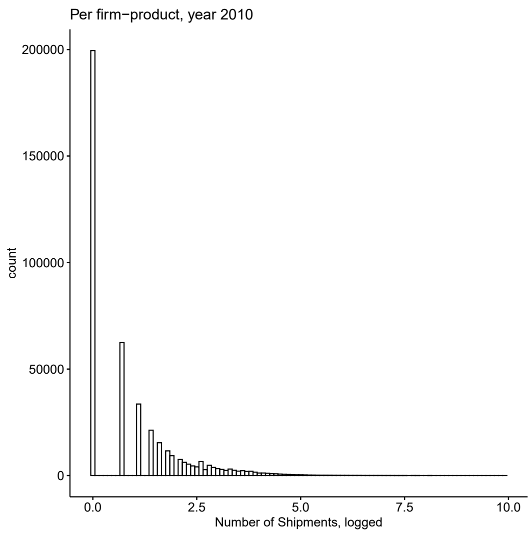
[3]HS is short for Harmonized System, a classification for products often used in international trade.

[4]Table 3.D.1 in Appendix 3.D provides information about data cleaning.

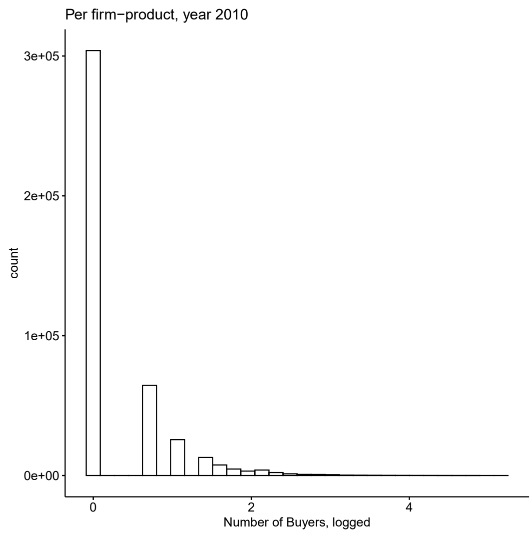
$$
\begin{aligned}
X_{fig} &= \sum_{b \in B_{fjg}} \sum_{s \in S_{fig}} q_{figbs} \\
&= N_{fig} \times \underbrace{\frac{\sum_{b \in B_{fig}} \sum_{s \in S_{fig}} q_{figb(s)}}{N_{fig}}}_{\bar{q}_{fig}} \\
&= B_{fig} \times \underbrace{\frac{N_{fig}}{B_{fig}}}_{\bar{N}_{fig}} \times \bar{q}_{fig}
\end{aligned} \tag{3.1}
$$

where N_{fig} is the actual total number of shipments of firm f from country i for product g, B_{fig} is the actual total number of buyers of firm f from country i for product g, q_{figbs} is the actual shipment size of firm f from country i for product g for each buyer b and shipment s, $\bar{N}_{fig}$ is the average number of shipments per buyer of firm f from country i for product g, $\bar{q}_{fig}$ is the average shipment size per shipment for seller f from country i of product g.

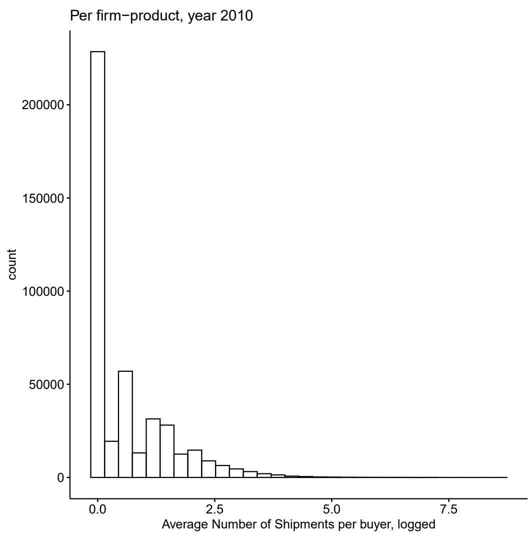
Figure 3.2 shows the histogram of shipment frequency (N_{fig}, $\bar{N}_{fig}$), number of buyers (B_{fig}), and average shipment size ($\bar{q}_{fig}$) as defined in equation (3.1) (all variables are in natural log form). Panel 3.2a shows that most sellers have only one transaction. Panel 3.2d shows that the average shipment size is mostly less than 150 tons per shipment. On average, a seller sends 8.3 shipments to 1.7 buyers yearly, which translates to 4.8 shipments yearly to a given buyer.



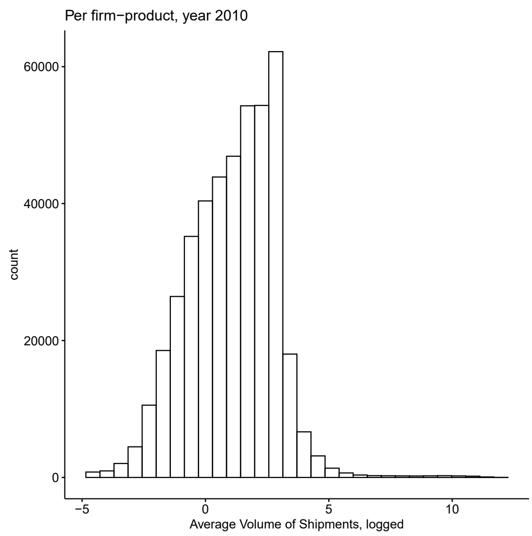
(a) Number of shipments



(b) Number of buyers



(c) A.N. Shipment per buyer

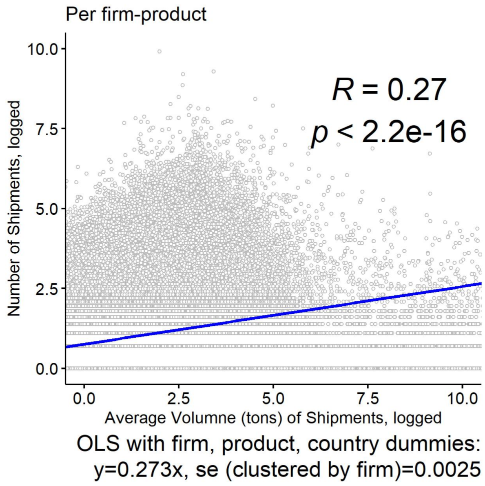


(d) Average Shipment Size

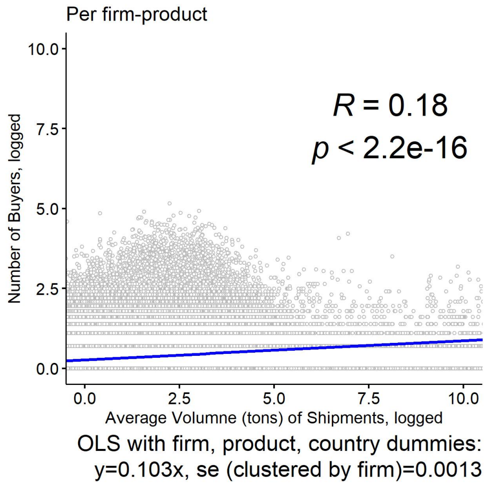
Figure 3.2: Different Trade Margins

Source: Author's calculation from US container import data

Panel 3.3a in figure 3.3 shows the correlation between shipment frequency (N_{fig}) and average shipment size ($\bar{q}_{fig}$), which is about 0.26 whether or not including the set of dummies (seller, buyer, country, product, pair). According to Kropf and Saure (2014), the significantly positive correlation between these two measures indicates the correct prediction of their model, in which shipment size and shipment frequency both increase in market sizes, and demand elasticity, and decrease with iceberg trade cost and production cost.[5]



(a) Number of Shipments



(b) Number of Buyers

Figure 3.3: Correlation between Shipment Frequency and Shipment Size

Source: Author's calculation from US container import data

Furthermore, the number of shipments can be expressed as the number of buyers and the average number of shipments per buyer as in equation (3.1). Panels 3.3b show the scatterplot between the number of buyers and the average shipment size. The fact that the average number of shipments per buyer significantly correlates with the average shipment size suggests a similar mechanism as in Kropf and Sauré (2014) at a seller-buyer level. In addition, the number of buyers also significantly correlates with the average shipment size. This suggests that the number of buyers also increases in the same set of variables, such as higher market size and demand elasticity, but decreases in higher iceberg trade cost and production cost. However, this has not been explored in the literature. The next section provides a theoretical framework to investigate this issue.

[5]However, an increase in fixed cost per shipment causes a decrease in shipment frequency but an increase in shipment size. In addition, shipment size decreases with a higher storage cost.

3.3 Theoretical Framework

3.3.1 Set Up

The model is based on Bernard, Moxnes, and Ulltveit-Moe (2018), where buyers from importing countries match sellers from exporting countries. The difference in this study is that they have to decide on the number of shipments after being matched.

Each country has L_i labor (provided in a perfect competition market), three sectors: homogeneous goods, traded intermediates goods (which require labor), and non-traded final goods (which require no labor). In the homogenous goods sector, goods are produced using 1-hour labor with a constant return to scale technology; the prices of these goods are normalized to 1 so that the wage rate becomes w_i.

Consumers' utility has two tiers: the upper tier is a Cobb Douglas utility function between homogeneous goods (with the budget share of $1-\mu$) and final goods (with the budget share of μ), and the lower tier is the final goods utility function with constant elasticity of substitution (CES) $\sigma > 1$.

In the intermediates sector, each variety is produced by a seller whose productivity z follows a Pareto distribution with the cumulative distribution function (CDF) function $F(z) = 1 - (z_L/z)^\gamma$ on $[z_L, \infty)$, where z_L is the scale parameter and γ is the shape parameter. A higher value of γ indicates less differences among seller's productivity (i.e. less heterogeneous).

In the final goods sector, each variety is produced by a buyer at time t whose technology is a CES function of all intermediates. This is a deviation from Bernard, Moxnes, and Ulltveit-Moe (2018) to incorporate the decision of shipment frequency.[6]

[6] This is similar to the idea in Kropf and Sauré (2014). However, in their model, consumers but not final producers, have to store goods if they are not arriving at exactly time t.

$$Z(b)\int_0^\infty e^{-\rho t}\left(\int_{\Xi_j(b)} c(t,\omega)^{(\sigma-1)/\sigma} d\omega\right)^{\sigma/(\sigma-1)} dt. \tag{3.2}$$

where ρ is a time discount factor.

Buyer b productivity follows a Pareto distribution with the cumulative distribution function $G(Z) = 1 - Z^{-\Gamma}$ on $[1, \infty)$, where Z_L is the scale parameter and Γ is the shape parameter. A higher value of Γ indicates less differences among buyer's productivity (i.e. less heterogeneous).

The variable $c(t, \omega)$ is the purchase of variety w and is used in production of final goods at time t. The set $\Xi_j(b)$ is a set of varieties available for buyer b in country j. The elasticity of substitution among intermediates is assumed to be the same as in the utility function ($\sigma > 1$).

Intermediate and final producers match with a fixed cost f_{ij}; hence there exists a lower bound $\underline{Z}_{ij}$ (and $\underline{z}_{ij}$), where the profit of a seller (and buyer) from a match is zero. There is no entry and positive profits are redistributed to consumers with dividend share ψ. Income in country i is then $w_i(1+\psi)L_i$. There is a fixed number of buyers (N_i) and sellers (n_i) in each country i. Intermediate goods are shipped with iceberg trade cost T_{ij}.

3.3.2 Prices

Let us denote $\bar{m} \equiv \sigma/(\sigma - 1)$ as the mark-up for intermediates. After a seller-buyer is matched, if intermediate goods were shipped at $t = 0$ but then used in production at $t = t'$, a storage cost is accrued at a rate $R > 0$ so the intermediate price becomes

$$p_{ij}(t') = e^{Rt'} \bar{m} T_{ij} w_i / z. \tag{3.3}$$

Final goods price is then $P_j = \bar{m}\phi_j(Z)/Z$ where $\phi_j(Z)$ is the price index for intermediate input in country j of final producer (buyer) Z.

$$
\begin{aligned}
\phi_j(Z)^{1-\sigma} &= \sum_k n_k \int_{\underline{z}} p_{kj}^{1-\sigma} dF(z) \\
&= \sum_k n_k (\bar{m} T_{kj} w_k)^{1-\sigma} \int_{\underline{z}} (e^{-Rt'_z} z)^{1-\sigma} dF(z).
\end{aligned}
\tag{3.4}
$$

Without the decision in shipment frequency, the two equations (3.3) and (3.4) collapse to those in Bernard, Moxnes, and Ulltveit-Moe (2018).

3.3.3 Match Pay-off

The intermediate goods export of a match is then

$$
r_{ij}(z, Z, t') = \left[\frac{p_{ij}(z,t')}{\phi_j(Z)}\right]^{1-\sigma} E_j(Z) = e^{(1-\sigma)Rt'} \underbrace{\left[\frac{\bar{m} T_{ij} w_i}{z \phi_j(Z)}\right]^{1-\sigma} E_j(Z)}_{\equiv r_{ij}(z,Z)}. \tag{3.5}
$$

where $E_j(Z)$ is the total spending on intermediate goods. It depends on the equilibrium pattern (see Appendix 3.A for the solution).

Let us denote Δ as the time between two shipments and $\zeta > 0$ as the interest rate. The net present value of a single shipment is

$$
\begin{aligned}
r_{ij}(z, Z, \Delta) &= \int_0^\Delta e^{-rt'} r_{ij}(z, Z, t') dt' \\
&= \int_0^\Delta e^{[(1-\sigma)R-\zeta]t'} r_{ij}(z, Z) dt' \\
&= \frac{1 - e^{-(\zeta+(\sigma-1)R)\Delta}}{\zeta + (\sigma-1)R} r_{ij}(z, Z).
\end{aligned}
\tag{3.6}
$$

If the reference period (a year) is normalized to one, the interval Δ is expressed as a fraction of a year. Therefore, the inverse of Δ (that is Δ^{-1}) is the number of shipments per year. Let us denote that $s_{ij} \equiv e^{-\Delta_{ij}}$. Because s_{ij} in increasing in Δ_{ij}^{-1}, it can also represent the number of shipments.

Assuming that F_{ij} is fixed cost per shipment, the net present value of all shipments is

$$r_{ij}^{NPV}(z,Z) = \frac{1}{1-s_{ij}^{\zeta}}\left[\frac{1-s_{ij}^{\zeta+(\sigma-1)R}}{\zeta+(\sigma-1)R} r_{ij}(z,Z) - F_{ij}\right]. \tag{3.7}$$

After intermediate goods sellers and the final producers are matched, if there is no shipment cost, intermediate goods sellers send a flow of shipments to the destination countries so that they arrive at exactly the date of producing the final goods. If they don't ship on the exact day of producing the final goods, they suffer a loss from storage costs. There is a trade-off between paying more fixed costs to ship at a higher frequency or paying more storage costs to ship more varieties at a time. This is the basic concept in Kropf and Sauré (2014).

3.3.4 Optimal Shipment Frequency

On the condition of a match, the optimal $\bar{\Delta}$ is the solution of the first order condition with respect to Δ_{ij} for the above revenue.

$$\frac{\zeta + (\sigma-1)Rs_{ij}^{\zeta+(\sigma-1)R} - [\zeta+(\sigma-1)R]s_{ij}^{(\sigma-1)R}}{\zeta+(\sigma-1)R} - \frac{\zeta F_{ij}}{r_{ij}(z,Z)} = 0. \tag{3.8}$$

This is similar to equation (12) in Kropf and Sauré (2014) but differs in the term $r_{ij}(z,Z)$. In their model, the equivalent term represents market condition (market size, iceberg trade cost, product wage, and price index) for a seller exporting to the whole market j. In this model, it is the market condition for an intermediate goods seller exporting to a final good producer in market j. Therefore, this model inherits all of their results about the optimal number of shipments ($\bar{s}_{ij}$ or $\bar{\Delta}^{-1}$).

The left-hand side of equation (3.8) decreases in $s_{ij} \in (0,1)$. It is negative when $s_{ij} = 1$ and non-negative when $s_{ij} = 0$ (because $r_{ij}(z,Z) > [\zeta+(\sigma-1)R]F_{ij}$ from equation (3.7)). This argument is true for all values of $r_{ij}(z,Z) > 0$.

Therefore, there is an unique solution $s_{ij} \in (0,1)$ for equation (3.8), denoted as $\bar{s}_{ij}$. Notice that $\bar{s}_{ij}$ increases with a higher value of the match $r_{ij}(z,Z)$, which decreases in iceberg trade cost T_{ij}. However, $\bar{s}_{ij}$ decreases with higher shipment costs F_{ij}.

Proposition 3.1. *Trade liberalization (decreasing iceberg trade costs and shipment costs) induces sellers to increase the number of shipments for each buyer.*

Proof. Rewrite equation (3.8) by bringing the second term to the right-hand side:

$$\frac{\zeta + (\sigma-1)Rs_{ij}^{\zeta+(\sigma-1)R} - [\zeta + (\sigma-1)R]s_{ij}^{(\sigma-1)R}}{\zeta + (\sigma-1)R} = \frac{\zeta F_{ij}}{r_{ij}(z,Z)}. \tag{3.9}$$

The derivative of the left-hand side of equation (3.9) with respect to s_{ij} is

$$\frac{\partial LHS}{\partial s_{ij}} = (\sigma-1)Rs_{ij}^{(\sigma-1)R-1}(s_{ij}^{\zeta}-1) < 0 \iff \sigma > 1 \quad \forall s_{ij} \in (0,1). \tag{3.10}$$

The derivative of the left-hand side of equation (3.9) with respect to T_{ij} is decreasing in s_{ij}.

$$\frac{\partial RHS}{\partial T_{ij}} = \underbrace{\frac{\partial RHS}{\partial r_{ij}}}_{<0} \underbrace{\frac{\partial r_{ij}}{\partial T_{ij}}}_{<0} > 0. \tag{3.11}$$

The second term follows the definition of $r_{ij}(z,Z)$ in equation (3.5). Therefore, the right-hand side is increasing in iceberg trade costs T_{ij}. A decrease in T_{ij} must be compensated by an increase in $\bar{s}_{ij}$.

It is straightforward to check that the right-hand side of the equation (3.9) increases in shipment cost F_{ij}. Similarly, a decrease in F_{ij} must be compensated by an increase in s_{ij}.

□

3.3.5 Shipment Size

The quantity for each shipment at time t' is $r_{ij}(z, Z, t')/p_{ij}(t')$. Using $\bar{s}_{ij} = e^{-\bar{\Delta}_{ij}}$ and equation (3.3), the quantity in present value terms is

$$\begin{aligned} q_{ij}(z, Z) &= \int_0^{\Delta_{ij}} \frac{[p_{ij}(t')]^{-\sigma}}{[q_j(Z)]^{1-\sigma}} E_j(Z) dt' \\ &= \int_0^{\Delta_{ij}} e^{-\sigma R t'} \underbrace{\left[\frac{\bar{m} T_{ij} w_i}{z q_j(Z)}\right]^{1-\sigma} E_j(Z)}_{r_{ij}(z,Z)} dt' \\ &= \frac{r_{ij}(z, Z)}{\bar{m} T_{ij} w_i R}(1 - \bar{s}_{ij}^{\sigma R}). \end{aligned} \tag{3.12}$$

The quantity of each shipment increases in the value of the match $r_{ij}(z, Z)$ but decreases with storage costs R.

The price index in (3.4) can be rewritten as

$$\begin{aligned} \phi_j(Z)^{1-\sigma} &= \sum_k n_k (\bar{m} T_{kj} w_k)^{1-\sigma} \int_{\underline{z}} \int_0^{\Delta_{kj}} (e^{-R t'_z} z)^{1-\sigma} dt' dF(z) \\ &= \frac{1}{(\sigma-1)R} \frac{\gamma z_L^{\gamma}}{\gamma_2} \sum_k n_k (\bar{m} T_{kj} w_k)^{1-\sigma} \underline{z}_{kj}(Z)^{\gamma_2} [1 - \bar{s}_{kj}^{(1-\sigma)R}] \end{aligned} \tag{3.13}$$

where $\gamma_2 \equiv \gamma - (\sigma - 1)$.

3.3.6 Sorting Function

Substituting the expression of fixed shipment cost F_{ij} in equation (3.8) into equation (3.7), the net present value of a match becomes

$$r_{ij}^{NPV}(z, Z) = \frac{r_{ij}(z, Z)}{\zeta} \bar{s}_{ij}^{(\sigma-1)R}. \tag{3.14}$$

The zero profit condition for a match is

$$\begin{aligned} \Pi_{ij}(z, Z, \bar{\Delta}) &= r_{ij}^{NPV}(z, Z)/\sigma - w_i f_{ij} \\ &= \frac{r_{ij}(z, Z)}{\sigma \zeta} \bar{s}_{ij}^{(\sigma-1)R} - w_i f_{ij} = 0. \end{aligned} \tag{3.15}$$

Substituting the expression of $r_{ij}(z, Z)$ as in (3.5), the implicit expression for $\underline{Z}_{ij}(z)$ (the lowest productivity of a buyer whose match gives seller z a zero profit) is

$$\phi_j(\underline{Z})^{\sigma-1}E_j(\underline{Z}) = \sigma w_i f_{ij}(\bar{m}T_{ij}w_i)^{\sigma-1}z^{1-\sigma}\zeta\bar{s}_{ij}^{(1-\sigma)R}. \tag{3.16}$$

Solving equation (3.15), lower bound $\underline{Z}_{ij}(z)$ is[7]

$$\underline{Z}_{ij}(z) = \frac{T_{ij}w_i\Omega_j}{z}(w_i f_{ij})^{1/(\sigma-1)}\frac{\zeta\bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R} \tag{3.17}$$

with

$$\Omega_j \equiv \kappa_2\left(\frac{\sigma}{\kappa_3}\frac{\gamma}{\gamma_2}\sum_k n_k'(T_{kj}w_k)^{-\gamma}(w_k f_{kj})^{-\gamma_2/(\sigma-1)}\bar{s}_{kj}^{\gamma_2 R}[1-\bar{s}_{kj}^{(1-\sigma)R}]\right)^{1/\gamma} \tag{3.18}$$

where $\kappa_2 \equiv [(\sigma\gamma)/(\kappa_3\gamma_2)]^{1/\gamma_2}$ and $\kappa_3 \equiv \mu(\Gamma-\gamma)/\gamma$. The variable n_i' is the normalized measure of sellers so that $n_i = z_L^{-\gamma}n_i'$. A lower productivity threshold (z_L) indicates that there are more potential sellers.

Without the terms that has $\bar{s}_{ij}$, the express of $\underline{Z}_{ij}(z)$ is the same as in Bernard, Moxnes, and Ulltveit-Moe (2018). $\underline{Z}_{ij}(z)$ decreases in z, so more efficient sellers match with less efficient buyers. $\underline{Z}_{ij}(z)$ also decreases in the optimal frequency $\bar{s}_{ij}$. Because $\bar{s}_{ij}$ is decreasing in fixed cost per shipment (F_{ij}), $\underline{Z}_{ij}(z)$ is increasing in fixed cost per shipment. Higher shipment costs reduce shipment frequency, which increases the price index. This in turn increases the potential profit of a given seller so they can match with more buyers.

The price index of the final producer Z becomes[8]

$$\phi_j(Z)^{1-\sigma} = Z^{\gamma_2}\bar{m}^{1-\sigma}\frac{\kappa_3}{\sigma\zeta}\left[\frac{\zeta\bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R}\right]^{\sigma-1}\Omega_j^{\sigma-1}. \tag{3.19}$$

[7]Details are in Appendix 3.A.

[8]Details are in Appendix 3.A.

E_i can be expressed as[9]

$$E_i(Z) = \kappa_3 Z^{\gamma}. \tag{3.20}$$

3.3.7 Seller-Buyer Level Trade

Substituting $E_j(Z)$ in equation (3.20) and $\phi_j(Z)$ in equation (3.19) into equation (3.5), $r_{ij}(z, Z)$ can be written as

$$\begin{aligned} r_{ij}(z, Z) &\equiv \left[\frac{\bar{m}T_{ij}w_i}{z\phi_j(Z)}\right]^{1-\sigma} E_j(Z) \\ &= \sigma\zeta\left(\frac{zZ}{T_{ij}w_i\Omega_j}\right)^{\sigma-1}\left[\frac{\zeta\bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R}\right]^{1-\sigma}. \end{aligned} \tag{3.21}$$

The value of the match (i.e. intermediate sale) is decreasing in iceberg trade costs T_{ij}. It increases in shipment frequency $\bar{s}_{ij}$ so it decreases in a fixed cost per shipment F_{ij}. This is achieved through an increase in the price index.

Shipment Size

Substituting the value of the match in equation (3.21) into equation (3.12), shipment size can be rewritten as

$$q_{ij}(z, Z) = \frac{\kappa_4\zeta^{2-\sigma}}{R}\frac{(zZ)^{\sigma-1}}{(T_{ij}w_i\Omega_j)^{\sigma}}\bar{s}_{ij}^{(\sigma-1)\gamma R}(1-\bar{s}_{ij}^{\sigma R}) \tag{3.22}$$

where $\kappa_4 \equiv \sigma(\sigma-1)^{\sigma-1}$.

It is unclear whether shipment size is decreasing or increasing in the number of shipments $\bar{s}_{ij}$. Taking the first derivatives of the log-transformed shipment size with respect to shipment frequency, it can be proved that the derivative is negative when $s_{ij} < \Lambda$ where $\Lambda \equiv \left[\frac{(\sigma-1)\gamma}{\sigma+(\sigma-1)\gamma}\right]^{1/(\sigma R)}$.

When the shipment frequency is small (i.e., less than the above threshold), shipment size and frequency move in opposite directions. The effects of shipment

[9] Details are in Appendix 3.A.

costs and iceberg trade costs are different. A decrease in shipment cost F_{ij} results in a positive adjustment from both shipment sizes and shipment frequency. However, a decrease in iceberg trade costs T_{ij} results in a positive adjustment in the value of the seller-buyer match (total sales of intermediates), but a negative adjustment through shipment frequency. As a result, the total effects of the iceberg trade costs, in this case, are ambiguous. Shipment size may decrease if the latter dominates the former.

On the other hand, shipment size and shipment frequency move in the same direction if the number of shipments is greater than the above threshold. In this case, a decrease in shipment cost F_{ij} or iceberg trade costs T_{ij} always increases in shipment size.

In addition, the productivity of sellers and buyers always increases the shipment size.

Proposition 3.2. *When the number of shipments is large enough (i.e., $s_{ij} > \Lambda$), trade liberalization (decreasing shipment costs or iceberg trade costs) induces sellers to increase shipment sizes for each buyer. When the number of shipments is small, the effect of decreasing iceberg trade costs is ambiguous. However, decreasing shipment costs always reduces shipment size.*

Proof. The total derivatives of log-transformed shipment size in equation (3.22) with respect to iceberg trade cost T_{ij} is

$$\begin{aligned} \frac{d \ln q_{ij}(z,Z)}{dT_{ij}} &= \frac{\partial \ln q_{ij}(z,Z)}{\partial T_{ij}} + \frac{\partial \ln q_{ij}(z,Z)}{\partial s_{ij}} \frac{\partial s_{ij}}{\partial T_{ij}} \\ &= -\sigma + \frac{\partial \ln q_{ij}(z,Z)}{\partial s_{ij}} \underbrace{\frac{\partial s_{ij}}{\partial T_{ij}}}_{<0}. \end{aligned} \tag{3.23}$$

The sign of the last term follows Proposition 3.1. The derivatives of $\ln q_{ij}(z, Z)$

with respect to s_{ij} is

$$\frac{\partial \ln q_{ij}}{\partial s_{ij}} = \frac{(\sigma-1)\gamma R}{s_{ij}} - \frac{\sigma R s_{ij}^{\sigma R-1}}{1-s_{ij}^{\sigma R}} > 0$$

when

$$\begin{aligned}
\frac{1-s_{ij}^{\sigma R}}{s_{ij}^{\sigma}} &< \frac{\sigma}{(\sigma-1)\gamma} \iff \\
\frac{1}{s_{ij}^{\sigma R}} &< \frac{\sigma+(\sigma-1)\gamma}{(\sigma-1)\gamma} \iff \\
s_{ij}^{\sigma R} &> \frac{(\sigma-1)\gamma}{\sigma+(\sigma-1)\gamma}.
\end{aligned}$$

Denote that $\Lambda \equiv \left[\frac{(\sigma-1)\gamma}{\sigma+(\sigma-1)\gamma}\right]^{1/(\sigma R)}$. When $s_{ij} > \Lambda$, the total derivative of log-transformed shipment size with respect to iceberg trade cost is negative. When $s_{ij} < \Lambda$, the sign of the derivative depends on the magnitude of each term in the right-hand side of equation (3.23).

The total derivatives of log-transformed shipment size in equation (3.22) with respect to shipment cost F_{ij} is

$$\begin{aligned}
\frac{d \ln q_{ij}(z,Z)}{dF_{ij}} &= \frac{\partial \ln q_{ij}(z,Z)}{\partial T_{ij}} + \frac{\partial \ln q_{ij}(z,Z)}{\partial s_{ij}} \frac{\partial s_{ij}}{\partial T_{ij}} \\
&= \frac{\partial \ln q_{ij}(z,Z)}{\partial s_{ij}} \underbrace{\frac{\partial s_{ij}}{\partial F_{ij}}}_{<0}.
\end{aligned} \tag{3.24}$$

The sign of the last term follows Proposition 3.1. The total derivative of log-transformed shipment size with respect to shipment cost is negative when $s_{ij} > \Lambda$ and vice versa. □

It may be counterintuitive for trade liberalization to reduce the average shipment size. The reason is that sellers reduce the shipment size but increase the number of shipments. The adjustment in the extensive margin (shipment frequency) may dominate the adjustment in the intensive margins (shipment size) when the number of shipments is small.

It is also worth noticing that the value of the threshold Λ is decreasing in storage costs R, but increasing in shape parameter of sellers' productivity distribution γ and the elasticity of substitution σ.[10]

Higher storage costs increase the seller's incentive to send smaller shipments more frequently. The smaller the shipment size, the less relevant it is to adjust shipment size when trade costs increase. At the same time, higher storage costs reduce the price index and potential profit. Shipment size does not have to increase to compensate for the reduction in shipment frequency. So the threshold Λ for such a case to happen becomes smaller.

Higher σ indicates a more competitive environment for sellers because demands for intermediates adjust strongly to an increase in price. Also, less frequent shipment increases prices for having to store more intermediate goods. Sellers can reduce the average price faced by buyers by increasing the shipment frequency and reducing shipment size. When trade costs increase, shipment size does not have to increase to compensate for the reduction in shipment frequency. So the threshold Λ for such a case to happen becomes smaller.

Higher γ indicates that sellers are not very different in their productivity. Demand for intermediates adjusts strongly to an increase in price under this circumstance. On the other hand, less frequent shipment increases price through storage costs. By increasing the shipment frequency and reducing shipment size, sellers can reduce the average price faced by buyers (by saving storage costs). Similar to the case of higher storage costs, shipment size does not have to increase to compensate for the reduction in shipment frequency. So the threshold Λ for such a case to happen becomes smaller.

To summarize, the threshold, under which shipment size and shipment fre-

[10] It can be proved that the derivative $\frac{\partial \ln \Lambda}{\partial \sigma} = \frac{1}{R}\left[- \frac{\ln \kappa_4}{\sigma^2} + \frac{1}{\sigma \kappa_4}\frac{\partial \kappa_4}{\partial \sigma}\right] > 0 \quad \forall \sigma > 1$ where $\kappa_4 \equiv \frac{(\sigma-1)\gamma}{\sigma+(\sigma-1)\gamma} < 1$.

quency move in the opposite direction, is smaller with higher storage costs, less heterogenous sellers, and more substitutability among intermediate goods.

Seller-level Yearly Trade Flow

The seller's yearly trade flow for each buyer is

$$X_{ij}(z,Z) = \underbrace{q_{ij}(z,Z)}_{ShipmentSize} \times \underbrace{\Delta_{ij}^{-1}}_{ShipmentFrequency} . \tag{3.25}$$

The derivative of seller-buyer trade volume with respect to shipment frequency Δ^{-1} is positive for all values of shipment frequency. However, the mechanisms are different depending on the value of shipment frequency. When $s_{ij} > \Lambda$, both shipment size and shipment frequency move in the same direction. A decrease in trade barriers increases both shipment size and shipment frequency, which increases seller-buyer trade volume. When s_{ij} is less than the above threshold, shipment size and frequency move in opposite directions. Reduction in iceberg trade costs may reduce shipment size through adjustment of shipment frequency. However, in this case, the gain from an increase in total sales always dominates this adjustment.

In addition, higher productivity of the seller and buyer increases the seller's yearly trade flow because this increases both shipment size and shipment frequency.

Proposition 3.3. *Trade liberalization (decreasing shipment costs and iceberg trade costs) increases seller-buyer-level trade.*

Proof. The total derivatives of log-transformed seller-buyer trade volume in equation (3.25) with respect to iceberg trade cost T_{ij} is

$$\frac{d \ln X_{ij}(z,Z)}{dT_{ij}} = \frac{\partial \ln X_{ij}(z,Z)}{\partial \Delta_{ij}^{-1}} \underbrace{\frac{\partial \ln \Delta_{ij}^{-1}}{\partial T_{ij}}}_{<0} . \tag{3.26}$$

The last term follows Proposition 3.1. The derivatives of $\ln X_{ij}(z, Z)$ with respect to Δ_{ij}^{-1} is

$$\frac{\partial \ln X_{ij}}{\partial \Delta_{ij}^{-1}} = \frac{\partial s_{ij}}{\partial \Delta_{ij}^{-1}} \left[\frac{(\sigma-1)\gamma R}{s_{ij}} - \frac{\sigma R s_{ij}^{\sigma R-1}}{1-s_{ij}^{\sigma R}} \right] + \frac{1}{\Delta_{ij}}.$$

Recall that $s_{ij} = e^{-\Delta_{ij}}$ so $\frac{\partial s_{ij}}{\partial \Delta_{ij}^{-1}} = s_{ij}$ and

$$\frac{\partial \ln X_{ij}}{\partial \Delta_{ij}^{-1}} = \underbrace{\left[(\sigma-1)\gamma R - \frac{\sigma R s_{ij}^{\sigma R}}{1-s_{ij}^{\sigma R}} \right]}_{\frac{\partial \ln q_{ij}}{\partial s_{ij}}} + \frac{1}{\Delta_{ij}}.$$

When $s_{ij} > \Lambda$, the first term is positive so $\frac{\partial \ln X_{ij}}{\partial \Delta_{ij}^{-1}} > 0$ (recall that $\Delta_{ij} > 0$). Hence, $\partial X_{ij}(z, Z)/\partial \Delta_{ij}^{-1} < 0$ so $\partial X_{ij}(z, Z)/\partial T_{ij} < 0$.

When $s_{ij} < \Lambda$, the first term is negative. However, the absolute value of this term is less than the second term as long as $\Delta^{-1}/R > 0$ (as proved below). Hence, $\partial X_{ij}(z, Z)/\partial \Delta_{ij}^{-1} < 0$ so $\partial X_{ij}(z, Z)/\partial T_{ij} < 0$.

$$\frac{1}{s_{ij}^{\sigma R}} > 1 + \frac{\sigma}{(\sigma-1)\gamma} > 1 + \frac{\sigma}{\Delta^{-1}/R + (\sigma-1)\gamma} \iff$$

$$\frac{1}{s_{ij}^{\sigma R}} - 1 > \frac{\sigma}{\Delta^{-1}/R + (\sigma-1)\gamma} \iff$$

$$\frac{s_{ij}^{\sigma} R}{1-s_{ij}^{\sigma R}} < \frac{\Delta^{-1} + (\sigma-1)\gamma R}{\sigma R} \iff$$

$$\frac{\sigma R s_{ij}^{\sigma R}}{1-s_{ij}^{\sigma R}} - (\sigma-1)\gamma R < \Delta^{-1}$$

The proof for shipment cost F_{ij} follows the same steps as above. □

3.3.8 Seller-Level Trade

Number of Buyers

The measure of buyers in country j for a seller in the country i with productivity $z < z_H$ (sellers that do not meet with every buyer in the market) is

$$
\begin{aligned}
b_{ij} &= N_j \int_{\underline{Z}_{ij}(z)} dG(Z) \\
&= Y_j (w_i f_{ij})^{-\Gamma/(\sigma-1)} \left(\frac{z}{T_{ij} w_i \Omega_j} \right)^{\Gamma} \left(\frac{r \bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R} \right)^{-\Gamma}.
\end{aligned} \tag{3.27}
$$

For these sellers, the number of buyers is increasing in $\bar{s}_{ij}$, which decreases in F_{ij}, so the number of buyers is decreasing in fixed cost per shipment. This is a new insight from this model. The intuition, however, is quite straightforward. When each shipment becomes more costly, sellers reduce the number of shipments, which incurs more storage costs. While this may increase shipment size, it always reduces the value of the match, which in turn drives out less productive buyers.

Because the number of shipments also decreases in iceberg trade costs T_{ij}, the number of buyers always decreases in iceberg trade costs.

Proposition 3.4. *The number of buyers decreases with shipment costs and iceberg trade costs.*

Proof. The total derivative of buyer numbers b_{ij} with respect to iceberg trade costs T_{ij} is

$$
\begin{aligned}
\frac{d \ln b_{ij}}{d T_{ij}} &= \frac{\partial \ln b_{ij}}{\partial T_{ij}} + \frac{\partial \ln b_{ij}}{\partial s_{ij}} \frac{\partial s_{ij}}{\partial T_{ij}} \\
&= -\Gamma + \frac{\partial \ln b_{ij}}{\partial s_{ij}} \underbrace{\frac{\partial s_{ij}}{\partial T_{ij}}}_{<0}.
\end{aligned} \tag{3.28}
$$

The last term follows Proposition 3.1. It is straightforward to check that $\partial b_{ij}/\partial s_{ij}$ is positive. Hence, $\partial b_{ij}/\partial T_{ij}$ is negative.

For shipment costs F_{ij}, the first term of equation (3.28) does not exist. The second term follows the same argument as iceberg trade costs. Hence, $\partial b_{ij}/\partial F_{ij}$ is negative. □

Seller-level Yearly Trade Flow

For sellers with $z < \underline{z}(Z_L) \equiv z_H$ (sellers that do not match with every buyer), total seller-level exports to the country j is

$$\begin{aligned} X_{ij}^{TOT}(z) &= N_j \int_{\underline{Z}_{ij}(z)} X_{ij}(z, Z) dG(Z) \\ &= \kappa_1 Y_j w_i f_{ij} \frac{\kappa_4 \zeta^{2-\sigma}}{R} \bar{s}_{ij}^{(\sigma-1)\gamma R} (1 - \bar{s}_{ij}^{\sigma R}) \underline{Z}_{ij}^{-\Gamma} \\ &= \kappa_1 Y_j (w_i f_{ij})^{1-\Gamma/(\sigma-1)} \frac{\kappa_4 \zeta^{2-\sigma-\Gamma}}{R^{1-\Gamma}} \left(\frac{z}{T_{ij} w_i \Omega_j} \right)^{\Gamma} \bar{s}_{ij}^{(\Gamma+\sigma-1)\gamma R} \frac{1 - \bar{s}_{ij}^{\sigma R}}{\Delta_{ij}} \end{aligned} \tag{3.29}$$

where $\kappa_1 \equiv (\sigma - 1)^{\Gamma} \sigma \Gamma / [\Gamma - (\sigma - 1)]$.

The derivative of seller trade volume with respect to shipment frequency Δ^{-1} is positive for all values of shipment frequency. However, the mechanisms are different depending on the value of shipment frequency. When $s_{ij}^{\sigma R} > \frac{(\Gamma+\sigma-1)\gamma}{(\Gamma+\sigma-1)\gamma+\sigma}$, both shipment size and shipment frequency move in the same direction, a decrease in trade barriers increase both shipment size and shipment frequency. When $s_{ij}^{\sigma R}$ is less than the above threshold, shipment size and frequency move in opposite directions. Reduction in iceberg trade costs may reduce shipment size through adjustment of shipment frequency. However, in this case, the gain from an increase in total sales always dominates this adjustment. Unlike the case of the seller-buyer level, there is a positive additional adjustment for the number of buyers. Therefore, there is no change in sign of trade barriers' effects due to the threshold.

The intuition for the change in the value of the thresholds is similar to Λ. The new element is Γ, the shape parameter of the buyer's productivity distribution. Higher Γ indicates a similarity in productivity among buyers. In this case, there are more potential buyers after a decrease in trade costs. With this new dimension, sellers do not have to sacrifice the reduction in shipment size in responding to an increase in shipment frequency.

Proposition 3.5. *Trade liberalization (decreasing shipment costs and iceberg trade costs) increases seller-level trade.*

Proof. The total derivative of yearly trade volume X_{ij}^{TOT} with respect to iceberg trade costs T_{ij} is

$$\begin{aligned} \frac{d \ln X_{ij}^{TOT}}{dT_{ij}} &= \frac{\partial \ln X_{ij}^{TOT}}{\partial T_{ij}} + \frac{\partial \ln X_{ij}^{TOT}}{\partial \Delta_{ij}^{-1}} \frac{\partial \Delta_{ij}^{-1}}{\partial T_{ij}} \\ &= -\Gamma + \frac{\partial \ln X_{ij}^{TOT}}{\partial \Delta_{ij}^{-1}} \underbrace{\frac{\partial \Delta_{ij}^{-1}}{\partial T_{ij}}}_{<0}. \end{aligned} \tag{3.30}$$

The last term follows Proposition 3.1.

The dervative of seller-level trade volume equation (3.29) with respect to Δ_{ij}^{-1} is

$$\frac{\partial \ln X_{ij}^{TOT}}{\partial \Delta_{ij}^{-1}} = \frac{\partial s_{ij}}{\partial \Delta_{ij}^{-1}} \underbrace{\left[\frac{(\Gamma+\sigma-1)\gamma R}{s_{ij}} - \frac{\sigma R s_{ij}^{\sigma R-1}}{1-s_{ij}^{\sigma R}} \right]}_{\frac{\partial \ln q_{ij}}{\partial s_{ij}} + \frac{\partial \ln b_{ij}}{\partial s_{ij}}} + \frac{1}{\Delta_{ij}}.$$

Recall that $s_{ij} = e^{-\Delta_{ij}}$ so $\frac{\partial s_{ij}}{\partial \Delta_{ij}^{-1}} = s_{ij}$ and

$$\frac{\partial \ln X_{ij}^{TOT}}{\partial \Delta_{ij}^{-1}} = \left[(\Gamma+\sigma-1)\gamma R - \frac{\sigma R s_{ij}^{\sigma R}}{1-s_{ij}^{\sigma R}} \right] + \frac{1}{\Delta_{ij}}.$$

When $s_{ij}^{\sigma R} > \frac{(\Gamma+\sigma-1)\gamma}{\sigma+(\Gamma+\sigma-1)\gamma}$, the first term is positive so $\partial \ln X_{ij}^{TOT}/\partial \Delta_{ij}^{-1} > 0$ (recall that $\Delta_{ij} > 0$). Hence, $\partial \ln X_{ij}^{TOT}/\partial T_{ij} < 0$.

When $s_{ij}^{\sigma R} < \frac{(\Gamma+\sigma-1)\gamma}{\sigma+(\Gamma+\sigma-1)\gamma}$, the first term is negative. However, the absolute value of this term is less than the second term as long as $\Delta^{-1}/R > 0$ (as proved below). So $\partial X_{ij}^{TOT}/\partial \Delta_{ij}^{-1} > 0$. Hence, $\partial \ln X_{ij}^{TOT}/\partial T_{ij} < 0$.

$$\frac{1}{s_{ij}^{\sigma R}} > 1 + \frac{\sigma}{(\Gamma+\sigma-1)\gamma} > 1 + \frac{\sigma}{\Delta^{-1}/R + (\Gamma+\sigma-1)\gamma} \iff$$

$$\frac{1}{s_{ij}^{\sigma R}} - 1 > \frac{\sigma}{\Delta^{-1}/R + (\Gamma+\sigma-1)\gamma} \iff$$

$$\frac{s_{ij}^{\sigma}R}{1-s_{ij}^{\sigma R}} < \frac{\Delta^{-1}+(\Gamma+\sigma-1)\gamma R}{\sigma R} \iff$$

$$\frac{\sigma R s_{ij}^{\sigma R}}{1-s_{ij}^{\sigma R}} - (\Gamma+\sigma-1)\gamma R < \Delta^{-1}$$

□

Average Shipment per Buyer

The average export per buyer per shipment (average shipment size per buyer) is

$$\bar{q}_{ij}(z) \equiv \frac{X_{ij}^{TOT}\Delta_{ij}}{b_{ij}} = \kappa_1 w_i f_{ij} \frac{\sigma r^{2-\sigma}}{R} \bar{s}_{ij}^{(\sigma-1)\gamma R}(1-\bar{s}_{ij}^{\sigma R}). \tag{3.31}$$

It is unclear whether the average export per buyer is decreasing or increasing in the number of shipments $\bar{s}_{ij}$. Taking the first derivatives of the log-transformed average export per buyer with respect to shipment frequency, it can be proved that the derivative is negative when $s_{ij} < \Lambda$. This is similar to Proposition 3.2. However, the difference is that there is no extra channel for iceberg trade costs T_{ij} to affect average export per buyer except for the adjustment in the shipment frequency. Therefore, the effects of T_{ij} on average shipment per buyer is not ambiguous when the shipment is small.

Proposition 3.6. *When the number of shipments is large enough (i.e., $s_{ij} > \Lambda$), trade liberalization (decreasing shipment costs and iceberg trade costs) induces an increase in average shipment size per buyer. When the number of shipments is small, trade liberalization (decreasing shipment costs and iceberg trade costs) reduces the average shipment size per buyer.*

Proof. The total derivatives of log-transformed average shipment size in equation (3.31) with respect to iceberg trade cost T_{ij} is

$$
\begin{aligned}
\frac{d \ln \bar{q}_{ij}(z)}{dT_{ij}} &= \frac{\partial \ln \bar{q}_{ij}(z)}{\partial T_{ij}} + \frac{\partial \bar{q}_{ij}(z)}{\partial s_{ij}} \frac{\partial s_{ij}}{\partial T_{ij}} \\
&= \frac{\partial \ln \bar{q}_{ij}(z)}{\partial s_{ij}} \underbrace{\frac{\partial s_{ij}}{\partial T_{ij}}}_{<0} .
\end{aligned} \tag{3.32}
$$

The first derivative is zero (which is different than the case of each seller-buyer shipment size). The sign of the last term follows Proposition 3.1. The derivative of $\ln \bar{q}_{ij}(z)$ with respect to s_{ij} is the same as the derivative of $\ln q_{ij}(z, Z)$ with respect to s_{ij}. The rest of the proof follows Proposition 3.2.

The proof for shipment cost F_{ij} follows the same logic. □

It is worth noticing that the adjustments on the number of buyers from both shipment costs and iceberg trade costs are canceled out so the intuition is similar to the case of buyer-seller shipment size. Sellers reduce the shipment size but increase the number of shipments. The adjustment in the extensive margin dominates the adjustment in the intensive margin when the number of shipments is small.

Similar to the case of seller-buyer level shipment size, the threshold, under which shipment size and shipment frequency move in the opposite direction, is smaller with higher storage costs, less heterogenous sellers, and more substitutability among intermediate goods.

3.4 Empirical Analysis

3.4.1 Hypothesis

This section explores the effects of shipment costs and iceberg trade costs on sellers' trade volume, the total number of shipments, average shipment size, number of buyers, shipment size per buyer, and number of shipments per buyer. From the analysis above, there are a few hypotheses that can be tested using the B/L data. Table 3.1 summarizes all the hypotheses and the corresponding propositions.

Table 3.1: Summary of Effects of Trade Barriers on Different Margins

	Seller Level				*Seller - Buyer Level*		
Hypothesis	1	2	3	4	5	6	7
Margins	Vol.	No.Ship.	Ave. Size	No.Buyer	Vol.	Size	No.Ship.
Iceberg Trade Costs							
Small No.Ship.	-	-	+	-	-	?	-
Large No.Ship.	-	-	-	-	-	-	-
Shipment Costs							
Small No.Ship.	-	-	+	-	-	+	-
Large No.Ship.	-	-	-	-	-	-	-
Proposition	3.5	3.1	3.6	3.4	3.3	3.2	3.1

The first four hypotheses look at the margins of trade at the seller level. Some of these have been examined in Kropf and Sauré (2014) and Bernard, Moxnes, and Ulltveit-Moe (2018). First, trade volume decreases with two trade costs (Hornok and Koren, 2015b; Kropf and Sauré, 2014). Second, the number of shipments decreases with two trade costs (Hornok and Koren, 2015b; Kropf and Sauré, 2014). Third, average shipment size increases with fixed shipment costs and iceberg trade costs when the number of shipments is small, and vice versa when the number of shipments is large. This differs from propositions 1 and 2 in Kropf and Sauré (2014) because there is an extra adjustment in the number of buyers in addition to the change in the number of shipments. Fourth, the number of buyers decrease with two trade costs (Bernard, Moxnes, and Ulltveit-Moe, 2018).

The last three hypotheses look at the margins of trade at the seller-buyer level. In the fifth hypothesis, trade volume at the seller-buyer-level decreases with both trade barriers (Bernard, Moxnes, and Ulltveit-Moe, 2018). The last two are newly investigated in this study. In the sixth hypothesis, shipment size per buyer increases with shipment costs when the number of shipments is small and vice versa, while iceberg trade costs may not decrease shipment size when the number of shipments is small. Finally, the number of shipments per buyer decreases with two trade barriers.

To test these hypotheses, the following equation is estimated for different margins of trade at seller level-product $y_{fig} = \{x_{fig}, S_{fig}, \bar{q}_{fig}, B_{fig}\}$ (seller's yearly trade volume, the number of shipments, average shipment size, and the number of buyers):

$$\ln y_{figt} = \alpha VC_{igt} + \beta FC_{it} + \delta_f + \delta_i + \delta_g + \delta_t + \varepsilon_{figt}. \tag{3.33}$$

VC_{it} and FC_{it} are iceberg trade costs and shipment costs. The variables δ_f are the seller fixed effect. The variables δ_i are the seller's country fixed effects. The variables δ_g are product (HS 2 digit) fixed effects. The variables δ_t are the year fixed effects.

For the seller-product-buyer level, the following equation is estimated for $y_{figb} = \{x_{figb}, S_{figb}, q_{figb}, \}$ (each buyer-seller match's yearly trade volume, the number of shipments for each match, and shipment size for each match):

$$\ln y_{figbt} = aVC_{igt} + bFC_{it} + \delta_f + \delta_i + \delta_g + \delta_t + \delta_b + \delta_{fb} + \epsilon_{figbt}. \tag{3.34}$$

The variables δ_b are the buyer fixed effects and δ_{fb} are the seller-buyer fixed effects.

The iceberg trade costs VC_{igt} contain the US's import tariff and CIF-FOB margin that represents the transportation cost.

The storage cost in the model is set to be the same for all buyers. Under this assumption, in equation (3.33) at the seller level, the time-fixed effects capture the time-variant components and product-fixed effects capture the time-invariant components of this cost.

In equation (3.33) at the buyer-seller level, the buyer-fixed effects additionally control for time-invariant components to be different among buyers. In section

3.5.1, the buyer-time fixed effects allow for the heterogeneity in storage costs for each buyer.

The variables y_{figt} and y_{figbt} are constructed from B/L data using equation (3.1). Taking natural log on both sides, the natural logarithm of bilateral trade (at seller level and seller-buyer level) is the log sum of these margins. Under the assumption that the error terms of the regression for each margin are not correlated with each other, the sum of coefficients α (β) for shipment frequency and shipment size should be equal to the coffecient α (β) for trade volume.

3.4.2 Data Summary

Table 3.2 describes the data summary being used to estimate equation (3.33) and (3.34) for different margins of trade at the country level, the seller level, and the seller-buyer level. These variables are constructed from B/L data using equation (3.1). For trade barriers, shipment costs are proxied by the administrative costs from the World Bank's Doing Business database (version 2015). To account for the change in US import costs, the shipment costs to the US are the sum of export costs to the US and import costs of the US.

Iceberg trade costs are proxied by the CIF-FOB margin from OECD's International Transport and Insurance Costs of Merchandise Trade database. Because the original data is reported at HS 4-digit, the margin is aggregated to HS 2-digit to match the B/L data using trade-weighted average and simple average. The simple average is reported here due to the many zeros in the trade-weighted average. The data is between 2010 and 2016.[11]

Tariff data is downloaded from World Integrated Trade Solution.[12] To match the container shipping data, the tariff is aggregated to HS 2-digit level using trade

[11]The CIF-FOB margin at the HS 4-digit level is calculated by dividing the gap between CIF and FOB price against the FOB price. These margins are used to aggregate from HS 4-digit to HS 2-digit data (denoted as CIF/FOB). In the regression (at HS 2-digit level), the aggregated margins are added by 1 before taking the natural logarithm, denoted as $\ln(1 + CIF/FOB)$.

[12]Details about tariff data sources are shown in Appendix 3.C.

share. However, because trade shares are not available for many country-pair-product triplets, results using the simple average is reported here. The summary statistics of both variables (simple average and weighted average) are shown in Table 3.2.

Table 3.2: Summary Statistics: the US' Container Imports

Statistic	N	Mean	Median	St. Dev.	Min	Max
At Country-Level						
Volume(t.tons)	42,601	40.9	0.3	391.6	0.000	19,901.8
N.Shipment	42,601	650.2	19	6,073.9	1	361,734
Avg Volume(t.tons)	42,601	0.4	0.01	3.7	0.000	138.9
N.Seller	42,601	78.7	5	808.3	1	53,751
N.Buyer	42,601	84.4	6	762.4	1	54,337
Avg N.Shipment	42,601	5.2	2.5	15.3	1.0	1,929.0
Avg N.Buyer	42,601	1.2	1.0	0.8	0.1	53.0
CIF/FOB (W.Avg)	42,601	0.1	0.1	0.04	0.0	0.6
CIF/FOB (S.Avg	42,601	0.1	0.1	0.03	0.0	0.5
Export Cost (USD)	42,601	2,344.5	2,201.7	753.5	1,522.7	10,656.1
W.Avg Tariff (%)	40,819	2.4	0.3	5.1	0.0	89.3
S.Avg Tariff (%)	40,819	2.5	0.8	5.0	0.0	89.3
At Seller-Level						
Volume(t.tons)	3,352,286	0.5	0.01	25.6	0.000	16,255.2
N.Shipment	3,352,286	8.3	2	58.8	1	20,133
Avg Volume(t.tons)	3,352,286	0.1	0.004	1.5	0.000	151.8
N.Buyer	3,352,286	1.7	1	2.5	1	320
Avg N.Shipment	3,352,286	3.9	1	25.6	1	15,145
At Seller-Buyer Level						
Volume(t.tons)	5,779,195	0.3	0.01	12.9	0.000	7,053.4
N.Shipment	5,779,195	4.8	1	39.9	1	19,973
Vol./Buyer(t.tons)	5,779,195	0.1	0.004	1.7	0.000	225.8
Single Buyer Dropped						
Volume(t.tons)	898,950	1.5	0.04	48.3	0.000	16,255.2
N.Shipment	898,950	21.4	7	102.9	2	20,133
Avg Volume(t.tons)	898,950	0.1	0.01	1.6	0.000	145.8
N.Buyer	898,950	3.7	2	4.2	2	320
Export Cost (USD)	898,950	1,956.1	1,810.0	375.7	1,522.7	10,656.1

As mentioned in section 3.2, many sellers only have one shipment (or one buyer). These sellers also have fewer shipments and less trade volume than those with more buyers. Table 3.3 shows that two-thirds of shipments in the sample are from single-buyer sellers. While the number of multiple-buyer sellers (M_{ig}) is only about one-fourth of single-buyer sellers, their number of shipments is more than twice, and their volume is three times higher, indicating the average shipment size is higher for multiple-buyer sellers. Dropping the single buyer from the sample, the medians of shipment frequency and average shipment sizes increase from 2 to 7 and 0.004 to 0.01, respectively. As shown in the theoretical part, sellers with a small number of shipments may respond to trade costs differently compared with sellers with a higher number of shipments. Equations (3.33) and (3.34) are also estimated for sub-sample with different thresholds on the number of buyers for each seller for the whole period.

Table 3.3: Single Buyer vs Multiple Buyers

B_{fig}	N.obs	S_{fig}	Q_{fig}	M_{ig}
1	2,453,336	8,480,093	428,295,549	845,957
>1	898,950	19,217,368	1,313,683,316	260,458

3.4.3 Empirical Results

Seller-Level Trade

Table 3.4 shows the results of equation (3.33) for the whole sample. Each column corresponds to the above hypothesis. Because half of the sample are sellers with single-buyers, the effects of trade barriers for small shipment frequency may dominate the total effects. All hypotheses are confirmed for columns 1, 2, and 4 because the coefficients for both trade barriers are significantly negative.

For average shipment size (column 3), iceberg transportation costs (CIF-FOB margin) are significantly positive but tariffs are significantly negative. This suggests that shipment frequency and shipment size move in the same direction when tariffs increase. On the other hand shipment frequency and shipment size move in the opposite direction when CIF-FOB margins increase. While it is puzzling for these two coefficients to be different, the total effects from iceberg trade costs (both tariff and CIF-FOB margin) are still positive.[13] This suggests that the positive adjustment through shipment frequency dominates the negative adjustment through sales when the number of shipments is small. However, the shipment costs (export costs) coefficients are not significant despite being positive.

Table 3.4: Seller-Level OLS (1 or more buyers)

Hypothesis No.:	1	2	3	4
Dependent Variables:	$\ln x_{fig}$	$\ln S_{fig}$	$\ln \bar{q}_{fig}$	$\ln B_{fig}$
$\ln(1 + Tariff)$	-1.152***	-0.4548***	-0.6975***	-0.0373
	(0.1270)	(0.0885)	(0.0715)	(0.0379)
$\ln(1 + CIF/FOB)$	-0.7198**	-2.030***	1.310***	-1.272***
	(0.2823)	(0.1938)	(0.1721)	(0.0951)
$\ln(ExportCost)$	-0.2245***	-0.2494***	0.0249	-0.1685***
	(0.0460)	(0.0286)	(0.0311)	(0.0147)
Observations	3,348,550	3,348,550	3,348,550	3,348,550
R^2	0.63286	0.41511	0.72590	0.38291
Within R^2	7.42×10^{-5}	0.00014	8.91×10^{-5}	0.00020

Clustered (Firm) standard-errors in parentheses
Fixed Effects: Firm, Year, Product, Country

The coefficients of tariff on the number of buyers show that the change in tariff does not affect the number of buyers. One potential explanation is that the value of Γ is very small (buyers are very different in productivity). So the change in tariff does not greatly affect the average prices faced by consumers. The potential revenue of the buyers (final goods producers) does not change much, hence there is less change in the number of buyers. Furthermore, the difference between the magnitude of the tariff and the CIF-FOB margin depends on the differences in the changes in shipment frequency due to the changes in

[13]Unfortunately, the theoretical model is not equipped to explain this situation. However, one could modify the model to separate tariff and ad-valorem shipping costs in a way so that the threshold Λ for the case of tariff is much smaller than for shipping costs. Conceptually, it is plausible that higher tariffs reduce both shipment frequency and shipment size in the same way but higher shipping costs may affect them in opposite ways.

these two variables (i.e. the last term in equation (3.28)). It is plausible that shipment frequency adjusts more strongly to an increase in transportation costs than tariff, so the coefficient of CIF-FOB margin is higher than that of the tariff.

The magnitudes of these coefficients also give more insight into how large each margin is. A one percent decrease in iceberg transportation costs increases the seller's yearly trade volume of an exporter by 0.7 percent, of which 177 percent is from the increase in average shipment size,[14] and −277 percent is from the decrease in the number of shipments[15] of which 62 percent[16] is from the reduction in the number of buyers.[17] The reduction in the number of buyers is large because the change in iceberg trade costs affects the number of buyers through adjustments in total sales and shipment frequency as discussed in Proposition 3.4.

To further test the idea that the coefficient of two types of trade barriers will not be positive when the number of shipments is higher, equation (3.33) is estimated for different sub-sample, where the average number of buyers for this whole period is greater or equal to a threshold (between 1 and 10). Figure 3.4 shows the magnitude of the coefficients and their 95% confidence interval.[18] The coefficients for iceberg transportation costs $\ln(1 + CIF/FOB)$ become insignificant when keeping sellers with more buyers (and more shipments). The coefficients for shipment costs $\ln(ExportCost)$ remain insignificant despite becoming negative. While this is not satisfactory because the coefficients do not become significantly negative, it at least shows that the positive effects of trade

[14] The share is calculated by dividing the coefficient of iceberg transportation costs for average shipment size against the coefficient of iceberg trade costs for yearly volume trade.

[15] The share is calculated by dividing the coefficient of iceberg transportation costs for the number of shipments against the coefficient of the iceberg transportation costs for yearly volume trade.

[16] The share is calculated by dividing the coefficient of iceberg transportation costs for the number of buyers against the coefficient of the iceberg transportation costs for the number of shipments.

[17] These shares can be interpreted this way because yearly trade volume can be decomposed into different margins by taking the natural log of both sides of equation (3.1). Furthermore, estimating these equations in OLS preserve the linear relationship among these margins.

[18] Detail estimation results are shown in the Appendix 3.B.

barriers on average shipment size diminish when there are more buyers (and more shipments).

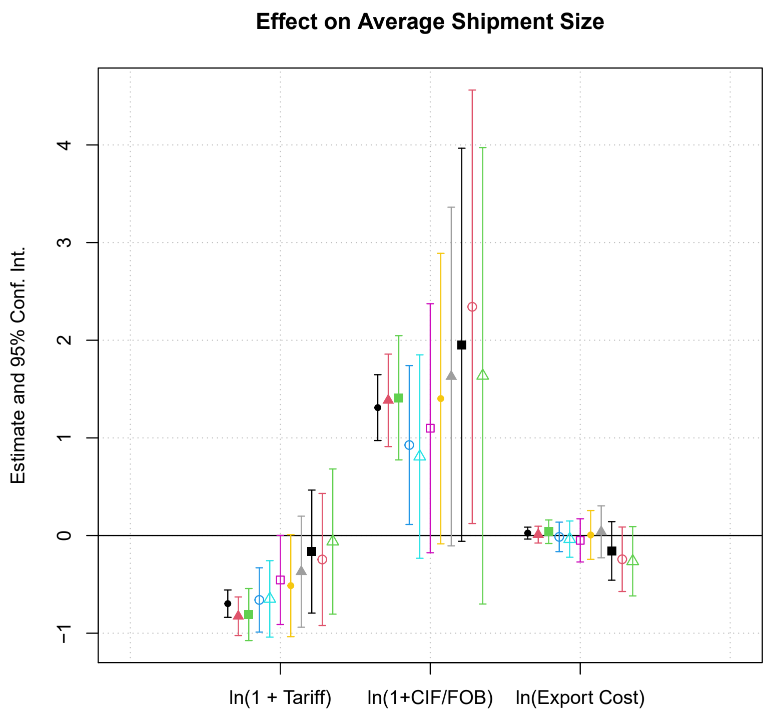


Figure 3.4: Estimation Results for Different Sub-samples with Tariff

From left to right: threshold for the number of buyers is from 1 to 10

The coefficients for tariff, however, remain negative for most cases. This shows that the effects of the CIF-FOB margin and tariff on average shipment size are different from each other for different sub-samples. However, the total effects from iceberg trade costs (the sum of both coefficients) reflect similar observations in the regression with the whole sample.

Seller-Buyer Level Trade

Table 3.5 shows the results of equation (3.34) for the whole sample. Each column corresponds to the above hypotheses. Because half of the whole sample are sellers with single buyer, the effects of trade barriers for small shipment frequency may dominate the total effects. For columns 1 and 2, all hypotheses are confirmed because the coefficients for both trade barriers are significantly negative. For column 3, both iceberg trade costs (CIF-FOB margin) and shipment costs (export costs) are insignificant.[19]

The coefficients for tariffs are significantly negative for trade volume and shipment frequency, which confirms the hypotheses derived from the theory. The coefficient of tariff on shipment size is significantly negative. This suggests that the negative effects of a higher tariff on sales dominate the positive effects through shipment size.

Table 3.5: Seller-Buyer-Level OLS

Hypothesis No:	5	6	7
Dependent Variables:	$\ln Q_{figb}$	$\ln S_{figb}$	$\ln \bar{q}_{figb}$
$\ln(1+Tariff)$	-0.6924***	-0.4434***	-0.2490***
	(0.1093)	(0.0834)	(0.0561)
$\ln(1+CIF/FOB)$	-1.690***	-1.534***	-0.1562
	(0.2314)	(0.1712)	(0.1284)
$\ln(ExportCost)$	-0.2476***	-0.2581***	0.0105
	(0.0390)	(0.0277)	(0.0242)
Observations	5,774,666	5,774,666	5,774,666
R^2	0.84556	0.60895	0.91370
Within R^2	8.7×10^{-5}	0.00013	1.23×10^{-5}

Clustered (Firm-Buyer) standard-errors in parentheses
Fixed Effects: Firm, Buyer, Year, Product, Country, Firm-Buyer

Examining the magnitudes of these coefficients, a one percent decrease in iceberg transportation costs increases the seller-buyer's yearly trade volume by 1.7 percent, most of which is from the increase in shipment frequency. A one percent decrease in shipment costs increases the seller-buyer's annual trade volume by 0.3 percent, most of which is from the increase in shipment frequency. The trade elasticities of iceberg trade costs (both CIF-FOB margin and tariff) are higher than the trade elasticity of shipment costs. This reflects the fact that iceberg

[19]Results for sub-sample (dropping buyer-seller pair whose average yearly numbers of shipments are less than 2 to 10) are similar to the case of the whole sample.

trade costs affect trade volume directly through an increase in intermediate sales and indirectly through an increase in shipment frequency. In contrast, shipment costs only affect these margins indirectly through an increase in shipment frequency.

3.4.4 Summary of Results and Discussion

The above results are summarized in table 3.6. All hypotheses are confirmed. Trade liberalization (decreasing iceberg trade costs or shipment costs) increases total trade volume at both the seller level as well as seller-buyer level. Controlling for buyer's characteristics, the effects are mostly from an increase in shipment frequency rather than shipment size. However, considering the adjustment of buyer margin, more than half of an increase in shipment frequency is from an increase in the number of buyers. This highlights the new insights of this study. It also justifies the theoretical model setting so that sellers choose the match before deciding shipment frequency and shipment sizes.

Table 3.6: Different Margins of Trade: Summary of Results

	Seller Level				*Seller - Buyer Level*		
Hypothesis	1	2	3	4	5	6	7
Margins	Vol.	No.Ship.	Ave. Size	No.Buyer	Vol.	Size	No.Ship.
All sample							
Tariff	-1.2	-0.5	-0.7	(-0.04)	-0.7	-0.4	-0.3
Transport	-0.7	-2.0	+1.3	-1.27	-1.7	(-0.2)	-1.5
Shipment	-0.22	-0.25	(0.02)	-0.17	-0.25	(0.01)	-0.26
4 or more buyers							
Tariff	-1.06	-0.4	-0.7	-0.13			
Transport	-6.2	-7.14	0.9	-5.2			
Shipment	(-0.1)	(-0.08)	(-0.01)	-0.2			
Proposition	3.5	3.1	3.6	3.4	3.3	3.2	3.1

Coefficients in bracket are insignificant.

The results also show that the elasticity of shipment costs is relatively small compared to the elasticity of iceberg transportation costs, which may sound puzzling. The measurement of shipment costs used in this estimate may underestimate the actual shipment costs. The data is from World Bank's Doing Business survey,[20] which excludes non-pecuniary costs. Under the reasonable assumption of storage costs ($R = 0.35$), Kropf and Sauré (2014) shows that country-average shipment costs can be twice higher as the average in the above survey. So with a rule of thumb, the elasticity of shipment costs could be twice higher, yet still smaller than the estimated elasticity of iceberg transportation costs.[21] This justifies the theoretical models where iceberg trade costs affect the seller's trade volume directly through sales and indirectly through the number of buyers and shipment frequency. On the other hand, the shipment costs' effects are only directly through shipment frequency.[22]

To interpret the threshold Λ empirically, let us rewrite the threshold in terms of shipment frequency in a year ($365/\Delta$). Using the definition $s_{ij} \equiv e^{-\Delta}$ so the value of Δ at the threshold Λ is $\Delta^{Threshold} = -\ln \Lambda$. The threshold for the number of shipments in a year is

$$\Lambda_2 \equiv \frac{365}{\Delta^{Threshold}} = -365 \ln \Lambda = \frac{-365}{\sigma R} \ln \left[\frac{(\sigma - 1)\gamma}{(\sigma - 1)\gamma + \sigma} \right]. \tag{3.35}$$

Table 3.7: Shipment Frequency Thresholds

$\sigma = 3$	$\sigma = 4$	$\sigma = 5$	$\sigma = 6$	$\sigma = 7$	$\sigma = 8$	$\sigma = 9$	$\sigma = 10$
110.70	75.00	56.72	45.60	38.13	32.76	28.72	25.56

The parameters in the previous studies are: the elasticity of substitution $\sigma \in (3, 10)$,[23] the shape parameter of sellers' productivity distribution $\gamma = 4$,[24] and annualized storage costs $R = 0.35$.[25] Using these parameters in equation

[20] See Hornok and Koren (2015b) and Hornok and Koren (2015a) for a detailed descriptions of the data.

[21] Using country-level trade data Hornok and Koren (2015a) estimated the elasticity of shipment costs and distance to be -1.063 and -1.298, respectively. Because distance can be a proxy for iceberg trade costs (Head and Mayer, 2014), the comments on the relative magnitudes of the two elasticities seem to be relevant.

[22] It is worth pointing out that the argument holds for the combined effects of both CIF-FOB margins and tariffs.

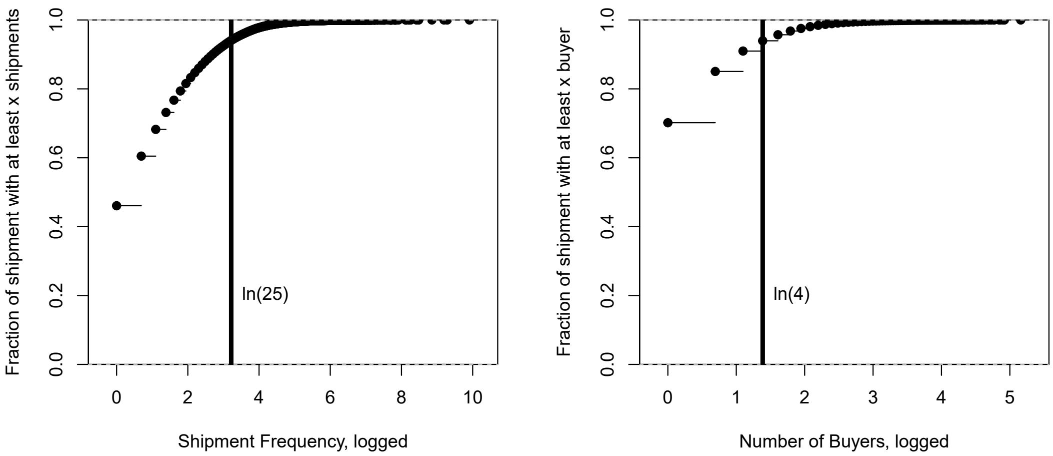
[23] See Anderson and Van Wincoop (2004, p. 716) and Hummels and Schaur (2013, p. 2949).

[24] See Bernard, Moxnes, and Ulltveit-Moe (2018, p. 437).

[25] See Kropf and Sauré (2014, p. 177).

(3.35), the thresholds for the number of shipments in a year is between 25 and 110 shipments (Table 3.7). For the minimum value of 25, more than 90 percent of the sample lie under this threshold (Figure 3.5a).

For the results from estimation, the thresholds are interpreted as the number of buyers. The point where the iceberg trade costs change sign is when the threshold is about four buyers. More than 90 percent of the sample lies under this threshold (Figure 3.5b). The threshold interpreted from the estimation covers a similar portion of the sample as the threshold calculated using paremeters in the literature.



(a) Shipment Frequency **(b)** Number of Buyers

Figure 3.5: Cummulative Distribution of Shipment Frequency and Number of Buyers in 2010

3.5 Robustness Checks

3.5.1 Heterogenous Storage Costs

In this section, regression equation (3.34) is modified to include the buyer-year fixed effects. This essentially relaxes the assumption that storage costs are the same for all buyers in equation (3.33) and that the storage costs differ only in time-invariant components in equation (3.34). Results are shown in Table 3.8.

The coefficients for export costs are smaller than in the case without buyer-year fixed effects. This suggests that the effects of shipment costs may have been contaminated by the effects of storage costs. In theory, higher shipment costs have similar effects to smaller storage costs (that is to reduce shipment frequency and trade volume). This explains the smaller coefficients in this case.

Table 3.8: Seller-Buyer-Level OLS with Buyer-Year Fixed Effects

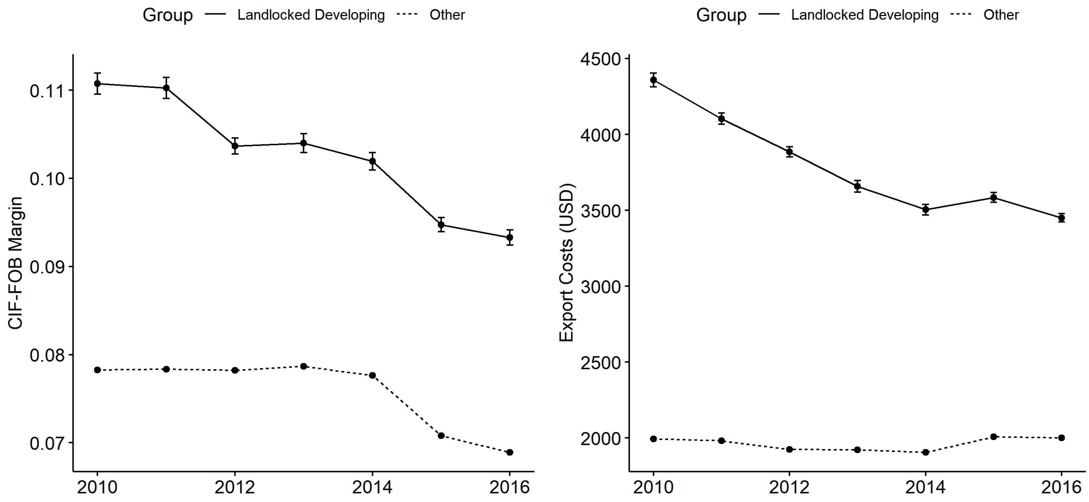
Hypothesis No:	5	6	7
Dependent Variables:	$\ln Q_{figb}$	$\ln S_{figb}$	$\ln \bar{q}_{figb}$
$\ln(1+Tariff)$	-1.110***	-0.7827***	-0.3275***
	(0.1441)	(0.1106)	(0.0716)
$\ln(1+CIF/FOB)$	-3.017***	-2.833***	-0.1838
	(0.3278)	(0.2501)	(0.1694)
$\ln(ExportCost)$	-0.1738***	-0.1772***	0.0034
	(0.0604)	(0.0427)	(0.0376)
Observations	5,774,666	5,774,666	5,774,666
R^2	0.87138	0.67092	0.93065
Within R^2	0.00017	0.00025	1.88×10^{-5}

Clustered (Firm-Buyer) standard-errors in parentheses
Fixed Effects: Firm, Buyer-Year, Product, Country, Firm-Buyer

3.5.2 Landlocked Developing Countries

This section addresses the concern that landlocked countries may be affected more by a change in shipping costs. Figure 3.6 shows the average CIF-FOB margin and export costs for landlocked developing countries.[26] The shipping costs of these countries are substantially higher than other countries despite a declining trend.

[26] These countries include: Afghanistan, Armenia, Azerbaijan, Bhutan, Bolivia, Botswana, Burkina Faso, Burundi, Central African Republic, Chad, Swaziland, Ethiopia (excludes Eritrea), Kazakhstan, Kyrgyz Republic, Lao PDR, Lesotho, Macedonia FYR, Malawi, Mali, Mongolia, Nepal, Niger, Paraguay, Moldova, Rwanda, Sudan, Tajikistan, Turkmenistan, Uganda, Uzbekistan, Zambia, and Zimbabwe.



(a) CIF-FOB Margin

(b) Export Costs

Figure 3.6: Shipping Costs for Landlocked Developing Countries

Table 3.9 shows the estimation results for two sub-samples: landlocked developing countries and other countries for the case of seller level. The results for the first sample show no significant variables. This is partly because there are very few observations.[27]

The other sub-sample for non-landlocked countries shows similar results to the main regression. Both trade barriers negatively trade volumes, shipment frequency, and the number of buyers. Iceberg transportation costs positively affect shipment size when all shipments with a small number of buyers are included in the regression sample. The trade elasticity of iceberg transportation costs is higher than the trade elasticity of shipment costs.

[27] Another specification with dummies for landlocked countries for the whole sample also shows no significant variables.

Table 3.9: Seller-Level OLS with Landlocked Developing Countries

Hypothesis No.:	1	2	3	4
Dependent Variables:	$\ln x_{fig}$	$\ln S_{fig}$	$\ln \bar{q}_{fig}$	B_{fig}
Landlocked Developing Countries				
$\ln(1 + CIF/FOB)$	1.250	3.296	-2.046	1.239
	(4.297)	(3.259)	(2.052)	(1.431)
$\ln(ExportCost)$	0.1263	0.0086	0.1178	0.0525
	(0.2709)	(0.1951)	(0.1720)	(0.1108)
Observations	5,899	5,899	5,899	5,899
R^2	0.85458	0.61648	0.92890	0.58799
Within R^2	0.00014	0.00073	0.00075	0.00045
Other Countries				
$\ln(1 + CIF/FOB)$	-0.7300***	-2.030***	1.300***	-1.272***
	(0.2832)	(0.1946)	(0.1725)	(0.0953)
$\ln(ExportCost)$	-0.2474***	-0.2610***	0.0136	-0.1717***
	(0.0469)	(0.0291)	(0.0318)	(0.0149)
Observations	3,346,387	3,346,387	3,346,387	3,346,387
R^2	0.63243	0.41471	0.72537	0.38253
Within R^2	1.67×10^{-5}	0.00012	3.75×10^{-5}	0.00020

Clustered (Firm) standard-errors in parentheses
Fixed Effects: Firm, Year, Product, Country

Table 3.10 shows the estimation results for two sub-samples for the case of seller-buyer level. While the results for the first sample show significance for trade volume and shipment frequency, the standard errors are very small. Similar to the case of seller level, this is partly because there are very few observations. The other sub-sample for non-landlocked countries shows similar results to the main regression.

Table 3.10: Seller-Buyer Level OLS with Landlocked Developing Countries

Hypothesis No:	5	6	7
Dependent Variables:	$\ln Q_{figb}$	$\ln S_{figb}$	$\ln \bar{q}_{figb}$
Landlocked Developing Countries			
$\ln(1 + CIF/FOB)$	-2.785***	-1.302***	-1.039
	(1×10^{-5})	(1×10^{-5})	(4.370)
$\ln(ExportCost)$	-0.5453***	-0.6367***	-0.1941
	(1×10^{-5})	(1×10^{-5})	(0.3613)
Observations	9,155	9,155	9,155
R^2	0.95986	0.84978	0.97704
Within R^2	0.00065	0.00074	0.00043
Other Countries			
$\ln(1 + CIF/FOB)$	-2.975***	-2.793***	0.9079***
	(0.3924)	(0.3153)	(0.1570)
$\ln(ExportCost)$	-0.1924***	-0.1889***	0.0120
	(0.0632)	(0.0452)	(0.0307)
Observations	5,770,040	5,770,040	5,770,040
R^2	0.87122	0.67058	0.81294
Within R^2	9.82×10^{-5}	0.00017	1.53×10^{-5}

Clustered (Firm-Buyer) standard-errors in parentheses
Fixed effects: Firm, Buyer-Year, Product, Country, Firm-Buyer

3.5.3 Different Estimation Method

This section examines the regression equation (3.33) and equation (3.34) by taking exponential transformation on both sides of these equations. The new equations are estimated using the pseudo-maximum likelihood (PPML) method.

Recent trade literature points out the benefits of estimating the gravity equations in multiplicative forms using the PPML method. This allows for including zero trade flows in the regression without having to manipulate the data (such as adding a small number to the zero trade flows in order to log-transform the dependent variables). In this chapter, it is not possible to consider the potential zero trade flows because the study uses data at the seller level. Another benefit of using PPML method is to address the potential heteroskedasticity (Silva and

Tenreyro, 2006). Therefore, it is worth checking whether the results using OLS hold in the case of PPML.

Results are shown in Table 3.11 for seller-level trade and Table 3.12 for seller-buyer level trade. Coefficients' signs change in the following case: tariff against the number of buyers and export costs against trade volume. Tariff against trade volume becomes insignificant. Other coefficients have the same sign as in the case of OLS. These remarks are the same for both regression equations.

Table 3.11: Seller-Level PPML

Hypothesis No.:	1	2	3	4
Dependent Variables:	$\ln x_{fig}$	$\ln S_{fig}$	$\ln \bar{q}_{fig}$	B_{fig}
$\ln(1 + Tariff)$	1.996	-0.7078	-2.564***	0.2066**
	(1.347)	(0.4650)	(0.7521)	(0.0888)
$\ln(1 + CIF/FOB)$	-5.295	-3.356***	4.226**	-2.729***
	(3.649)	(0.8464)	(1.722)	(0.2526)
$\ln(ExportCost)$	1.032***	-0.6442***	0.0842	-0.3647***
	(0.3267)	(0.0910)	(0.1351)	(0.0297)
Observations	3,348,550	3,348,550	3,348,550	3,348,550
Squared Correlation	0.81111	0.35478	0.84570	0.36916
Pseudo R^2	0.90563	0.56466	0.95040	0.18852

Clustered (Firm) standard-errors in parentheses
Fixed Effects: Firm, Year, Product, Country

Table 3.12: Seller-Buyer-Level PPML

Hypothesis No:	5	6	7
Dependent Variables:	$\ln Q_{figb}$	$\ln S_{figb}$	$\ln \bar{q}_{figb}$
$\ln(1 + Tariff)$	1.531	-0.2855	-1.088*
	(1.454)	(0.6737)	(0.6444)
$\ln(1 + CIF/FOB)$	-9.899***	-2.503**	-0.8060
	(3.798)	(1.114)	(1.332)
$\ln(ExportCost)$	0.7867*	-0.6009***	0.0207
	(0.4202)	(0.1306)	(0.1267)
Observations	5,774,666	5,774,666	5,774,666
Squared Correlation	0.73739	0.42136	0.94957
Pseudo R^2	0.93896	0.62494	0.98342

Clustered (Firm-Buyer) standard-errors in parentheses
Fixed effects: Firm, Buyer-Year, Product, Country, Firm-Buyer

It is puzzling that the effects of export costs on trade volume are positive in PPML estimation.[28] Silva and Tenreyro (2006) argues that the difference between OLS and PPML for positive trade can come from heteroskedasticity (the variance of error terms depends on the independent variable non-linearly) but the OLS estimator can still be valid when the error terms are proportional to the independent variable. However, it is not straightforward to examine the types of heteroskedasticity.

Silva and Tenreyro (2006) also argues that including the fixed effects in the gravity regression makes the bilateral trade costs coefficients vary for both PPML and OLS. Table 3.13 shows the results for yearly trade volume on the left-hand side of equation (3.33) with different combinations of fixed effects. The omission of product and year-fixed effects causes a greater change in the coefficients for both OLS and PPML. In particular, without controlling for product fixed effects, tariff coefficients become negative for PPML. This suggests that tariffs do not vary much within each product group. The coefficient of export cost in PPML becomes insignificant without controlling for firm-fixed effects. This suggests that firm-fixed effects capture some firm-specific export costs.

Changing fixed costs also affects the results in OLS. Coefficients of both tariff and CIF-FOB margin change greatly when removing product-fixed effects, year-fixed effects, and firm-fixed effects. It shows the importance of controlling for product and firm heterogeneity when estimating the effects of both iceberg trade costs (CIF-FOB margins and tariffs). Because these costs are only available at the country-product level, it is not possible to include the year interaction terms.

[28] Increasing the threshold for the number of buyers produces similar results.

Table 3.13: Seller-Level OLS and PPML with Differenchanget Fixed Effects for $\ln x_{fig}$

Model:	(1)	(2)	(3)	(4)	(5)
PPML					
$\ln(1 + Tariff)$	1.996	1.996	-5.248***	2.815**	1.991
	(1.347)	(1.347)	(1.191)	(1.422)	(1.300)
$\ln(1 + CIF/FOB)$	-5.295	-5.295	26.12***	-2.520	-3.369
	(3.648)	(3.648)	(3.920)	(2.776)	(6.015)
$\ln(ExportCost)$	1.032***	1.032***	1.087***	1.291***	0.3237
	(0.3267)	(0.3267)	(0.4141)	(0.3460)	(0.2920)
OLS					
$\ln(1 + Tariff)$	-1.152***	-1.152***	0.3700***	-0.7814***	-1.529***
	(0.1270)	(0.1270)	(0.0733)	(0.1193)	(0.1372)
$\ln(1 + CIF/FOB)$	-0.7198**	-0.7198**	4.773***	1.460***	0.8222***
	(0.2823)	(0.2823)	(0.1840)	(0.2461)	(0.3009)
$\ln(ExportCost)$	-0.2245***	-0.2245***	-0.0978**	-0.3150***	-0.5388***
	(0.0460)	(0.0460)	(0.0460)	(0.0399)	(0.0497)
Fixed-effects					
Firm	Yes	Yes	Yes	Yes	
Year	Yes	Yes	Yes		Yes
Product	Yes	Yes		Yes	Yes
Country	Yes		Yes	Yes	Yes

Clustered (Firm) standard-errors in parentheses

3.6 Conclusion

This chapter shows how sellers adjust shipment frequency and size in response to a change in trade barriers. By using a novel data set of B/L data that records container shipments into the US, a detailed investigation of the shipment size and the number of shipments at the buyer level, which was not available in previous literature, becomes available.

The data shows a positive correlation between average shipment size and shipment frequency, as has been explored by Kropf and Sauré (2014). It also shows a positive correlation between the number of buyers and average shipment size, which is a new observation. These suggest that these three variables are likely to correlate to the same set of variables, such as market size, trade costs, and production costs, in the same manner.

A theoretical model considering seller-buyer heterogeneity and shipment frequency was examined to provide insights into these decisions at the seller-buyer level. In the model, iceberg trade costs affect trade directly through sales and indirectly through shipment frequency and the number of buyers. Shipment costs' effects on trade volume are only through adjusting shipment frequency. Sellers with a small number of shipments may reduce shipment size even when iceberg trade costs decrease. This happens if the adjustment through shipment frequency dominates the sales effects. The theoretical models also provide testable hypotheses, which are examined using the B/L dataset.

The empirical analysis is carried out by estimating the gravity-like equation for the seller level and the buyer-seller level. The empirical results resonate Hornok and Koren (2015b), Hornok and Koren (2015a), and Kropf and Sauré (2014) in that trade barriers reduce the total shipment frequency of one seller from all buyers. It also shows that trade costs reduce the shipment frequency of each seller-buyer match. The number of buyers for each seller also decreases in trade barriers. Furthermore, the empirical results show that shipment costs may increase the average shipment size of a seller when the number of shipments is small, as predicted from the theoretical model. It is worth noting that both tariff and CIF-FOB margin are used as proxies for iceberg trade costs. While the effects of tariff and CIF-FOB margin on shipment size are not the same, the total effects from these two variables reflect the argument for iceberg trade costs.

In terms of magnitude, the effects of both trade barriers on trade volume are mostly from an increase in shipment frequency rather than shipment size; more than half of an increase in shipment frequency is from an increase in the number of buyers. The results at both the seller level and buyer-seller level show that the elasticity of iceberg trade costs can be twice higher as that of shipment costs. This reflects that iceberg trade costs affect trade volume directly through an increase in intermediate sales and indirectly through an increase in shipment frequency. In contrast, shipment costs only affect indirectly through an increase in shipment frequency.

In the robustness check, controlling for the possible differences in buyer-specific storage costs, the effects of shipment costs become smaller. While land-locked developing countries have higher shipment costs, their effects on different margins of trade are not confirmed for this group of countries. Finally, using the PPML method instead of OLS results in similar results except for the case of tariff and export costs on yearly trade volume.

For future research, more detailed data on the heterogeneity of buyers and sellers could provide more insights into how different types of firms adjust these margins. The potential relationship between buyers and sellers (such as an intra-firm relationship) could shed light on different patterns of shipment decisions.

Appendices

3.A Solution for Sorting Function

First we guess $\underline{z}_{ij}(Z) = W_{ij}Z^u$ and the inverse $\underline{Z}_{ij} = (z/W_{ij})^{1/s}$ and try to solve for W_{ij} and u. Similarly, we guess $E_j(Z) = \kappa_3 Z^\gamma$ and try to solve for κ_3. Inserting this expression into the expressions of the price indices in equation (3.13) and (3.16).

$$\frac{1}{\sum_k n_k(\bar{m}T_{kj}w_k)^{1-\sigma}\underline{z}_{kj}(Z)^{\gamma_2}[1-\bar{s}_{kj}^{(1-\sigma)R}]} = \frac{\gamma z_L^\gamma}{\gamma_2}\frac{\sigma w_i f_{ij}}{E_j(Z)}(\bar{m}T_{ij}w_i)^{\sigma-1}z^{1-\sigma}\frac{\zeta\bar{s}_{ij}^{(1-\sigma)R}}{(\sigma-1)R} \iff$$

$$\frac{\underline{Z}^{u\gamma_2+\gamma}}{\sum_k n_k(\bar{m}\tau_{kj}w_k)^{1-\sigma}W_{kj}^{-\gamma_2}[1-\bar{s}_{kj}^{(1-\sigma)R}]} = \frac{\gamma z_L^\gamma}{\gamma_2}\frac{\sigma w_i f_{ij}}{\kappa_3}(\bar{m}T_{ij}w_i)^{\sigma-1}z^{1-\sigma}\frac{\zeta\bar{s}_{ij}^{(1-\sigma)R}}{(\sigma-1)R} \iff$$

$$\frac{z^{1/u}}{\sum_k n_k(\bar{m}T_{kj}w_k)^{1-\sigma}W_{kj}^{-\gamma_2}[1-\bar{s}_{kj}^{(1-\sigma)R}]}\left(\frac{1}{W_{ij}}\right)^{1/u} = \frac{\gamma z_L^\gamma}{\gamma_2}\frac{\sigma w_i f_{ij}}{\kappa_3}(\bar{m}T_{ij}w_i)^{\sigma-1}z^{\frac{1-\sigma}{s\gamma_2+\gamma}}\frac{\zeta\bar{s}_{ij}^{(1-\sigma)R}}{(\sigma-1)R}.$$

To equate the expression that has z on both sides, it must be that

$$\frac{1}{u} = \frac{1-\sigma}{u(\gamma_2+\gamma/s)} \iff \frac{1}{u} = -1.$$

Then, we have

$$\left(\frac{1}{W_{ij}}\right)^{1/u} = \left[\frac{\gamma z_L^\gamma}{\gamma_2}\frac{\sigma w_i f_{ij}}{\kappa_3}(\bar{m}T_{ij}w_i)^{\sigma-1}\frac{\zeta\bar{s}_{ij}^{(1-\sigma)R}}{(\sigma-1)R}\times \sum_k n_k(\bar{m}T_{kj}w_k)^{1-\sigma}W_{kj}^{-\gamma_2}[1-\bar{u}_{kj}^{(1-\sigma)R}]\right]^{1/(u\gamma_2+\gamma)} \iff$$

$$W_{ij} = \left[\frac{\gamma z_L^\gamma}{\gamma_2} \frac{\sigma w_i f_{ij}}{\kappa_3} (T_{ij} w_i)^{\sigma-1} \frac{\zeta \bar{s}_{ij}^{(1-\sigma)R}}{(\sigma-1)R} \sum_k n_k (T_{kj} w_k)^{1-\sigma} W_{kj}^{-\gamma_2} [1 - \bar{s}_{kj}^{(1-\sigma)R}] \right]^{1/(\sigma-1)}. \tag{3.A.1}$$

The cut-off is

$$\underline{z}_{ij}(Z) = \frac{W_{ij}}{Z}. \tag{3.A.2}$$

Substituting the expression for cut off into equation (3.13), we have

$$\phi_j(Z)^{1-\sigma} = Z^{\gamma_2} \bar{m}^{1-\sigma} \frac{1}{(\sigma-1)R} \frac{\gamma z_L^\gamma}{\gamma_2} \sum_k n_k (T_{kj} w_k)^{1-\sigma} W_{kj}^{-\gamma_2} [1 - \bar{s}_{kj}^{(1-\sigma)R}].$$

Substituting the expression for W_{kj} from equation (3.A.1), the price index becomes

$$\phi_j(Z)^{1-\sigma} = Z^{\gamma_2} \bar{m}^{1-\sigma} \frac{\kappa_3}{\sigma w_i f_{ij} \zeta} \left[\frac{W_{ij}}{T_{ij} w_i \bar{s}_{ij}^{-R}} \right]^{\sigma-1}.$$

This must hold for all i, so that

$$(w_i f_{ij})^{-1/(\sigma-1)} \frac{W_{ij}}{T_{ij} w_i \bar{s}_{ij}^{-R}} = (w_k f_{kj})^{-1/(\sigma-1)} \frac{W_{kj}}{T_{kj} w_k \bar{s}_{kj}^{-R}}.$$

The expression for W_{ij} can be transformed as the following

$$W_{ij}^{\sigma-1} = \frac{\gamma z_L^\gamma}{\gamma_2} \frac{\sigma w_i f_{ij}}{\kappa_3} (T_{ij} w_i)^{\sigma-1} \frac{\zeta \bar{s}_{ij}^{(1-\sigma)R}}{(\sigma-1)R} \times$$
$$\sum_k n_k (T_{kj} w_k)^{1-\sigma} (T_{kj} w_k \bar{s}_{kj}^{-R})^{-\gamma_2} f_{kj})^{-\gamma_2/(\sigma-1)} \left((w_k f_{kj})^{-1/(\sigma-1)} \frac{W_{kj}}{T_{kj} w_k \bar{s}_{kj}^{-R}} \right)^{-\gamma_2} [1 - \bar{s}_{kj}^{(1-\sigma)R}]$$

$$\iff$$

$$W_{ij}^{\sigma-1} = \frac{\gamma z_L^\gamma}{\gamma_2} \frac{\sigma w_i f_{ij}}{\kappa_3} (T_{ij} w_i)^{\sigma-1} \frac{\zeta \bar{s}_{ij}^{(1-\sigma)R}}{(\sigma-1)R} \left((w_i f_{ij})^{-1/(\sigma-1)} \frac{W_{ij}}{T_{ij} w_i \bar{s}_{ij}^{R}} \right)^{-\gamma_2} \times$$
$$\sum_k n_k (T_{kj} w_k)^{-\gamma} (w_k f_{kj})^{-\gamma_2/(\sigma-1)} \bar{s}_{kj}^{-\gamma_2 R} [1 - \bar{s}_{kj}^{(1-\sigma)R}].$$

Recall that $\gamma_2 \equiv \gamma - (\sigma - 1)$, so

$$W_{ij}^\gamma = (T_{ij} w_i)^\gamma \frac{\sigma}{\kappa_3} (w_i f_{ij})^{\gamma/(\sigma-1)} \frac{\gamma z_L^\gamma}{\gamma_2} \frac{\zeta \bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R} \times$$

$$\sum_k n_k(T_{kj}w_k)^{-\gamma}(w_k f_{kj})^{-\gamma_2/(\sigma-1)}\bar{s}_{kj}^{\gamma_2 R}[1-\bar{s}_{kj}^{(1-\sigma)R}] \iff$$

$$W_{ij} = T_{ij}w_i(w_i f_{ij})^{1/(\sigma-1)} z_L \frac{\zeta \bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R} \times$$

$$\left(\frac{\sigma}{\kappa_3}\frac{\gamma}{\gamma_2}\sum_k n_k(T_{kj}w_k)^{-\gamma}(w_k f_{kj})^{-\gamma_2/(\sigma-1)}\bar{s}_{kj}^{\gamma_2 R}[1-\bar{s}_{kj}^{(1-\sigma)R}]\right)^{1/\gamma}.$$

We define

$$\Omega_j \equiv \kappa_2\left(\frac{\sigma}{\kappa_3}\frac{\gamma}{\gamma_2}\sum_k n'_k(T_{kj}w_k)^{-\gamma}(w_k f_{kj})^{-\gamma_2/(\sigma-1)}\bar{s}_{kj}^{\gamma_2 R}[1-\bar{s}_{kj}^{(1-\sigma)R}]\right)^{1/\gamma}$$

where $\kappa_2 = \left(\frac{\sigma}{\kappa_3}\frac{\gamma}{\gamma_2}\right)^{1/\gamma_2}$ and $n_i = z_L^{-\gamma} n'_i$ (follow the normalization as in Bernard, Moxnes, and Ulltveit-Moe (2018)). The solution for the sorting function is

$$\underline{z}_{ij}(Z) = \frac{T_{ij}w_i\Omega_j}{Z}(w_i f_{ij})^{1/(\sigma-1)}\frac{\zeta \bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R}.$$

The price index becomes

$$\begin{aligned}
\phi_j(Z)^{1-\sigma} &= Z^{\gamma_2}\bar{m}^{1-\sigma}\frac{\kappa_3}{\sigma w_i f_{ij}\zeta}\left[\frac{W_{ij}}{T_{ij}w_i\bar{s}_{ij}^{-R}}\right]^{\sigma-1} \\
&= Z^{\gamma_2}\bar{m}^{1-\sigma}\frac{\kappa_3}{\sigma w_i f_{ij}\zeta}\left[\frac{T_{ij}w_i(w_i f_{ij})^{1/(\sigma-1)}z_L}{T_{ij}w_i\bar{s}_{ij}^{-R}}\frac{\zeta \bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R}\Omega_j\right]^{\sigma-1} \qquad (3.A.3) \\
&= Z^{\gamma_2}\bar{m}^{1-\sigma}\frac{\kappa_3}{\sigma\zeta}\left[\frac{\zeta \bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R}\right]^{\sigma-1}\Omega_j^{\sigma-1}.
\end{aligned}$$

The revenue of final goods producers is

$$R_i = \left(\frac{P_i}{\Phi_i}\right)^{1-\sigma}\mu Y_i = \left(\frac{\bar{m}\phi_i(Z)}{Z\Phi_i}\right)^{1-\sigma}\mu Y_i.$$

where $P_i = \bar{m}q_i(Z)/Z$ is the price and Φ_i is the price index for final goods as the following

$$
\begin{aligned}
\Phi_i^{1-\sigma} &= N_i \int_1^{\infty} P_i(Z)dG(Z) \\
&= N_i \int_1^{\infty} (\bar{m}q_i(Z)/Z)^{1-\sigma} dG(Z) \\
&= Y_i \frac{\bar{m}^{2(1-\sigma)}\kappa_3}{\sigma\zeta} \frac{\Gamma}{\Gamma-\gamma} \left[\frac{\zeta \bar{s}_{ij}^{-\gamma R}}{(\sigma-1)R}\right]^{\sigma-1} \Omega_j^{\sigma-1}.
\end{aligned} \tag{3.A.4}
$$

Rewrite $R_i = \bar{m}E_i$ and substitute prices indices into the above revenue, we have $\kappa_3 = \mu(\Gamma-\gamma)/\gamma$ and

$$
E_i(Z) = \kappa_3 Z^{\gamma}. \tag{3.A.5}
$$

3.B Estimation Results for Different Sub-samples

This section shows detailed estimation results for equation (3.33) for different sub-samples. These coefficients are used to plot figure 3.4.

Table 3.B.1: Seller-Level OLS (2 or more buyers)

Hypothesis No.:	1	2	3	4
Dependent Variables:	$\ln x_{fig}$	$\ln S_{fig}$	$\ln \bar{q}_{fig}$	$\ln B_{fig}$
2 or more buyers				
$\ln(1 + Tariff)$	-1.440***	-0.6137***	-0.8264***	-0.1909***
	(0.1810)	(0.1261)	(0.1008)	(0.0611)
$\ln(1 + CIF/FOB)$	-1.617***	-3.002***	1.384***	-2.018***
	(0.4061)	(0.2897)	(0.2416)	(0.1567)
$\ln(ExportCost)$	-0.1996***	-0.2089***	0.0093	-0.2033***
	(0.0662)	(0.0419)	(0.0439)	(0.0238)
Observations	1,624,490	1,624,490	1,624,490	1,624,490
R^2	0.48891	0.29987	0.59134	0.27474
Within R^2	0.00010	0.00022	0.00010	0.00032
3 or more buyers				
$\ln(1 + Tariff)$	-1.505***	-0.6956***	-0.8091***	-0.2675***
	(0.2523)	(0.1783)	(0.1360)	(0.0939)
$\ln(1 + CIF/FOB)$	-3.613***	-5.024***	1.411***	-3.551***
	(0.6065)	(0.4654)	(0.3245)	(0.2721)
$\ln(ExportCost)$	-0.1253	-0.1645***	0.0391	-0.2315***
	(0.0951)	(0.0621)	(0.0616)	(0.0373)
Observations	871,458	871,458	871,458	871,458
R^2	0.42011	0.25095	0.52811	0.23924
Within R^2	0.00017	0.00044	9.81×10^{-5}	0.00065
4 or more buyers				
$\ln(1 + Tariff)$	-1.058***	-0.3983*	-0.6594***	-0.1330
	(0.3178)	(0.2292)	(0.1680)	(0.1253)
$\ln(1 + CIF/FOB)$	-6.209***	-7.136***	0.9269**	-5.171***
	(0.8433)	(0.6744)	(0.4151)	(0.4190)
$\ln(ExportCost)$	-0.0934	-0.0797	-0.0138	-0.2010***
	(0.1211)	(0.0805)	(0.0771)	(0.0496)
Observations	559,379	559,379	559,379	559,379
R^2	0.38819	0.23552	0.49132	0.23484
Within R^2	0.00027	0.00072	5.55×10^{-5}	0.00105

Clustered (Firm) standard-errors in parentheses
Fixed effects: Seller, Year, Product, Country
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3.B.2: Seller-Level OLS (5 or more buyers)

Hypothesis No.:	1	2	3	4
Dependent Variables:	$\ln x_{fig}$	$\ln S_{fig}$	$\ln \bar{q}_{fig}$	$\ln B_{fig}$
5 or more buyers				
$\ln(1+Tariff)$	-0.3690	0.2798	-0.6488***	0.2629*
	(0.3799)	(0.2762)	(0.1998)	(0.1594)
$\ln(1+CIF/FOB)$	-9.374***	-10.18***	0.8088	-7.338***
	(1.142)	(0.9268)	(0.5314)	(0.6011)
$\ln(ExportCost)$	-0.1310	-0.0944	-0.0366	-0.2148***
	(0.1482)	(0.0987)	(0.0948)	(0.0626)
Observations	396,764	396,764	396,764	396,764
R^2	0.36769	0.23120	0.46308	0.23961
Within R^2	0.00047	0.00123	4.9×10^{-5}	0.00172
6 or more buyers				
$\ln(1+Tariff)$	0.2130	0.6678**	-0.4548*	0.3911**
	(0.4423)	(0.3217)	(0.2327)	(0.1889)
$\ln(1+CIF/FOB)$	-11.40***	-12.50***	1.099*	-9.311***
	(1.409)	(1.122)	(0.6505)	(0.7392)
$\ln(ExportCost)$	-0.0997	-0.0503	-0.0495	-0.1895**
	(0.1769)	(0.1173)	(0.1128)	(0.0754)
Observations	301,717	301,717	301,717	301,717
R^2	0.35622	0.23122	0.44384	0.24675
Within R^2	0.00063	0.00172	3.81×10^{-5}	0.00246
7 or more buyers				
$\ln(1+Tariff)$	0.1738	0.6872*	-0.5133*	0.4759**
	(0.5068)	(0.3715)	(0.2664)	(0.2193)
$\ln(1+CIF/FOB)$	-12.81***	-14.21***	1.402*	-10.74***
	(1.683)	(1.365)	(0.7588)	(0.9169)
$\ln(ExportCost)$	-0.1186	-0.1246	0.0059	-0.2462***
	(0.2002)	(0.1354)	(0.1274)	(0.0890)
Observations	239,340	239,340	239,340	239,340
R^2	0.35019	0.23354	0.42971	0.25428
Within R^2	0.00077	0.00212	5.02×10^{-5}	0.00306

Clustered (Firm) standard-errors in parentheses
Fixed effects: Seller, Year, Product, Country
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3.B.3: Seller-Level OLS (8 or more buyers)

Hypothesis No.:	1	2	3	4
Dependent Variables:	$\ln x_{fig}$	$\ln S_{fig}$	$\ln \bar{q}_{fig}$	$\ln B_{fig}$
8 or more buyers				
$\ln(1+Tariff)$	0.8842*	1.254***	-0.3703	0.8195***
	(0.5257)	(0.3893)	(0.2902)	(0.2355)
$\ln(1+CIF/FOB)$	-14.39***	-16.02***	1.628*	-12.12***
	(1.985)	(1.627)	(0.8846)	(1.106)
$\ln(ExportCost)$	-0.1475	-0.1858	0.0383	-0.2877***
	(0.2219)	(0.1541)	(0.1357)	(0.1017)
Observations	194,355	194,355	194,355	194,355
R^2	0.34676	0.23709	0.41774	0.26092
Within R^2	0.00094	0.00259	4.69×10^{-5}	0.00367
9 or more buyers				
$\ln(1+Tariff)$	1.278**	1.442***	-0.1643	1.043***
	(0.5586)	(0.4105)	(0.3214)	(0.2520)
$\ln(1+CIF/FOB)$	-17.09***	-19.04***	1.953*	-14.37***
	(2.329)	(1.934)	(1.027)	(1.368)
$\ln(ExportCost)$	-0.3296	-0.1721	-0.1575	-0.3276***
	(0.2531)	(0.1779)	(0.1529)	(0.1180)
Observations	161,775	161,775	161,775	161,775
R^2	0.34324	0.24127	0.40678	0.26920
Within R^2	0.00121	0.00333	6.49×10^{-5}	0.00465
10 or more buyers				
$\ln(1+Tariff)$	1.346**	1.591***	-0.2452	1.165***
	(0.6027)	(0.4460)	(0.3449)	(0.2729)
$\ln(1+CIF/FOB)$	-18.18***	-20.52***	2.342**	-15.19***
	(2.590)	(2.185)	(1.132)	(1.565)
$\ln(ExportCost)$	-0.4193	-0.1764	-0.2429	-0.3359**
	(0.2785)	(0.1974)	(0.1686)	(0.1315)
Observations	138,963	138,963	138,963	138,963
R^2	0.33439	0.24383	0.38790	0.27579
Within R^2	0.00136	0.00383	0.00011	0.00512

Clustered (Firm) standard-errors in parentheses
Fixed effects: Seller, Year, Product, Country

3.C Data Source

Data used in this chapter includes the bill of lading data, CIF-FOB margin, and the World Bank's "Doing Business" report.

Bill of Lading Data

The data contain bills of lading for many countries, but I only have access to the US's bill of lading.[29] Each bill of lading includes information on exporters, importers, products, trade volume, shipping companies, and the ship's identification. However, there is no information about the value of the products.

CIF-FOB margin

This data was also used in chapter 2 for all country pairs. The difference is that products are grouped at the 2-digit level in this chapter.

World Bank's "Doing Business" report

The data[30] contains time costs, administrative costs to ship one container, etc. The administrative costs are used as fixed shipment costs per shipment. See Hornok and Koren (2015b) for the description of the data.

Tariff

Data on tariff is downloaded from World Integrated Trade Solution.[31] The original data has tariff information at each tariff line. There are different types of tariffs but the selection of tariff rates is as the following. First, preferential rates are used whenever available. If it is not the case, then the Most Favoured Nation Rate (MFN) is used. When both are not available, the Bound Tariff Rate (the

[29] https://www.datamyne.com/countries-covered-global-trade-data
[30] https://archive.doingbusiness.org/en/doingbusiness
[31] https://wits.worldbank.org

maximum rate a country commits upon joining the World Trade Organization) is used.

3.D B/L Data Cleaning Process

This section shows steps in cleaning the B/L dataset. The raw data contain more than 54 million records, which amount to about 4 million tons of cargo imported by the US between 2009 and 2015. Data containing no information about products, buyers, or sellers are removed (steps 2, 3, 4). Using the list of buyers and sellers compiled by the same data provider, those records whose senders and receivers are not companies or logistics companies are removed (steps 5, 6, 7). After removing the above, the data has about 25 million records and 1.6 million tons of cargo. This data is merged with distance, gravity controls, and administrative costs (steps 9, 10, 11, 12, 13). Records with multiple products in one shipment (about 2 million records with 40 thousand tons of cargo) are removed (step 14). Records that belong to countries without complete data and duplicated records are removed (steps 15, 16). Some CIF-FOB margins are not available for all commodities; they are filled with the average of all commodities' CIF-FOB margins for the same country pairs. The remaining NAs are removed (steps 17 and 18). Records with zero tons are removed (19). Non-commercial products (HS2=98) are removed (step 20). Some outliers are removed (step 21).

Table 3.D.1: Note on Cleaning Data

	Note	N	Tons
1	Raw	54,741,936	4,308,229
2	Company type in buyer type	12,069,856	1,691,755
3	Empty/not declared container product	543,549	16,719
4	Buyer/supplier not available	15,466,435	908,995
5	Non-company in buyer name	100,577	669
6	Non-company in supplier name	23,518	2,214
7	Logistics company	966,688	20,978
8	Removed the above	25,667,413	1,671,417
9	Sea distance not available	25,529,144	1,651,280
10	Bunker price	25,529,144	1,651,280
12	Gravity data	25,497,611	1,639,391
13	Administrative export cost	25,419,034	1,577,559
14	Single product	23,581,925	1,532,370
15	Country with full data	23,578,176	1,528,925
16	Remove duplicated	23,578,176	1,528,925
17	Transport (CIF-FOB) cost not available	4,812,405	213,308
18	Removed not available transport cost after filled with average	23,546,726	1,493,678
19	Removed zero tons after filled with teu ratio	23,536,240	1,493,718
20	Removed HS2=98	28,078,738	1,749,385
21	Removed outlier supplier	27,697,464	1,741,978
22	Summarized by buyer-seller-year	5,779,195	1,741,978

Chapter 4

Trade Costs and Multinational Firms' Location Decisions

4.1 Introduction

Recent decades have witnessed a great change in the flow of trade and investments thanks to the reduction in trade barriers. Multinational corporations have become the main players in global integration by expanding their operations in multiple countries. Their complex organization has shaped the flows of goods and tasks both within and between the boundary of multinational firms. While many studies have focused on manufacturing plants and headquarters, few have studied the location of wholesale. This study investigates the decision on the location of manufacturing plants and wholesale subsidiaries of Japanese manufacturing firms in the EU. Even though the Single Market has dramatically reduced trade and investment barriers, the existence of regional clusters indicates that frictions are still large. Studies of these frictions are not only interesting academically but also important for political implications.

There is a vast literature on the relationship between trade costs and FDI.[1]

[1] See Alfaro and Chen (2018) for a thorough review.

Due to the complexity of multinational firms, the trade literature has divided the FDI into different motives to examine the effects of trade costs on FDI. Because horizontal FDI is a substitution strategy for export, trade costs tend to increase horizontal FDI (Markusen, 1984). Other types of FDI (vertical, complex FDI, export platform) exploit the comparative advantage of the host country in labor input and tend to complement exports. Higher trade costs result in a decrease in these types of FDI (Helpman, 1984; Yeaple, 2003; Ekholm, Forslid, and Markusen, 2007). In general, this literature studies the expansion in the boundary of firms when domestic firms switch to multinational firms.

While traditionally, the agglomeration economy treats the boundary of firms as an exogenous variable, there is a strand of literature integrating the agglomeration economy in the location decision of FDI (Head, Ries, and Swenson, 1995; Strauss-Kahn and Vives, 2009). Transportation costs play an important role in the agglomeration economy because proximity induces market externalities. Agglomeration and congestion can motivate firms to separate different functions in specialized regions (Duranton and Puga, 2005; Defever, 2006; Defever, 2012). Evidence has shown that headquarters are most agglomerative among different types of subsidiaries (Alfaro and Chen, 2018). However, little has been shown about the relative location between wholesale and manufacturing.[2]

Recent studies also highlight the spatial interdependence among FDI flow (Blonigen, Davies, et al., 2007). Head and Mayer (2004) stresses the importance of market access for Japanese FDI firms investing in the EU. Chen (2011) shows that supplier access is crucial for French firms investing abroad. These studies highlight the role of third-country effects. Not only the trade costs between home and host countries but the costs among suppliers and markets seem to greatly affect the profitability of multinational firms. These studies often use distance

[2]Defever (2012) has shown that the location decision of the sales unit does not depend on the location of manufacturing plants of the same firm. However, they did not explain why that is the case.

as proxied for transportation costs.

The distinct between FDI and trade is the separability of knowledge capital (Helpman, 1984). A recent study by Keller and Yeaple (2013) points out that the flow of knowledge also faces some frictions and seems to follow the law of gravity. The cost of communication tends to decrease with distance (Head and Ries, 2008), and direct communication can substitute for the transfer of technology through intermediates trade (Keller and Yeaple, 2013). While communication costs can be reduced by the improvement in information infrastructure (Alfaro and Chen, 2018), they are also related to the interaction of global services firms.

The business literature has shown that a well-connected city can greatly reduce communication costs for firms located there (Taylor, 2001). Belderbos, Du, and Goerzen (2017) and Goerzen, Asmussen, and Nielsen (2013) shows that headquarters are likely to locate in a well-connected global city.[3] Given the focus of this paper is Japanese firms, financial groups play an important role in co-location decisions for information sharing (Head, Ries, and Swenson, 1995; Blonigen, Ellis, and Fausten, 2005). Exploiting this fact, this study constructs a network index to measure the connectivity of a region in the EU in the Japanese firms' network. The information-sharing mechanism is shown to be different between manufacturing plants and wholesale subsidiaries.[4]

While the attractiveness of well-connected regions for FDI has been explored in Belderbos, Du, and Goerzen (2017) for the decision of headquarters, this has not been studied for other types of subsidiaries. While manufacturing plants may not need the level of connectedness as headquarter, wholesale may need these services because they often require transactions with customers.

[3]In the context of this study, connectivity means the interlocking model as in (Taylor, 2001), where firms connect cities through their activities. Two cities are connected if they have subsidiaries of the same firm, counted as one link. A city with more active firms is more likely to have more links, hence more likely to be well-connected.

[4]Belderbos, Du, and Goerzen (2017) use a composite measure of connectivity based on producer services, airport passenger flows, and the intensity of international co-invention for city connectivity index.

The results show that wholesale affiliates tend to spread to a region with a strong presence in other Japanese service firms. They are more likely to locate in more connected regions to enjoy the externalities of Japanese services firms as well as the reduction in trade costs among regions. This means that "information sharing" among multinational firms' subsidiaries does not need to happen in close proximity.

The study also confirms the agglomeration force for manufacturing plants within the industry and financial groups. The industry agglomeration effects are less than half of the within-financial-group agglomeration effects. While the manufacturing subsidiaries tend to be located in regions with a strong presence of financial groups, they need not cluster in a well-connected city. However, the findings are only robust for the agglomeration index for different regional levels. The network index becomes significant at the country level and states level.

The relative location between manufacturing plants and wholesale subsidiaries is shown in the analysis. It turns out that manufacturing plants and wholesale subsidiaries of the same industry and the same firms are not located near each other. On the other hand, manufacturing firms locate their plants in regions with a strong presence in their financial group's wholesale subsidiaries, which stresses the importance of sharing information beyond industry agglomeration.

The remainder of this paper includes an analytical framework, data and descriptive statistics, empirical results, discussion and conclusion.

4.2 Analytical Framework

4.2.1 General Concept

In general, the potential profit firm f investing in the region r in sector k has the following form

$$\pi^{s}_{fr} = \mathbf{Y}_r\alpha + \mathbf{X}_{fr}\beta + \mathbf{Z}_f\gamma. \tag{4.1}$$

$\mathbf{Y_r}$ contains variables that are specific to region r. From the trade literature, they may include market size, market potential, and production cost (Markusen, 1984; Head and Mayer, 2004; Helpman, 1984). Market size is related to the motivation of horizontal and export-platform FDI, while production costs are considered in vertical FDI. Recent studies highlight the importance of information technology infrastructure in servicing multinational firms' inside communication (Alfaro and Chen, 2018). In this paper, time-invariant regional characteristics are controlled by regional dummies. Time-variant characteristics are captured through population density (inhabitants per square kilometers) which reflects the market size, human capital (people with tertiary education and/or employed in science and technology as a percentage of the total population), and average value added for employees in industries (one thousand Euros per person).

$\mathbf{X_{fr}}$ contains variables specific to firms f in region r. In the gravity context, they are bilateral trade barriers between the host and the home country. Since the focus is FDI between Japan and the EU, gravity-like controls such as common language or distance are not included.

$\mathbf{X_{fr}}$ also contains variables relating the characteristics of firms f with region r. The benefits from agglomeration may increase the firms' profits. Regarding agglomeration economies, region-specific variables may include agglomeration through production linkage or financial group connection (Head and Mayer, 2004; Alfaro and Chen, 2014). The source of agglomeration is the proximity between suppliers and final good producers (vertical linkages) or the externalities in the factor market. In this study, the industry agglomeration is captured by the number of subsidiaries in the same industry of the investing firm. Because the paper uses Japanese data, which tends to cluster within *keiretsu* to share information (Head and Mayer, 2004; Blonigen, Ellis, and Fausten, 2005), an ownership-ratio-weighted number of subsidiaries in the same *keiretsu* is used. These two variables are also used in the literature (Defever, 2012; Mayer, Mejean, and Nefussi, 2010).

In addition, this paper argues that "information sharing" among multinational firms' subsidiaries does not need to happen in close proximity. A firm needs all types of services, such as transportation, marketing, accounting, or legal aid to operate smoothly. These services are particularly less easy to access for foreign than domestic firms. Therefore, it is more likely for Japanese firms to be situated in a region well connected to other Japanese services firms. This variable is temporarily called the "Network Index" to distinguish it from the agglomeration index described above. It is not just the number of services firms but the region's connectedness, thanks to the operation of these firms, that make the region easier to access from other regions. While more clustered services can create agglomeration-like effects for firms in certain regions, the connectedness potentially reduces trade costs between regions.[5] Distant regions (with high transportation costs) may become more accessible.

Furthermore, congestion costs due to agglomeration may induce firms to separate tasks in different cities (Duranton and Puga, 2005). This suggests that different types of subsidiaries may respond differently to the agglomeration index. For example, manufacturing subsidiaries may cluster with other manufacturing but not with other service subsidiaries (Defever, 2012). In this study, manufacturing and wholesale subsidiaries are considered. Both types are affected by agglomeration forces, but wholesale is more likely to disperse in space to reach more customers. However, to compensate for the lack of agglomeration benefits from proximity with their groups, connectivity with other Japanese services firms is likely to be an important factor.

$\mathbf{Z_f}$ are firm-specific variables. They may include firm productivity or firm size (Helpman, Melitz, and Yeaple, 2004; Yeaple, 2003). Firms with higher

[5] This also includes communication costs because it contains friction in moving knowledge within and between firms. Both costs, however, are likely to follow gravity law. The longer the distance, the higher the costs (Alfaro and Chen, 2018; Keller and Yeaple, 2013).

productivity can pay more fixed costs to set up a complex FDI, including export platform FDI (Ekholm, Forslid, and Markusen, 2007). Bigger firms tend to invest in more regions and more in one region (Yeaple, 2009). Firms' characteristics are controlled by firms' fixed effects in this study.

4.2.2 Constructed Variables

This section provides details of the variables discussed in the previous section. Table 4.1 summarizes the list of variables used in the analysis.

Table 4.1: List of Variables

Variable	Definition	Sign	Cross-Sector
$\ln(1 + IndustryCount)_{ir}$	Number of affiliates of same industry of investor	+	?
~ [Man/Who]	~ for each sector (Manufacturing or Wholesale)	+	?
$\ln(W.GroupCount)_{fr}$	Ownership-weighted count of affiliates of same group	+	?
~ [Man/Who]	~ for each sector (Manufacturing or Wholesale)	+	?
$\ln(Affi.Network)_{ir}$	Number of non-manufacturing affiliates of big manufacturing firms	+	
$\ln(FirmsNetwork)_{fr}$	Number of affiliates of non-manufacturing group members	+	
$\ln(HumanCap.)$	Population percentage of people with tertiary education (%)	+/-	
$\ln(VA/emp)$	Average value added for industry employees (1000Eur/person)	+/-	

Details of these variables are as the following.

The first two variables represent the industry agglomeration index. The first one is the number of affiliates in the same industry. Those affiliates that belong to the same group are excluded to avoid correlation with the group agglomeration index. In the baseline specification, affiliates of all sectors are counted. To consider the effect of prior affiliates of different sectors on the location choice of a specific sector, a similar index is calculated for each affiliate sector (the cross-sector effect).

The third variable is the group agglomeration index. It is calculated as

$$W.GroupCount^s_{fgr(t-1)} = \sum_{k \in G_f, k \neq f} \sum_{a_{kr} \in A_{kr}} share_{kg} share_{ka} D^s_{a_{kr}(t-1)} \tag{4.2}$$

where g denotes the firm's group parent company of firm f, and k denotes a related company of f is the same group g. G_f is the set of firms belong to group g, and A_{kr} is the set of affiliates of firm k in region r. $D^s_{a(t-1)}$ is a dummy that equals 1 if firm k has an affiliate a of sector s in region r at time $t-1$. $share_{kg}$ is the ownership ratio of group g over firm k, and $share_{ka}$ is the ownership ratio of firm k over affiliate a. The product of these two is the ownership ratio of group g in affliate a. Weighting the ownership ratio instead of just counting the number of affiliates of group g considers the inter-relationship in terms of "ownership control" among different affiliates a that belong to group g.

Similar to within-industry indices, $W.GroupCount^s_{gr(t-1)}$ for all sectors are used in the baseline specification and for each sector in the cross-sector specification. Affiliates that belong to firm f are excluded.

The fifth and sixth variable represent the network indices. The first one is calculated as the log count of non-manufacturing affiliates of manufacturing firms that have at least two affiliates in more than two regions.[6] A higher index value indicates more services from Japanese firms investing in the region. Specifically, the non-manufacturing affiliates network index for industry i in region r is

$$AffiNet_{ir(t-1)} = \sum_{k \in I_{ir}} \sum_{a_{ikr} \in A^{NonManuf}_{kr}} D^{NonManuf}_{a_{ikr}(t-1)} \tag{4.3}$$

[6]This borrows from the network literature. The FDI link between firms and cities can be expressed in a two-mode network so the network index in this study is the log of in-degree of cities, which is one simple measure of centrality in network analysis (for example, Borgatti and Everett (1997)). In the context of global city as in Taylor (2001), if the firm's value is 1, the network index here is the city nodal connection.

where $D^s_{a_{ik}(t-1)}$ is a dummy that equals 1 if firm k of industry i has a non-manufacturing affiliate a of sector s in region r at time $t-1$, and A^{NM}_{ik} is a set of of such affiliates. I_i is the set of firms belong to industry i that has affiliates in region r.

The second variable is the log of ownership-weighted-count of affiliates from non-manufacturing members from the same financial group of the firm.[7] These firms are not included in the first components to avoid a correlation between the two. Similarly, a higher index value indicates a stronger presence of the firm's group in the region. Specifically, the non-manufacturing firm's network index for industry i in region r is

$$FirmNet^{NonManuf}_{fr(t-1)} = \sum_{k \in G^{NM}_f} \sum_{a \in A_{kr}} share_{a_k} D^{NonManuf}_{a_{kr}(t-1)} \tag{4.4}$$

where G^{NM}_f is the set of non-manufacturing member firms in group company of firm f.

4.2.3 Estimation Equations

Assuming that firm f chooses to open affiliates of sector s in the region r at time t if $\pi^s_{frt} > \pi^s_{fr't}$ for $r, r' \in R$ where π^s_{frt} is the expected profit function of the firm as in equation 4.1 but with time index t. If error terms follow type I extreme distribution, the choice probability for $y_{rft} = 1$ can be calculated as in equation (4.5).

$$P_{frt}(y_{frt} = 1) = \int \frac{\exp(\mathbf{Y}_r \alpha_{\mathbf{i}} + \mathbf{X}_{fr}\beta + \mathbf{Z}_f \gamma)}{\sum_{fqt} \exp(\mathbf{Y}_r \alpha_{\mathbf{i}} + \mathbf{X}_{fr}\beta + \mathbf{Z}_f \gamma)} f(\alpha|\mu, \sigma) d\beta. \tag{4.5}$$

[7] This is different from the Group Agglomeration index because these are non-manufacturing firms while the Group Agglomeration index counts affiliates of manufacturing firms. The number of these firms and their affiliates are shown in table 4.A.2 in the Appendix.

Assuming that profits coming from different sectors do not correlate with each other and profits are independent among periods, the likelihood function takes into account all periods for each firm f as in equation (4.6). Notice that the coefficients α can be modelled as random variables.

$$S_f = \int \prod_{t=1}^{T_f} \prod_{r=1}^{R} \left[\frac{\exp(\mathbf{Y}_r \alpha_{\mathbf{i}} + \mathbf{X}_{fr}\beta + \mathbf{Z}_f \gamma)}{\sum_{fqt} \exp(\mathbf{Y}_r \alpha_{\mathbf{i}} + \mathbf{X}_{fr}\beta + \mathbf{Z}_f \gamma)} \right]^{y_{frt}} f(\beta|\mu, \sigma) d\beta. \tag{4.6}$$

This can be estimated using a simulation approach, assuming $f(\alpha|\mu, \sigma)$ follows normal distribution (Train, 2002) .

4.3 Data and descriptive statistics

Data is from the Overseas Japanese Companies Data and Domestic Affiliates Database from the Toyo Keizai Database to collect information about Japanese affiliates abroad and groups of Japanese firms. Only affiliates with at least 10 percent ownership from Japanese firm(s) are considered direct investments in a foreign country. The entry year of affiliates is the established year of the affiliate.

Information about affiliates' spatial locations is retrieved using postal codes from affiliates' addresses to calculate their coordinates. In the case where postal codes are unavailable, cities in addresses are used instead.[8] The raw data has 4,890 firms and 26,748 affiliates. The final sample has 527 firms and 1,233 foreign affiliates, excluding those with less than 10 percent ownership by Japanese firms, those outside the EU, and those whose parent companies are not manufacturing firms. These firms are linked with the Domestic Affiliates Database to create information about groups of firms. There are 49 groups that cover 147 firms (see Table 4.A.1). The Japan Company Handbook is used to get informa-

[8]The geonames.org database is used to search for coordinates from postal codes and cities. Results are manually checked to make sure they are correct.

tion about each firm's capital, year of establishment, and employees. The data is from 1991 to 2015. It is an unbalanced panel data.

The regression sample only includes observations of which investments were made after 1990. However, all possible investments before 1990 to calculate the initial state of each industry, group, and firm agglomeration. Figure 4.1 shows affiliates' locations from firms with more than one affiliate in the EU before and after 1990. Manufacturing and Headquarters seem to be concentrated in certain regions, such as the UK, Belgium, Netherlands, and some western countries, while wholesale affiliates are more scattered over many regions.

The network firms index is constructed using Japanese firms in four sectors: Finance and Insurance, Service Activities, Transportation and Communications, Wholesale and Retail Trade, using only firms that have more than two or three affiliates in the EU.[9] Specifically, there are 45 finance and insurance firms, 101 services firms, 53 transportation and communication firms, and 119 wholesale and retail firms.

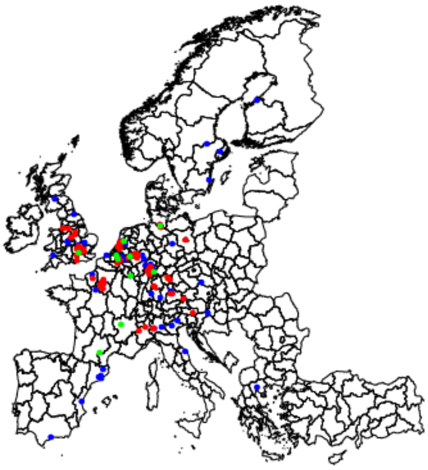
Industry classification follows Toyo Keizai Database's own classification. These industries can be grouped into SNA sectors.[10]

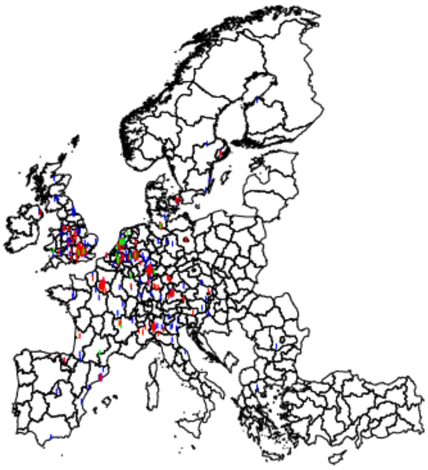
Regional characteristics and region maps are from Eurostat. These maps and firms' coordinates are used to identify the regions where firms invested. Other data include population density, land, GDP, value added per employee, and human capital (people with tertiary education and/or employed in science and technology as a percentage of the total population).

The choice of a location may differ based on region size. Alfaro and Chen (2014) raises the issue that spatial boundaries may affect the agglomeration levels in that these boundaries are not continuous like distance. Blonigen, Davies, et al.

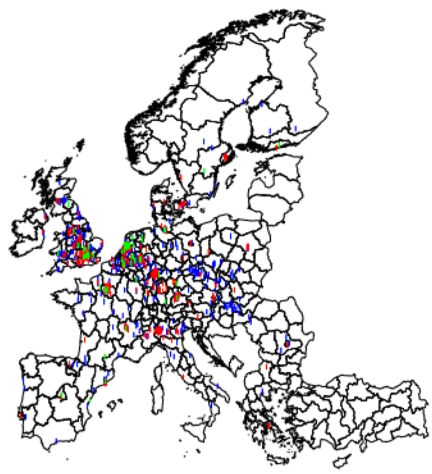
[9]See Table 4.A.2 in Appendix 4.A. The criteria to choose these firms is the top 25 percentile of the distribution of the number of affiliates of each firm within each sector. Only firms that can create connections among cities are included.

[10]Table 4.A.3 in Appendix 4.A shows a list of these industries and their corresponding SNA sectors.





(a) 1980 (b) 1990



(c) 2000 (d) 2010

Figure 4.1: Accumulated Japanese Affiliates in EU in different years

Manufacture(Blue), Wholesale(Red), Headquarters(Green)

(2007) also suggests aggregation of country-level FDI may affect their results on the US's FDI to European countries. Therefore, the robustness check is done for different administrative levels. Because detailed coordinates of affiliates make it possible to identify the location at different administrative levels, including country, group of states, and states. The baseline is regional level 1 (groups of states).

4.4 Empirical Results

4.4.1 Baseline Specifications

Table 4.2 shows the estimation results of the baselined specification for equation (4.5) for the location decision of manufacturing and wholesale affiliates.[11]

Table 4.2: Location Choice at Region Level

	Manufacturing	Wholesale
ln(1+ Industry Count)	2.120***	2.538***
	(0.151)	(0.153)
ln(1+ Group Count)	4.135***	6.905***
	(1.155)	(1.440)
Non-manuf Affiliates Network	-0.032	-0.113***
	(0.041)	(0.031)
Non-manuf Firms Network	0.288	0.651***
	(0.211)	(0.150)
ln(Human Cap.)	-2.024	1.687
	(1.174)	(1.370)
ln(VA/emp)	0.185	0.075
	(0.547)	(0.844)
Observations	21688	24330

Clustered (Firm) Standard errors in parentheses
All regressions include firm fixed effects and region fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first column shows results for Manufacturing affiliates. The region-specific variables, including value-added per employee and human capital, do

[11] The mixed logit with random parameters was not properly converged for some cases so they are omitted.

not show significant results.[12] The agglomeration index within industry and within financial groups are both significantly positive. The magnitude of the latter is twice as high as the former. This shows that manufacturing subsidiaries tend to be located in regions with a strong presence of financial groups. On the contrary, the network indices do not show any significance. This confirms the finding in the literature that Japanese manufacturing subsidiaries are clustered in regions with strong industry linkages and financial groups.

The second column shows results for Wholesale affiliates. The region-specific variables, including value-added per employee and human capital, do not show significant results. Similar to Manufacturing affiliates, the agglomeration index within industry and within financial groups are both significantly positive, and the latter is twice as high as the former. This shows that wholesale subsidiaries also tend to be located in regions with a strong presence of financial groups. Unlike manufacturing affiliates, the affiliates network index is significantly negative, while the firm network index shows a positive sign. This is mostly because firms find it beneficial to spread their wholesale networks to reach more customers. This has also been shown in Defever (2012).

The new insight is that these wholesale affiliates tend to spread to a region with a strong presence in other Japanese service firms. This highlight the mechanism described in the previous section. Japanese firms situate their wholesale in a region well connected to other Japanese services firms to utilize their services. It is not just the number of services firms but the region's connectedness, thanks to the operation of these firms, that makes the region easier to access from other regions. While more clustered services can create agglomeration-like effects for firms in certain regions, the connectedness potentially reduces trade costs between regions. The wholesale affiliates are more likely to locate in more connected regions to enjoy the externalities of Japanese services firms as well as

[12] Density variable is highly correlated with other variables, so it is omitted.

the reduction in trade costs among regions.

4.4.2 Cross-sector effects

Table 4.3 shows the estimation results for equation (4.5) where the agglomeration indices are calculated for wholesale and manufacturing affiliates. This exercise shows how manufacturing and wholesale affiliates cluster differently.

Table 4.3: Firm Location Decision: Cross-Sector Effects

	Manufacturing	Wholesale
ln(1+ Industry Count)[Man]	3.468***	1.809
	(0.230)	(1.030)
ln(1+ Industry Count)[Who]	-23.283***	2.741***
	(1.181)	(0.178)
ln(1+ Group Count)[Man]	-2.094	-29.585***
	(5.074)	(3.284)
ln(1+ Group Count)[Who]	36.437***	10.089***
	(4.756)	(1.894)
Non-manuf Affiliates Network	-0.031	-0.061
	(0.046)	(0.037)
Non-manuf Firms Network	0.233	0.610***
	(0.210)	(0.171)
ln(Human Cap.)	-3.408**	2.434
	(1.200)	(1.405)
ln(VA/emp)	0.141	0.237
	(0.595)	(0.952)
Observations	21828	24330

Clustered (Firm) Standard errors in parentheses
All regressions include firm fixed effects and region fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first columns show the results for Manufacturing. While the industry count of manufacturing is significantly positive, the count of wholesale affiliates is significantly negative. The fact that manufacturing plants cluster with each other for exploit industrial spilled over effects has been well-established in the literature. However, it is rather new that manufacturing firms disperse the location of manufacturing plants and wholesale subsidiaries.

Notice that the count of wholesale subsidiaries in the same group is significantly positive. Still, the count of manufacturing plants in the group is not significant.[13] It means that Japanese firms locate their wholesale subsidiaries with other groups' wholesale subsidiaries to utilize the group's knowledge about customers, thanks to the lower cost of information sharing and private information.[14]

The first columns show the results for Wholesale. While the industry count of manufacturing is insignificant, the count of wholesale affiliates is significantly positive. Similar to manufacturing plants, wholesale subsidiaries tend to cluster with each other for information sharing. Furthermore, the count of wholesale subsidiaries in the same group is significantly negative, but the count of manufacturing plants in the group is significantly positive. Information sharing among wholesale subsidiaries is also present within financial groups.

The above results confirm that the same type of subsidiaries in the same industry of Japanese firms cluster with each other for agglomeration effects. Furthermore, group agglomeration is more likely to happen in wholesale functions, which stresses the importance of sharing information beyond industry agglomeration. The importance of information sharing among wholesale subsidiaries is further confirmed because wholesale firms are more likely to locate in a well-connected region.

4.4.3 Robustness Check

Results at different regional levels

Equation (4.5) are estimated for different regional levels: country, group of states (level 1), and states (level 3). This exercise addresses the concern that the

[13]It is concerning that the magnitudes of the wholesale count are much larger than other coefficients. This may be due to some collinearity among variables. However, Table 4.A.4 in the Appendix shows that the correlation among variables is not large. The two network indices have a high correlation. However, an estimation dropping either network indices shows the same results.

[14]This has been explored theoretically in Blonigen, Ellis, and Fausten (2005), and examined for manufacturing plants in Head and Mayer (2004).

boundary of agglomeration is not always clear and choosing the wrong level of administrative regions may alter the results(Alfaro and Chen, 2014; Blonigen, Davies, et al., 2007; Defever, 2012).[15]

Table 4.4: Mixed Logit at Different Regional Levels for Manufacturing

	Country	Level 1	Level 2
ln(1+ Industry Count)	1.510***	2.120***	2.666***
	(0.123)	(0.151)	(0.227)
ln(1+ Group Count)	3.241***	4.135***	5.387***
	(0.667)	(1.155)	(1.162)
Non-manuf Affiliates Network	-0.066*	-0.032	-0.138**
	(0.029)	(0.041)	(0.053)
Non-manuf Firms Network	0.409*	0.288	0.961***
	(0.200)	(0.211)	(0.182)
ln(Human Cap.)	-0.655	-2.024	
	(1.431)	(1.174)	
ln(VA/emp)	0.456	0.185	
	(0.657)	(0.547)	
Observations	6654	21688	48660

Clustered (Firm) Standard errors in parentheses
All regressions include firm fixed effects and region fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.4 shows the results for Manufacturing. Both industry and group agglomeration indices are significantly positive for all regional levels. The ratio between the two coefficients is also roughly the same, and group agglomeration index is twice as high as the industry agglomeration index. The affiliates' network indices become significantly negative, while the firm's network indices become significantly positive for country and state levels. For both cases, the latter is nine times higher than the former. This shows two things. The difference between these results and the baseline results indicates that the "network index" is sensitive to the choice of the regional unit.

Table 4.5 shows the results for Wholesale. Both industry and group agglomeration indices are significantly positive for all regional levels. The ratio between

[15]There are a few variations for Human Captial and Value Added per Employee at regional level 3 so they are omitted from the model.

Table 4.5: Location Choice at Different Regional Levels for Wholesale

	Country	Level 1	Level 2
ln(1+ Industry Count)	2.170***	2.538***	2.546***
	(0.187)	(0.153)	(0.156)
ln(1+ Group Count)	5.213*	6.905***	10.262***
	(2.269)	(1.440)	(2.978)
Non-manuf Affiliates Network	-0.101***	-0.113***	-0.107*
	(0.024)	(0.031)	(0.045)
Non-manuf Firms Network	0.418*	0.651***	0.722***
	(0.166)	(0.150)	(0.197)
ln(Human Cap.)	3.002	1.687	
	(1.748)	(1.370)	
ln(VA/emp)	-0.157	0.075	
	(1.052)	(0.844)	
Observations	7582	24330	54258

Clustered (Firm) Standard errors in parentheses
All regressions include firm fixed effects and region fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the two coefficients increases from roughly two at the country level to about five at the state level. This means there is a lot of fluctuation in the wholesale agglomeration index among regions when the regional boundary becomes smaller. One explanation is that wholesale subsidiaries are so spread out that there may be more wholesale clusters in one region. In contrast, there seem to be fewer manufacturing plant clusters in one region.

Similarly, the "network index" also retains the significant sign as in the baseline specification. However, the ratio between Firms Network Index and Affiliates Network Index increased from about four at the country level to seven at the regional level 2. The latter may have fewer clusters in one region than the latter.

4.4.4 Continuous Regional Boundary

This section takes one step further to relax the boundary of regions by defining regions as the area inside a circle of different diameters from the regional center. The regional center is in the geographical centroids (middle point) of the most populated cities of that region to calculate the vicinity. In the baseline case, the

Table 4.6: Distances among Regional Centroids and Affiliates

Distance	N	mean	sd	min	max
1. Economic Center to Economic Center	320	42.12	42.85	0.00	368.50
2. Affiliates to Economic Center	1261	28.69	30.26	*0.41*	*245.30*
3. Affiliates to Geographical Center	1261	31.60	29.38	0.35	474.26

Note: In 2 and 3, only chosen regions are counted. Economic Center indicates the centroid of the densest city. The geographical Center indicates geographical centroids. Data is calculated using Eurostat.

diameter is about $d = 76km$. To put this in perspective, the minimum distance of an affiliate to a region's centroid is $0.41km$, and the furthest is $244km$ (as shown in Table 4.6).

Equation (4.5) is estimated by adding the dummy showing whether an affiliate of sector k is in the region or not, where the region boundary is changed as above. The objective of this exercise is to check how close/far the cluster for both within and across sectors.

The results are in Tables 4.7 and 4.8. Table 4.7 shows the results for manufacturing affiliates. The results for the industry index and group index are kept at baseline. The parameters of interest are the dummies indicating the existence of previous manufacturing plants or wholesale subsidiaries in the region. The effect of an existing manufacturing plant on setting up a manufacturing subsidiary has two peaks at $51km$ and $126km$. This confirms the conjecture that manufacturing plants have few clusters in each region. The effects of existing wholesale on opening a new wholesale subsidiary are significantly negative for almost all distances, which confirms the intuition that locations of wholesale spread out.

Table 4.7: Cross-sector Effect of Prior Presence in the Vicinity for Manufacturing Location

	26km	51km	76km	101km	126km	151km	176km
ln(1+ Industry Count)	2.637***	2.636***	2.631***	2.632***	2.633***	2.630***	2.627***
	(0.223)	(0.221)	(0.223)	(0.223)	(0.222)	(0.223)	(0.224)
ln(1+ Group Count)	5.144***	5.182***	5.161***	5.172***	5.249***	5.221***	5.185***
	(1.120)	(1.122)	(1.124)	(1.128)	(1.125)	(1.122)	(1.124)
Previous Manufacturing	1.068	1.131*	0.648	0.728	1.061**	0.792*	0.375
	(0.734)	(0.497)	(0.521)	(0.426)	(0.361)	(0.365)	(0.365)
Previous Wholesale	3.472	2.635	2.063	1.751	1.506	0.803	0.570
	(1.830)	(1.607)	(1.195)	(1.238)	(1.298)	(1.398)	(1.246)
Observations	48660	48660	48660	48660	48660	48660	48660

Clustered (Firm) Standard errors in parentheses

Man: Manufacturing, Who: Wholesale.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.8: Cross-sector Effect of Prior Presence in the Vicinity for Wholesale Location

	26km	51km	76km	101km	126km	151km	176km
rid							
ln(1+ Industry Count)	2.489***	2.494***	2.494***	2.494***	2.500***	2.506***	2.504***
	(0.152)	(0.153)	(0.153)	(0.153)	(0.155)	(0.155)	(0.156)
ln(1+ Group Count)	10.798**	10.778**	10.754**	10.747**	10.744**	10.707**	10.643**
	(3.451)	(3.465)	(3.443)	(3.443)	(3.450)	(3.462)	(3.459)
Previous Manufacturing	-21.398***	-20.070***	1.866	1.599	1.493	1.128	0.957
	(1.005)	(0.941)	(2.733)	(2.628)	(2.706)	(2.464)	(2.376)
Previous Wholesale	-0.432	-0.880*	-0.876*	-0.988**	-1.094**	-1.225***	-1.104***
	(0.479)	(0.407)	(0.371)	(0.363)	(0.361)	(0.342)	(0.305)
Observations	54258	54258	54258	54258	54258	54258	54258

Clustered (Firm) Standard errors in parentheses

Man: Manufacturing, Who: Wholesale.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.5 Conclusion

This study examined the decision of manufacturing and wholesale subsidiaries' location of Japanese manufacturing firms in the EU. The paper highlight the geographical patterns difference between wholesale subsidiaries and manufacturing plants. The new insight is that these wholesale affiliates tend to spread to a region with a strong presence in other Japanese service firms. The wholesale affiliates are more likely to locate in more connected regions to enjoy the externalities of Japanese services firms and the reduction in trade costs among regions.

The study also confirms the agglomeration force for manufacturing plants within the industry and financial groups. The industry agglomeration effects are less than half of the within-financial-group agglomeration effects. While the manufacturing subsidiaries tend to be located in regions with a strong presence of financial groups, they need not cluster in a well-connected city. However, the findings are only robust for different regional-level agglomeration indexes. The network index becomes significant at the country level and states level.

The relative location between manufacturing plants and wholesale subsidiaries is shown in the analysis. It turns out that manufacturing plants and wholesale subsidiaries of the same industry and the same firms are not located near each other. On the other hand, manufacturing firms locate their plants in regions with a strong presence in their financial group's wholesale subsidiaries, which stresses the importance of sharing information beyond industry agglomeration.

This study shows that "information sharing" among multinational firms' subsidiaries does not need to happen in close proximity. A firm needs all types of services to operate smoothly, which are particularly less easy to access for foreign than domestic firms. Japanese firms situate their wholesale in a region well connected to other Japanese services firms to utilize their services. It is not just the number of services firms but the region's connectedness, thanks to the operation of these firms, that makes the region easier to access from other regions. While more clustered services can create agglomeration-like effects for firms in certain regions, the connectedness potentially reduces trade costs between regions.

To improve the study, more regions can be added to examine the regional characteristics. Furthermore, data on intra-firm trade can show more insight into the interaction between firms' subsidiaries.

Appendices

4.A Extra Materials for Empirical Analysis

Table 4.A.1 shows different steps in cleaning the data for Japanese affiliates in the EU. Table 4.A.2 shows the number of non-manufacturing for Japanese big firms in the EU. This data is used to construct the Non-manufacturing firms' network index ($\ln(FirmsNetwork)$) described in section 4.2.2. Table 4.A.4 shows the correlation matrix for variables used in the baseline estimation.

Table 4.A.1: Data Cleaning

	Note	No. Firm	No. Affiliates
1	All countries	4966	29083
2	Europe	1119	3212
3	FDI	1035	1862
4	With year information	1035	1862
5	With geocode information	1000	1763
6	Only manufacturing parent	688	1233
7	Firms invested before 1990	161	223
8	Firms invested before and after 1990	143	604
9	Firms invested after 1990	384	605

Table 4.A.2: Numbers of Affiliates of each Non-manufacturing Big Firm in the EU

	Parent's SNA sector	Nfirm	min	max	Q25	Q50	Q75
1	Agriculture, Forestry and Fishing	5	1	6	1.00	2	4
2	Construction	14	1	6	1.00	1	2
3	Electricity, gas and water supply	10	1	7	1.25	3	3
4	Finance and insurance	45	1	13	1.00	2	3
5	Manufacturing	761	1	46	1.00	1	3
6	Mining	6	1	2	1.25	2	2
7	Real estate	5	1	3	1.00	1	2
8	Service activities	101	1	15	1.00	1	2
9	Transport and communications	53	1	22	1.00	1	3
10	Wholesale and retail trade	119	1	69	1.00	1	2

Table 4.A.3: List of Industries and Sectors

Manufacturing	Wholesale
Chemistry	Chemical wholesale
Drug	Electric equipment wholesale
Electric equipment	Fiber clothes wholesale
Fiber clothes	General Wholesale
Glass soil and stone	Glass soil and stone wholesale
Grocery	Grocery wholesale
Iron and steel	Machinery wholesale
Machine	Other wholesale
Metal products	Petroleum fuel wholesale
Non-ferrous metal	Pharmaceutical wholesale
Other manufacturing industry	Precision equipment wholesale
Petroleum coal	Steel & Metal Wholesale
Precision mechanical equipment	Transportation equipment wholesale
Pulp paper	
Rubber product	
Transport equipment	

Services	Finance
Advertisement	Bank
Architectural Design	Commodity futures
Building management security	Investment Management
Communication broadcasting	Investment Services , etc.
Consulting	Lease
Hotel	Life insurance
Information system software	Money Lending Credit Card
Leisure entertainment	Other financial
Machinery repair	Property and casualty insurance
Newspaper publishing	Securities
Other services	Trust bank
Real estate	
Temporary staffing business contract	
Travel	
Video and Music	

Note: These industries are classified by Toyo Keizai Database. Sectors are in concordance with SNA sectors except for Wholesale because SNA does not separate Wholesale and Retail. Headquarters only contain Headquarters.

Table 4.A.4: Correlation Between Agglomeration Index and Network Index

	ln(1+ Industry Count)	ln(1+ Group Count)	ln(1+ Industry Count) [Man]	ln(1+ Industry Count) [Who]	ln(1+ Group Count) [Man]	ln(1+ Group Count) [Who]	Non-manuf Firms Net-work	Non-manuf Affi. Net-work
ln(1+ Industry Count)	1							
ln(1+ Group Count)	0.176***	1						
ln(1+ Industry Count)[Man]	0.362***	0.178***	1					
ln(1+ Industry Count)[Who]	0.533***	0.0934***	0.00528***	1				
ln(1+ Group Count)[Man]	0.0859***	0.447***	0.248***	-0.000965	1			
ln(1+ Group Count)[Who]	0.0833***	0.497***	-0.000603	0.150***	-0.000162	1		
Non-manuf Firms Network	0.292***	0.180***	0.165***	0.142***	0.155***	0.0514***	1	
Non-manuf Affiliates Network	0.333***	0.0626***	0.129***	0.152***	0.00105	0.0299***	0.726***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 5

Conclusion

This dissertation has examined the role of trade costs in international trade and investment. While there has been extensive literature on this topic, the availability of microdata brought new insights into these matters. The second and third chapters address the role of friction in the flow of goods, while the fourth chapter discusses the friction in the flow of information.

The second chapter studies the role of the shipping sector in international trade. Utilizing a spike in fuel prices during the first ten years of the 2000s, the study provides a new estimate for the value of timeliness in international container trade. The theoretical framework explains how fuel price affects trade through speed adjustment of the shipping sector (slow steaming). The ship can adjust delivery days and freight costs when fuel prices increase. The change in delivery days affects consumers as if there is a decrease in the quality of goods.

The model is then tested using data on ship movement. Estimates of the elasticity of delivery days with respect to fuel prices were shown to depend on ship sizes. The average elasticity at the sample means $4,437$ TEU is 0.5. On average, a 10 percent increase in fuel prices increases delivery days by five percent. This translates to a delay of one day on the trade route between East Asia and North America (with an average of 20 days).

The value of timeliness is estimated by using the IV method. The IV considers the response of ships with respect to fuel prices and the composition of ship sizes for different trade routes. The elasticity of trade with respect to delivery days is estimated to be about -0.0814. A one percent delay in delivery is equivalent to an additional one percent in tariff. On the East Asia - North America trade route (20 days), one delayed day adds a five percent tariff equivalent. Between 2010 and 2011, fuel prices increased by 36 percent. This causes delays of 22 percent on average (for the East Asia - North America trade route, an equivalent of four days).

The study also shows geographical differences in ships' responses to fuel prices. Slow steaming does not happen in regional trade routes because it's easier to adjust schedules to respond to high fuel prices due to shorter distances. Among inter-regional trade routes, the slow steaming effects in the North-South group are twice as high as in the East-West group. It is because ship sizes on the East-West route are slightly bigger than that of the North-South, and the fuel consumption coefficient is higher for larger ships.

Changing functional forms of delivery days from constant to variable elasticity shows slightly different results. The elasticity of trade with respect to delivery days is $-0.0087Day$. On the East Asia - North America trade route (20 days), the elasticity equals -0.175. However, the time value estimates are smaller than in the literature in general. This is because previous literature measures the premium between air and ocean transportation, whereas this study measures the variation within only the container shipping sector.

This result by no means discredits the importance of time costs in international trade. On the contrary, it shows that time costs are present even after controlling for the quality of goods. Consumers are very sensitive to delivery time. After all, who does not enjoy one-day delivery service from Amazon Prime? Not only final consumers but final goods producers are sensitive to time costs. For these producers, faster (and more reliable) delivery times can save storage costs and, in some cases, get rid of the storage system (i.e., the "just-in-time" logistics system). This is the topic of the study in chapter three.

The third chapter explores the value of timeliness for final producers. Unlike final consumers, producers buy intermediate goods not to consume but to produce and sell them to the customers. This involves decisions on how much input to stock, how much to produce, and which buyers to sell to. The time costs for these producers are related to the cost of storing intermediate inputs and the associated opportunity costs of idle capital. If the costs per shipment for these intermediate inputs are minimal, they can have them shipped continuously and get rid of storage costs and reduce opportunity costs. While some multinational firms have succeeded by integrating the logistics system into their business models, most firms still have to deal with shipment costs and balance shipment frequency and size. This chapter shows how these decisions are affected by different types of trade costs: namely iceberg trade costs and shipment costs. The new insight of this study is to consider the role of buyers' heterogeneity in these decisions.

In the theoretical model, iceberg trade costs affect trade directly through sales and indirectly through shipment frequency and the number of buyers. Shipment cost effects are only through the adjustment of shipment frequency. Sellers with a small number of shipments may reduce shipment size when iceberg trade costs decrease. This happens if the adjustment through shipment frequency dominates the sales effects. The theoretical models also provide testable hypotheses, which are examined using the B/L data set.

The empirical analysis is carried out by estimating the gravity-like equation for the seller level and the buyer-seller level. The empirical results confirm that trade barriers reduce firms' trade volume, shipment frequency, and the number of buyers. Higher shipment costs may increase the average shipment size of a seller when the number of shipments is low, as predicted from the theoretical model.

The results also highlight the importance of buyer margins. While the effects of iceberg trade costs and fixed shipment costs on trade volume are mainly from an increase in shipment frequency rather than shipment size, more than half of an increase in shipment frequency is from an increase in the number of buyers. This new insight has not been explored in the literature on shipment frequency.

Indeed, looking for buyers is essential for any operating firm. This requires much effort in marketing and collecting customer information, especially for firms operating abroad. The producers may locate their subsidiaries closer to the market to acquire this information. However, being in the market does not automatically result in having relevant information. Having connections is one way to get access to helpful information. This is the topic of chapter four.

The fourth chapter shows how the need for "information sharing" affects the decision of manufacturing and wholesale subsidiaries' location of Japanese manufacturing firms in the EU. This chapter highlights the difference in the geographical patterns between wholesale subsidiaries and manufacturing plants. The new insight is that these wholesale affiliates tend to spread to a region with a strong presence of other Japanese service firms. While more clustered services can create agglomeration-like effects for firms in certain regions, the connectedness potentially reduces trade costs between regions. In this case, "information sharing" among multinational firms' subsidiaries does not need to happen in close proximity.

Regarding the relative location between firms' subsidiaries, manufacturing plants, and wholesale subsidiaries of the same industry and the same firms are not located near each other. On the other hand, manufacturing firms locate their plants in regions with a strong presence in their financial group's wholesale subsidiaries, which stresses the importance of sharing information beyond industry agglomeration. The study also confirms the agglomeration force for manufacturing plants within the industry and financial groups. The industry agglomeration effects are less than half of the within-financial-group agglomeration effects.

Many improvements can be made for future analysis. In chapter two, data on freight costs can improve the accuracy of the estimation. In particular, container transportation cost is proxied by the CIF-FOB margins that include all modes of transportation in the study. While an increase in fuel costs is likely to increase all modes of transportation costs, the magnitudes are different due to the different types of fuels and different shares of air and ocean transportation usage for different country pairs. In chapter three, more detailed data on the heterogeneity of buyers and sellers could show the differences in adjusting these margins for different types of firms. Investigating the potential relationship between buyers and sellers (such as intra-firm relationships) could shed light on different patterns of shipments. In chapter four, a global firm dataset could show different behaviors due to firms' nationalities and regional characteristics. Furthermore, data on intra-firm trade could help examine the interaction between firms' subsidiaries.

Finally, even though the dissertation has explored different aspects of trade costs, more studies are needed to perfect our understanding of the topic in the context of an ever-changing global economy.

Bibliography

Alessandria, George, Joseph P. Kaboski, and Virgiliu Midrigan (2010). "Inventories, Lumpy Trade, and Large Devaluations". In: *American Economic Review* 100.5, pp. 2304–2339.

Alfaro, Laura and Maggie Xiaoyang Chen (2014). "The global agglomeration of multinational firms". In: *Journal of International Economics* 94.2, pp. 263–276.

— (2018). "Transportation cost and the geography of foreign investment". In: *Handbook of International Trade and Transportation*. Edward Elgar Publishing, pp. 369–406.

Anderson, James E. and Eric Van Wincoop (2003). "Gravity with Gravitas: A Solution to the Border Puzzle". In: *American Economic Review* 93.1, pp. 170–192.

— (2004). "Trade Costs". In: *Journal of Economic Literature* 42.3, pp. 691–751.

Arkolakis, Costas (2010). "Market Penetration Costs and the New Consumers Margin in International Trade". In: *Journal of Political Economy* 118.6, pp. 1151–1199.

Baldwin, Richard and Javier Lopez-Gonzalez (2015). "Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses". In: *World Economy* 38.11, pp. 1682–1721.

Békés, Gábor and Balázs Muraközy (2012). "Temporary trade and heterogeneous firms". In: *Journal of International Economics* 87.2, pp. 232–246.

Belderbos, René, Helen S. Du, and Anthony Goerzen (2017). "Global Cities, Connectivity, and the Location Choice of MNC Regional Headquarters". In: *Journal of Management Studies* 54.8, pp. 1271–1302.

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott (2005). *Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods*. Working Paper 11404. National Bureau of Economic Research.

Bernard, Andrew B., Andreas Moxnes, and Karen Helene Ulltveit-Moe (2018). “Two-Sided Heterogeneity and Trade”. In: *Review of Economics and Statistics* 100.3, pp. 424–439.

Blonigen, Bruce A., Ronald B. Davies, Glen R. Waddell, and Hellen T. Naughton (2007). “FDI in space: Spatial autoregressive relationships in foreign direct investment”. In: *European Economic Review* 51.5, pp. 1303–1325.

Blonigen, Bruce A., Christopher J. Ellis, and Dietrich Fausten (2005). “Industrial groupings and strategic FDI”. In: *Japan and the World Economy* 17.2, pp. 125–150.

Borchert, Ingo and Yoto V. Yotov (2017). “Distance, globalization, and international trade”. In: *Economics Letters* 153, pp. 32–38.

Borgatti, Stephen P. and Martin G. Everett (1997). “Network analysis of 2-mode data”. In: *Social Networks* 19.3, pp. 243–269.

Bradford, Scott (2003). “Paying the Price: Final Goods Protection in OECD Countries”. In: *The Review of Economics and Statistics* 85.1, pp. 24–37.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou (2020). “Geography, Transportation, and Endogenous Trade Costs”. In: *Econometrica* 88.2, pp. 657–691.

Chaney, Thomas (2008). “Distorted gravity: The intensive and extensive margins of international trade”. In: *American Economic Review* 98.4, pp. 1707–1721.

Chen, Maggie Xiaoyang (2011). “Interdependence in multinational production networks”. In: *Canadian Journal of Economics* 44.3, pp. 930–956.

Defever, Fabrice (Sept. 2006). “Functional fragmentation and the location of multinational firms in the enlarged Europe”. In: *Regional Science and Urban Economics* 36.5, pp. 658–677.

— (2012). “The Spatial Organization of Multinational Firms”. In: *Canadian Journal of Economics* 45.2.

Duranton, Gilles and Diego Puga (2005). “From sectoral to functional urban specialisation”. In: *Journal of Urban Economics* 57.2, pp. 343–370.

Ekholm, Karolina, Rikard Forslid, and James R. Markusen (2007). “Export-Platform Foreign Direct Investment”. In: *Journal of the European Economic Association* 5.4, pp. 776–795.

Ferrari, Claudio, Francesco Parola, and Alessio Tei (2015). “Determinants of slow steaming and implications on service patterns”. In: *Maritime Policy & Management* 42.7, pp. 636–652.

Feyrer, James (2019). “Trade and Income Exploiting Time Series in Geography”. In: *American Economic Journal: Applied Economics* 11.4, pp. 1–35.

Goerzen, Anthony, Christian Geisler Asmussen, and Bo Bernhard Nielsen (2013). "Global cities and multinational enterprise location strategy". In: *Journal of International Business Studies* 44.5, pp. 427–450.

Head, Keith and Thierry Mayer (2004). "Market Potential and the Location of Japanese Investment in the European Union". In: *Review of Economics and Statistics* 86.4, pp. 959–972.

— (2014). "Gravity Equations: Workhorse,Toolkit, and Cookbook". In: *Handbook of International Economics*. Vol. 4. Elsevier, pp. 131–195.

Head, Keith and John Ries (2008). "FDI as an outcome of the market for corporate control: Theory and evidence". In: *Journal of International Economics* 74.1, pp. 2–20.

Head, Keith, John Ries, and Deborah Swenson (1995). "Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States". In: *Journal of International Economics* 38.3, pp. 223–247.

Helpman, Elhanan (1984). "A Simple Theory of International Trade with Multinational Corporations". In: *Journal of Political Economy* 92.3, pp. 451–471.

Helpman, Elhanan, Marc J. Melitz, and Stephen Ross Yeaple (Mar. 2004). "Export versus FDI with heterogeneous firms". In: *American Economic Review* 94.1, pp. 300–316.

Hornok, Cecília and Miklós Koren (2015a). "Administrative barriers to trade". In: *Journal of International Economics* 96.1, pp. 110–122.

— (2015b). "Per-Shipment Costs and the Lumpiness of International Trade". In: *Review of Economics and Statistics* 97.2, pp. 525–530.

Hummels, David (2007). "Transportation Costs and International Trade in the Second Era of Globalization". In: *Journal of Economic Perspectives* 21.3, pp. 131–154.

Hummels, David, Volodymyr Lugovskyy, and Alexandre Skiba (2007). *The Trade Reducing Effects of Market Power in International Shipping*. Working Paper 12914. National Bureau of Economic Research.

Hummels, David and Georg Schaur (2013). "Time as a Trade Barrier". In: *American Economic Review* 103.7, pp. 2935–2959.

IMO (2015). *Third IMO GHG Study 2014 Executive Summary and Final Report*.

Keller, Wolfgang and Stephen Ross Yeaple (2013). "The Gravity of Knowledge". In: *American Economic Review* 103.4.

Kosmas, Vasileios and Michele Acciaro (2017). "Bunker levy schemes for greenhouse gas (GHG) emission reduction in international shipping". In: *Transportation Research Part D: Transport and Environment* 57, pp. 195–206.

Kropf, Andreas and Philip Sauré (2014). "Fixed costs per shipment". In: *Journal of International Economics* 92.1, pp. 166–184.

Larmet, Vincent (2019). *cppRouting: Fast Implementation of Dijkstra Algorithm in R.*

Markusen, James R. (1984). "Multinationals, multi-plant economies, and the gains from trade". In: *Journal of International Economics* 16.3-4, pp. 205–226.

Mayer, Thierry, Isabelle Mejean, and Benjamin Nefussi (2010). "The location of domestic and foreign production affiliates by French multinational firms". In: *Journal of Urban Economics* 68.2, pp. 115–128.

Melitz, Marc J. (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity". In: *Econometrica* 71.6, pp. 1695–1725.

Miao, Guannan and Fabienne Fortainier (2017). *Estimating Transport and Insurance Costs of International Trade.* OECD Statistics Working Papers 2017/04.

Mirza, Daniel and Habib Zitouna (2009). *Oil Prices, Geography and Endogenous Regionalism: Too Much Ado About (Almost) Nothing.*

Notteboom, Theo E. and Bert Vernimmen (2009). "The effect of high fuel costs on liner service configuration in container shipping". In: *Journal of Transport Geography* 17.5, pp. 325–337.

OECD (2015). *World Energy Outlook 2015.* Washington: Organisation for Economic Co-operation and Development.

Psaraftis, Harilaos N. and Christos A. Kontovas (2013). "Speed models for energy-efficient maritime transportation: A taxonomy and survey". In: *Transportation Research Part C: Emerging Technologies* 26, pp. 331–351.

Rauch, James E. and Vitor Trindade (2002). "Ethnic Chinese Networks in International Trade". In: *The Review of Economics and Statistics* 84.1. Publisher: The MIT Press, pp. 116–130.

Sequeira, Sandra, Nathan Nunn, and Nancy Qian (2020). "Immigrants and the Making of America". In: *Review of Economic Studies* 87.1, pp. 382–419.

Silva, J. M. C. Santos and Silvana Tenreyro (2006). "The Log of Gravity". en. In: *Review of Economics and Statistics* 88.4, pp. 641–658.

South, Andy (2011). *rworldmap: A New R package for Mapping Global Data.*

Stopford, Martin (2009). *Maritime economics.* 3rd ed. London: Routledge.

Strauss-Kahn, Vanessa and Xavier Vives (2009). "Why and where do headquarters move?" In: *Regional Science and Urban Economics* 39.2, pp. 168–186.

Taylor, Peter J. (2001). "Specification of the World City Network". In: *Geographical Analysis* 33.2, pp. 181–194.

Train, Kenneth E. (2002). *Discrete choice methods with simulation*. Cambridge University Press.

UNCTAD (2010). *Oil Prices and Maritime Freight Rates: An Empirical Investigation*. UNCTAD/DTL/TLB/2009/2. UNCTAD secretariat, p. 40.

Von Below, David and Pierre-Louis Vezina (2016). "The Trade Consequences of Pricey Oil". In: *IMF Economic Review* 64.2, pp. 303–318.

Yeaple, Stephen Ross (2003). "The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment". In: *Journal of International Economics* 60.2, pp. 293–314.

— (2009). "Firm heterogeneity and the structure of U.S. multinational activity". In: *Journal of International Economics* 78.2, pp. 206–215.