Endogenous Transportation Costs*

Jose Asturias
Georgetown University Qatar

April 7, 2018

Abstract

Quantitative trade models used to evaluate the effects of policy changes have largely abstracted away from modeling the transportation industry. This paper extends a standard Armington trade model to incorporate an oligopolistically competitive transportation industry in which shippers endogenously choose a transportation technology. I collect detailed data on the containerized maritime transportation industry to calibrate the parameters of the model. I then conduct quantitative experiments in which there is a symmetric increase in tariffs. On average, changes in transportation costs account for almost half of the changes in welfare. These findings suggest that the endogeneity of transportation costs can substantially affect the estimated welfare effects of such a policy change.

JEL classifications: F13, F15, D43

Keywords: Trade costs, transportation costs, trade liberalization

*Email: jga35@georgetown.edu. I would like to thank Kerem Cosar and Kensuke Teshima for thoughtful discussions of this paper. I would also like to thank the following individuals for suggestions and comments: Cristina Arellano, Thomas Holmes, Timothy Kehoe, Bernabe Lopez-Martin, Thierry Mayer, Fabrizio Perri, Scott Petty, Terry Roe. Many thanks to the participants of the International Trade and Development Workshop at the University of Minnesota, Society for Economic Dynamics (Cyprus), Workshop on Dynamic Macroeconomics (Vigo, Spain), Congressional Budget Office, U.S. Government Accountability Office, Drexel University, University of Missouri (Columbia), Davidson College, Georgetown University in Qatar, NEUDC (Harvard University), Rocky Mountain Empirical Trade Conference (FREIT), North American Meeting of the Econometric Society, Spanish Economics Association Symposium, and the European Winter Meeting of the Econometric Society. I would also like to thank David Hummels and Douglas McDonald at the US Army Corp of Engineers for their help in acquiring data.

1 Introduction

Quantitative trade models have been used to study significant issues in international trade, such as determining the welfare effects of enacting policy changes. These models have largely simplified the modeling of the transportation industry through the use of iceberg costs. This notion dates back to Samuelson (1954), in which the rationale is stated as: "The simplest assumption is the following: to carry each good across the ocean you must pay some of the good itself. Rather than set up elaborate models of a merchant marine,. . . we can achieve our purpose by assuming that just as only a fraction of ice exported reaches its destination as unmelted ice..." Since then, little work has been done to quantify how welfare predictions would change if policy changes were evaluated using a model that matches micro-level facts about the transportation industry (e.g. market structure of the transportation industry or economies of scale in the transportation technology). This issue matters all the more as transportation costs are a major barrier faced by firms when exporting their products.[1]

In order to address this issue, the goal of this paper is twofold. First, this paper aims to build a general equilibrium trade model with a transportation industry that matches micro-level facts. The second goal is to calibrate the model to measure welfare changes as a result of a reduction in tariffs.

In order to achieve these goals, I first collected detailed data on the containerized maritime transportation industry. Focusing on one form of transportation allows the collection of detailed data on that industry. This particular form of transportation has two advantages. First, it is an important mode of transportation in international trade. Second, this form of transportation uses standardized shipping containers. The standardization of shipping containers provides a clean form to compare transportation costs across destinations.

I collected data on the cost to ship a standardized container from U.S. to foreign ports. I also collected data on the number of shippers that operate between these ports, the value of containerized trade flows, and distance. I do not find a strong relationship between transportation costs and distance. However, larger markets consistently have lower costs. There is suggestive evidence that the market structure of the transportation industry and the scale of shippers play a role in explaining why these larger markets have lower costs: larger markets tend to have more shipping firms and these shipping firms tend to be significantly larger.

I build a two-country Armington (1969) trade model with a transportation industry. The transportation industry has two key features. The first feature is that the industry is characterized by oligopolistic competition among shippers. The second feature, which is novel to the literature, is that shippers face a technological trade-off in which they can choose a lower marginal cost in exchange for a higher fixed cost. This paper shows that these features are consistent with the data. Furthermore, these features are consistent with research findings in the transportation literature, as well as with

[1] Anderson and van Wincoop (2004) indicate that the ad-valorem all-commodities arithmetic average is 10.7 percent for the United States. Furthermore, the transportation costs I have collected show that there is significant variation, implying that this number is significantly larger in some locations.

anecdotal evidence from the containerized maritime transportation industry. The result is that, as market size grows, there are two reasons why transportation costs may decline. First, more shippers enter, which lowers the markups that firms charge. Second, firms find it optimal to invest in a technology that has a lower marginal cost.

Another novel feature of the model is that each country can have multiple ports, and that goods can be transported between countries using different ports. The result is that there is competition across shippers operating in different port pairs. Ports are imperfectly substitutable and the parameter that governs the substitutability is estimated in the data.

I find analytical expressions that characterize the transportation industry equilibrium. Markups are a function of the number of shippers that operate between two ports as well as of the importance of the trade flows between these two ports. For example, all else equal, if a port pair accounts for a small fraction of overall trade between two countries, then markups are lower than for a port pair that accounts for a large fraction of overall trade. The expression for the markups, along with the pricing data, also allows me to find the marginal cost. I also derive analytical expressions that characterize the profitability (before the payment of fixed costs) of shippers. Since a continuous number of shippers is assumed, along with a free entry condition, the profitability of shippers allows to find the fixed cost of shippers.

Next, the model is taken to the data. I derive a condition from the model that allows me to estimate the parameter that governs the substitutability across port pairs between two countries. This condition takes the form of a gravity equation. I also use the analytical expressions that characterize the transportation industry equilibrium to find the marginal and fixed costs in each market. Larger markets consistently have higher fixed costs and lower marginal costs relative to smaller markets. These findings support the idea that shippers face a technological trade-off. As market size grows, they find it worthwhile to adopt technologies with a higher fixed cost and lower marginal cost.

The two-country model is calibrated separately for each U.S.-foreign country pair. I use the calibrated model to simulate a symmetric one percent increase in tariffs in both the United States and the foreign country. I find that changing transportation costs due to the smaller trade flows account for almost half of the welfare losses due to higher tariffs: the increases in transportation costs account for 46 percent of the losses in real income in the United States on average; in the foreign country, transportation costs account for 43 percent of the losses in real income on average. Thus, endogenously changing transportation costs are almost as important as the effects arising directly from tariffs.

I also investigate some of the model's predictions about the transportation industry after an increase in tariffs. The model predicts that, on average, a 1 percent increase in trade flows is associated with 0.25 percent decline in transportation costs from U.S. to foreign ports. This prediction of the model is compared with estimates using an instrumental variables (IV) approach from the empirical literature. This approach predicts that a 1 percent increase in trade flows is associated with a 0.24 percent decline in transportation costs. The model also predicts that increases in trade flows lead primarily to an increase in the scale of shippers. On average, a 1 percent increase in trade flows in the model is associated with a 0.85 percent increase in average shipper size. There is also a 0.16 percent increase in the number of shippers. Methods that use an IV approach imply that a 1 percent increase in trade flows is associated with a 0.90 percent increase in the average size of shippers and 0.10 percent increase in the number of shippers.

Finally, I find that markups decline after an increase in tariffs. To explain this finding, consider two contradictory forces that affect markups. First, the increase in tariffs reduces the number of shippers, which increases markups. Second, shippers also choose a technology with a lower fixed cost and higher marginal cost. The increase in the marginal cost lowers the markups of shippers due to imperfect pass-through: only a fraction of the increase in marginal cost is reflected in the change in prices. The latter effect dominates and shippers have lower markups after the increase in tariffs.

2 Related Literature

This paper contributes to various strands of literature in international trade. First, there is a large literature that uses quantitative trade models to study the welfare effects of policy changes, such as tariffs. This literature includes Eaton and Kortum (2002) and Alvarez and Lucas (2007), and is also related to the workhorse models considered by Arkolakis, Costinot, and Rodriguez-Clare (2012) since these models can be used to perform quantitative experiments on policy changes. In contrast to the existing literature, this is the first paper to study how welfare predictions change if a model of the transportation industry is incorporated into a quantitative trade model. The quantitative results show that, on average, changes in transportation costs account for almost half of the changes in real income when tariffs change. Thus, these results show that the inclusion of a transportation industry can have quantitatively relevant effects on the estimated welfare effects of a policy change.

This paper also contributes to the literature that studies the determinants of transportation costs.[2] Hummels and Skiba (2004b) and Skiba (Forthcoming) focus on "scale economies" in the maritime transportation industry. They find empirical evidence that transportation costs decline as trade flows increase. Another set of papers study the market power of shippers. Hummels, Lugovskyy, and Skiba (2009) study the quantitative importance of market power among shippers and its effect on trade flows. Francois and Wooton (2001) study theoretically how the market structure of the transportation industry impacts the estimated welfare effects of tariff changes. Fink, Mattoo, and Neagu (2002) and Moreira, Volpe, and Blyde (2008) find evidence consistent with oligopolistic competition in the shipping industry.

This paper contributes to this literature along two dimensions. This is the first paper that models

[2]Other papers focus on how port quality affects transportation costs, including Bougheas, Demetriades, and Morgenroth (1999), Clark, Dollar, and Micco (2004), and Wilmsmeier, Hoffmann, and Sanchez (2006). Holmes and Singer (2017) study the indivisibility constraint that firms face, which is that they must typically ship entire containers, and consolidation strategies that they can implement.

the technological choice of shippers in which there is a trade-off between fixed cost and marginal cost. This technological choice plays an important role in explaining the negative relationship between transportation costs and the size of trade flows. For example, when tariffs increase, the main driver of changing transportation costs is the technological choice of firms. A second contribution is that this paper extends the existing literature by allowing for goods to be transported between countries using different ports, which alters the competition shippers face. In this model, the level of competition that shippers face when operating between a given port pair (e.g. Los Angeles to Callo, Peru) is determined by the shippers that operate between these two ports, as well as by the shippers that operate along other port pairs (e.g. Oakland to Callo). The markups of shippers are characterized under this new framework. Furthermore, this paper shows how the parameter governing the substitutability across port pairs, which determines the level of competition that arises across shippers operating in different port pairs, can be estimated through the use of a gravity equation.

Other contemporaneous works have built trade models in which transportation costs and trade flows are jointly determined.[3] Wong (2018) studies how trade imbalances affect transportation costs using data from containerized maritime shipping.[4] Brancaccio, Kalouptsidi, and Papageorgiou (2017) develop a model based on dry bulk shipping, which is the shipping of non-containerized goods, with an emphasis on search costs between exporters and shippers. In contrast to these paper, I aim to understand the quantitative importance of modeling the transportation industry when evaluating the welfare effects of a change in tariffs. Since a change in tariffs affects the size of trade flows, my work emphasizes the relationship between the size of trade flows and transportation costs. I build a model with the aim of capturing this relationship, so as to conduct quantitative experiments.

This paper also contributes to the literature in international trade that identifies transportation costs. Indeed, as Anderson and van Wincoop (2004) state in their survey paper on trade costs: "An important theme is the many difficulties faced in obtaining accurate measures of trade costs..." Few papers in the literature use a direct measure of transportation costs as I do, such as a price quote to ship a container. One exception is the work of Wong (2018) who uses similar data in this paper on transportation costs. Another exception is Limao and Venables (2001), who obtain the price to ship a standard container from Baltimore to 64 destinations. Most commonly, the literature has relied on differences in the price of a good across locations (e.g. differences in salt prices across locations in India), including Asturias, García-Santana, and Ramos (2016), Atkin and Donaldson (2015), and Donaldson (Forthcoming). The work of Atkin and Donaldson (2015) points to three shortcomings of using price differences: 1) there may be unobserved product differences across locations (such

[3] There is a set of papers in economic geography that incorporates a transportation industry in which transportation costs are endogenously determined. These papers include Behrens, Gaigné, Ottaviano, and Thisse (2006), Behrens, Gaigne, and Thisse (2009), Behrens and Picard (2011), and Mori and Nishikimi (2002). Their goal is to theoretically understand how the inclusion of a transportation industry into a model of economic geography affects the concentration and location of economic activity.

[4] Jonkeren, Demirel, van Ommeren, and Rietveld (2011) empirically study the role of trade imbalances in determining transportation costs.

as quality), 2) it is necessary to know which locations trade in order for price differences to be informative, and 3) price differences contain both transportation costs and markups which may vary across locations.[5] The price quotes that I use allow to circumvent all of the issues faced by the literature that uses price differences across locations.[6]

3 Data

I collected detailed data on the containerized maritime transportation industry at the port-to-port level. The data includes transportation costs, number of shippers, value of trade flows, and distance. The data shows that large markets have lower transportation costs. There is suggestive evidence that market structure and the scale of transportation firms play a role in explaining these lower prices: large markets have more shippers and the average shipper size is larger.

I focus on maritime transportation since it is a leading mode of transportation in international trade. Maritime shipping is used most intensively when countries do not rely on land trade (non-contiguous countries). For example, when excluding land trade, approximately 60 percent of U.S. manufacturing imports by value are transported via maritime shipping and 40 percent by air. The reliance on maritime shipping for manufacturing imports is stronger in Latin American countries. Excluding land trade, maritime shipping accounts for 74 percent of trade in Argentina, 66 percent in Colombia, 64 percent in Mexico, 82 percent in Peru, and 75 percent in Venezuela by value, according to trade data from the Inter-American Development Bank. It is important to note that by weight, maritime shipping accounts for 98 percent of U.S. manufacturing imports.

Finally, containerization is a key method to transport these manufactured goods. Containerized shipping makes use of the shipping container, which is a standardized metal box that can be easily transported across multiple modes of transportation including ships, trains, and trucks.[7] The fact that it is standardized lowers the cost of loading and unloading. In the case of the United States, 72 percent of manufactured goods using maritime transportation are containerized (by value) according to the Waterborne Databank issued by the U.S. Maritime Administration.

3.1 Transportation Costs

Data on transportation costs are obtained from the freight forwarder Air Parcel Express (APX). A freight forwarder is a third-party logistics provider that helps arrange shipments and related paperwork for exporters. Freight forwarders advise exporters on transportation costs and other fees (port charges,

[5]The idea that higher quality products tend to be shipped further is known as the Alchian-Allen Conjecture, first found in Alchian and Allen (1964). If this is indeed the case, the use of pricing data would tend to bias the measurement of transportation costs. Hummels and Skiba (2004a) find empirical evidence for this conjecture.

[6]Other papers have used CIF/FOB measures from the IMF and UN, which has been criticized by Hummels and Lugovskyy (2006).

[7]Levinson (2008) provides an introduction to the invention and diffusion of the use of the shipping containers in international trade.

consular fees, costs of special documentation, insurance costs, and handling fees), as well as on import rules and regulations, methods of shipping, and the necessary documents.

The transportation costs are for transporting a 20-foot container from major U.S. ports to over 300 destinations abroad in October 2014. I use data from the top 10 U.S. ports in terms of number of containers loaded in the port. These ports account for approximately 85 percent of U.S. container traffic each year and include: Charleston, Houston, Los Angeles/Long Beach, New York/Newark, Norfolk, Oakland, Savannah, Seattle, and Tacoma.[8]

3.2 Number of Shippers

I acquired data collected in October 2014 on the number of shippers operating between ports from the Journal of Commerce (JOC), a respected trade publication in the transportation/logistics industry. The data comes from the JOC Global Sailings Schedule, which includes information on shipping schedules for containerized shippers.[9] For example, the shipping schedules indicate that A.P. Møller-Maersk, the largest shipping firm in the world, operates the ship Maersk Wolfsburg. This ship picked up cargo in Los Angeles on October 16, 2014 and delivered it to Puerto Quetzal, Guatemala, on October 23, 2014.

Much like a bus system in a city, shippers have lines that make multiple sequential stops. This is due to the economies of scope in containerized shipping. In this case, after Los Angeles, the Maersk Wolfsburg stopped in Lazaro Cardenas, Mexico (October 21) to pick up and drop off cargo before arriving in Puerto Quetzal. After Puerto Quetzal, the ship stopped in Acajutla, El Salvador (October 23), Corinto, Nicaragua (October 25), and finally Balboa, Panama (October 28). The data also contains schedules that make use of regional shipping hubs in transporting shipping containers to their final destination. This is analogous to transferring onto another bus to reach a final destination.

3.3 Containerized Trade Flows

Data on port-to-port containerized trade flows comes from the Waterborne Databank issued by the U.S. Maritime Administration for the years 2000-2005. The dataset contains information about U.S. international maritime trade on a port-to-port level. The data is broken down into the Harmonized System (HS) 6-product level, and includes the cost of transportation and insurance, weight, and whether the shipment is containerized.

The Waterborne Databank was discontinued in 2005, and there are no other available sources that report containerized trade at the port-to-port level. For this reason, I use the reported value of containerized trade in 2005 between ports, and adjust it for the percentage increase in bilateral trade over the 2005-2014 period between the United States and the countries where the ports are located.

[8] See U.S. Department of Transportation (2011), table 3.

[9] When determining the number of shippers, I combine shippers that operate in the same alliance. See the Appendix for more details.

This measure of trade flows is expected to be highly correlated to the true value of trade flows if data was available for 2014. To illustrate this point, I compute the correlation of the value of containerized trade flows in 2000 and 2005, and find it to be 0.85.

3.4 Port-to-Port Distance

I use the geospatial information system (GIS) to construct the distance between U.S. ports and foreign ports. I calculate the shortest navigable distance between the ports. For example, the distance between Los Angeles and Rotterdam incorporates the fact that ships can use the Panama Canal to minimize distance.

To do so, I combine two separate sets of geospatial data. First, I use the Global Shipping Lane Network shapefile provided by the Oak Ridge National Labs CTA Transportation Network Group. The shapefile contains information on global trading lanes used by maritime shippers. Second, I use the World Port Index provided by the National Geospatial Intelligence Agency to provide information on the location of ports. After combining these two sets of geospatial data, I use Network Analyst on ArcGIS to find the shortest path from the origin to destination port along the commonly used trading lanes. The final distance measure is given in nautical miles.[10]

3.5 Summary Statistics

Table 1 shows summary statistics for the assembled data. Each observation is a U.S.-foreign port pair. In the analysis, I only use observations for which all four data sources are available. Finally, each observation is referred to as a market.

First, there is a significant dispersion in transportation costs: the ratio between the 90th and 10th percentile is 2.45 (2,749/1,120). Second, the industry is characterized by a high degree of concentration among shippers. For example, the median market has only 2 shippers present. In fact, 46 percent of markets are serviced by monopolists and even the markets in the 90th percentile have only 4 shippers. This finding is consistent with the work of Hummels, Lugovskyy, and Skiba (2009), who also document high levels of concentration in the containerized maritime transportation industry. Third, there is large variation in the size of trade flows across markets: the ratio of the 90th and 10th percentile is 285 (1,708/6). Fourth, I define the average shipper size, which is the value of port-to-port containerized trade flows divided by the number of shippers. The data shows significant variation in the average size of shippers: the ratio of the 90th and 10th percentile is 189 (756/4).

3.6 How are Transportation Costs Related to Distance and Market Size?

I first study the relationship between distance and transportation costs since distance is a commonly used variable in the gravity literature to explain variation in trade costs.[11] I find that distance does

[10] A nautical mile is equivalent to 1,852 meters, and is a widely used unit in marine navigation.

[11] Anderson (2011) and Head and Mayer (2014) survey the use of gravity models.

a poor job of explaining the overall variation in transportation costs. To see why this is the case, I estimate the following regression

$$\log Freight_{i'j'} = \delta_0 + \delta_1 \log Dist_{i'j'} + \mathrm{FE}_{i',R(j')} + \epsilon_{i'j'}, \tag{1}$$

where $Freight_{i'j'}$ is the price to transport a container from port i' to j', $Dist_{i'j'}$ is the distance between the ports, $\mathrm{FE}_{i',R(j')}$ is a fixed effect for the origin port/destination region.[12] Table 2 reports the results of the estimation. The coefficient on distance is positive and statistically significant when fixed effects are not included. The result implies that a 10 percent increase in distance is correlated to a 1.0 percent increase in transportation costs. To understand the importance of distance in explaining transportation costs, I sort the predicted values of prices from column 1. The ratio between the 90th and 10th percentiles is 1.12 (1,857/1,662). Thus, ports that are far away (in the 90th percentile in terms of distance) have only 12 percent higher transportation costs than those that are nearby (in the 10th percentile in terms of distance). The second column of Table 2 reports the estimation when fixed effects are included. The significance of the distance coefficient disappears once including the fixed effects. Overall, distance does a poor job of explaining the large variation in transportation costs.

Next, I find that market size does a better job of explaining the variation in transportation costs. Since all markets in the dataset originate in the United States, I use the GDP of the destination country where the port is located as the measure of market size. I do not use the value of containerized trade flows since I mechanically expect lower transportation costs to be correlated with higher trade flows. The correlation of GDP and containerized trade flows is 0.93, so countries with larger GDPs also have larger trade flows in the dataset. I estimate equation 1 and include GDP as an independent variable. The results of this estimation can be seen in Table 3. The coefficient on GDP is negatively and statistically significant across all specifications. Furthermore, distance is statistically significant and positive once GDP is included in the regression.

I also investigate the relationship between the number of shippers and the average shipper size with respect to the market size. I estimate the following regression:

$$\log NumShippers_{i'j'} = \beta_0 + \beta_1 \log GDP_{C(j')} + \beta_2 \log Dist_{i'j'} + \mathrm{FE}_{i',R(j')} + \epsilon_{i'j'}, \tag{2}$$

where $C(j')$ indicates the country where port j' is located. I similarly estimate equation 2 except that average shipper size is the independent variable. The results can be found in Table 4 and 5. Larger markets consistently have more shippers and the average shipper size is larger. The results are robust across specifications and remain significant with the inclusion of distance.

[12] I use regional dummies defined by the World Bank. These regions are: East Asia and Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, North America, South Asia, and Sub-Saharan Africa.

Using an IV approach To further understand the relationship between transportation costs and market size, I adopt methods used by the empirical literature that studies the determinants of transportation costs. I estimate the following regression

$$\log Freight_{i'j'} = \phi_0 + \phi_1 \log TradeFlows_{i'j'} + \phi_2 \log Dist_{i'j'} + \mathrm{FE}_{i',R(j')} + \epsilon_{i'j'}, \tag{3}$$

where $TradeFlows_{i'j'}$ are the total trade flows between ports i' and j'. Notice that this regression suffers from issues of simultaneity bias that arise when estimating price on quantity. A similar expression was estimated in Hummels and Skiba (2004b) and Skiba (Forthcoming) in which these authors use population as an IV for trade flows in order to address this endogeneity issue. I follow this approach and use the population of the country where the destination port is located as an IV. There is a correlation of 0.19 between the population of the destination country and trade flows. Table 6 reports the results of the estimation of equation 3. Columns 1-3 report the results without the use of the IV and columns 4-6 report the same specifications with the use of the IV. All of the estimates of the coefficient are negative and these estimates become more negative once the IV is included. The last column is the preferred specification in which the elasticity of transportation costs with respect to trade flows is -0.24 percent.

I now study how the average shipper size and the number of shippers change with the size of trade flows. I decompose the cross-sectional differences in trade flows using the following equation

$$tradeflows = \frac{tradeflows}{shippers} * shippers. \tag{4}$$

Thus, changes in trade flows can be decomposed into those accounted for by the change in the average shipper size and the change in the number of shippers. For example, suppose that a 1 percent increase in trade flows is observed and that this increase is accompanied by a 1 percent increase in the trade flows per shipper. This implies that the number of shippers remains constant. Conversely, if a 1 percent increase in trade flows is accompanied by a 1 percent increase in the number of shippers then the trade flows per shipper remains constant.

I estimate the regression in equation 3 except that the number of shippers is replaced as independent variable. The results of the estimation can be found in Table 7. In the case of shippers, the IV does not significantly affect the results. The last column is the preferred specification, which implies an elasticity of the number of shippers with respect to trade flows of 0.10. Finally, I estimate equation 3 with the average shipper size as independent variable, and the results are reported in Table 8. The last column, which is the preferred specification, indicates an elasticity of 0.90 .

The results using the IV approach show additional evidence that transportation costs decline with trade flows. This suggests that increases in trade flows lead to both increases in the average shipper size and the number of shippers. Quantitatively, however, larger trade flows are primarily accompanied by increases in the average shipper size.

4 Model

I build a two-country Armington (1969) trade model that incorporates an oligopolistically competitive transportation industry. I use the Armington (1969) model because it has a simple CES demand structure for the consumption good. This simple demand structure allows finding analytical characterizations of the solution of the transportation industry. These expressions give a better understanding of how variables, such as markups, are determined in equilibrium. Furthermore, these expressions are useful when taking the model to the data as will be shown in Section 5.

This model of the transportation industry has two novel features relative to the existing literature. The first is that shippers face a technological trade-off: they can choose a lower marginal cost in exchange for a higher fixed cost. I discuss that this is consistent both with research in the transportation literature and with anecdotal evidence from the industry. This setup is similar to what is used by Sutton (1991) in which sunk costs are endogenous. I find that this technological choice, and not markups, is the main driver of differences in transportation costs that are documented in Section 3: high transportation costs are driven by a high marginal costs and not by high markups. Another novel feature relative to the existing literature is that I model competition across ports. For example, a container that is shipped from the United States to Peru can travel from Los Angeles to Callao or from Oakland to Callao. Thus, shippers face competition from firms that operate in other ports. I discuss how incorporating multiple ports affects the estimates of firm markups.

4.1 Consumer

There are two countries, i and j. Country j is populated by identical consumers of measure L_j. Each agent supplies one unit of labor to the market and spends her income on goods from both countries. The representative consumer of country j chooses the quantity of the good purchased from country j, c_{jj}, and from country i, c_{ij}, to solve

$$\max_{c_{jj},c_{ij}} \left(c_{jj}^{\frac{\sigma-1}{\sigma}} + \zeta^{\frac{1}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{5}$$

$$\text{subject to } p_{jj}c_{jj} + \tau_{ij}p_{ij}c_{ij} = w_j L_j + R_j, \tag{6}$$

where $\zeta \in [0,1]$ is a home bias parameter; σ is the elasticity of substitution of goods across countries; τ_{ij} is the ad-valorem tariff on goods from country i traveling to j; p_{jj} is the price of the country j good paid by country j and similarly for p_{ij}; w_j is the wage in country j; and R_j is tariff revenue that is rebated to the household lump sum. Consumers in j have the following demand for goods from i

$$c_{ij} = \frac{\zeta \left(w_j L_j + R_j \right)}{\tau_{ij}^{\sigma} p_{ij}^{\sigma} P_j^{1-\sigma}}, \tag{7}$$

and similarly for c_{jj}. The price index, P_j, is

$$P_j = \left(p_{jj}^{1-\sigma} + \zeta\tau_{ij}^{1-\sigma}p_{ij}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}. \tag{8}$$

I now allow for the possibility that a country has multiple ports. Suppose that country i has Ω_i ports indexed by i' and country j has Ω_j ports indexed by j'. The composite good from i to j, c_{ij}, is defined to be

$$c_{ij} = \left(\sum_{i'=1}^{\Omega_i}\sum_{j'=1}^{\Omega_j}\mathbb{I}_{i'j'}\beta_{i'}^{\frac{1}{\gamma}}\beta_{j'}^{\frac{1}{\gamma}}c_{i'j'}^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}}, \tag{9}$$

where $\mathbb{I}_{i'j'}$ is an indicator function that equals 1 if there is trade between ports i' and j'; $\beta_{i'}$ is an expenditure weight that is characterized by the importance of port i' in bilateral trade and similarly for $\beta_{j'}$; $c_{i'j'}$ is the consumption good transported from port i' to port j'; and γ is the elasticity of substitution across port pairs.[13] Goods shipped from Los Angeles to Callo and Oakland to Callo are differentiated products, with γ governing their substitutability. The demand for goods from port i' going to port j', $c_{i'j'}$, is

$$c_{i'j'} = \mathbb{I}_{ij,i'j'}\frac{\beta_{i'}\beta_{j'}c_{ij}}{p_{i'j'}^{\gamma}p_{ij}^{-\gamma}}, \tag{10}$$

where $p_{i'j'}$ is the price of $c_{i'j'}$, and the price index, p_{ij}, is defined as

$$p_{ij} = \left(\sum_{i'=1}^{\Omega_i}\sum_{j'=1}^{\Omega_j}\mathbb{I}_{i'j'}\beta_{i'}\beta_{,j'}p_{i'j'}^{1-\gamma}\right)^{\frac{1}{1-\gamma}}. \tag{11}$$

Defining c_{jj} using an expression similar to 9 is not necessary since consumers are assumed to receive the consumption good from the domestic producer with no transportation costs.

4.2 Production Firm

Country i has a representative production firm that operates under perfect competition. The firm has a linear production technology with a constant labor requirement, $1/x_i$. Since there are no domestic transportation costs, the firm in country i charges domestic consumers the factory-gate price, $p_{ii} = w_i/x_i$. In order to sell in the foreign country, the firm must pay $T_{i'j'}$ per unit of good to the shipper for goods transported from port i' to port j'. Thus, the firm charges $p_{i'j'} = w_i/x_i + T_{i'j'}$ for the product shipped from port i' to j'. The conditions are similar for goods produced by country j.

[13] Notice that $\mathbb{I}_{i'j'}$ is exogenous in the model. When going to the data, I will let this indicator function be 1 if trade between the two ports is observed, and the data needed to carry out the analysis is available.

4.3 Transportation Industry

The transportation industry is characterized by a three-stage entry game. In the first stage, N identical shippers enter the market to transport goods between ports i' and j' by paying a fixed cost, f, which is denominated in units of labor. Paying the fixed cost gives shippers access to a transportation technology with a labor requirement of $1/\phi$ to ship one unit of consumption good. In the second stage, shippers choose an optimal technology that has a lower unit labor requirement, $1/\Phi < 1/\phi$, and higher fixed cost, $F > f$. In the third stage, shippers compete a la Cournot in each direction: the quantity to transport from i' to j' and from j' to i'. I consider the pure strategy Nash equilibrium in which, in a given market, all shippers choose the same technology in the second stage and are symmetric in the third stage.

In this Section 4.3, I characterize the problem of the transportation industry between two ports. For ease of exposition, the notation for these two ports is omitted unless it is necessary to indicate the direction of trade.

I suppose that the shipper uses labor from country i to pay both the marginal and fixed cost. In Section 6.1, I discuss why this assumption does not appear to be an important determinant of the quantitative results. I also assume that there is a continuous number of shippers that satisfies a zero profit condition.[14]

4.3.1 Total Cost Function of Shippers

Suppose N shippers operating between ports i' and j' who have chosen a transportation technology with marginal cost w_i/Φ and fixed cost $w_i F$. The total cost function of the shipper n that operates between ports i' and j' is

$$c^n_{i'j'}\frac{w_i}{\Phi} + c^n_{j'i'}\frac{w_i}{\Phi} + w_i F \tag{12}$$

where $c^n_{i'j'}$ is the amount of goods transported by firm n from i' to j' and similarly for $c^n_{j'i'}$. Notice that by paying the fixed cost, the firm can transport goods from i to j and vice versa.

4.3.2 Third Stage for Shipper's Problem

I work backwards through the entry game. In the third stage, shipping firm n takes the following as given: the total quantity supplied by competitors, $c^{-n}_{i'j'}$ and $c^{-n}_{j'i'}$; the nominal income of the representative consumer in each country, characterized by the right-hand side of the budget constraint in equation 6; the factory-gate prices of the goods in each country, p_{ii} and p_{jj}; price $p_{i'j'}$ of all other port

[14]Notice that a model with a discrete number of shippers and my model give the same answer when my model predicts a whole number of shippers (e.g. 1, 2, 3, etc...). My model tends to *overpredict* the welfare effects relative to a model with discrete shippers if there is *no entry* in the latter model. Similarly, my model tends to *underpredict* the welfare effects relative to a model with discrete shippers if there is *entry* in the latter model. Thus, I consider the quantitative results to be an average effect of a change in tariffs.

pairs. The shipper chooses $c^n_{i'j'}$ and $c^n_{j'i'}$ to maximize profit

$$\pi = \max_{c^n_{i'j'}, c^n_{j'i'}} \quad c^n_{i'j'} \left(T_{i'j'} - \frac{w_i}{\Phi} \right) + c^n_{j'i'} \left(T_{j'i'} - \frac{w_i}{\Phi} \right), \tag{13}$$

where $c^n_{i'j'} + c^{-n}_{i'j'} = c_{i'j'}$ and similarly from j' to i'.

4.3.3 Second Stage for Shipper's Problem

In the first stage, N shippers have entered each market with a marginal cost of w_1/ϕ to transport a unit of consumption good. In the second stage, shippers choose a lower marginal cost by investing in a more productive technology, $\Phi \geq \phi$, which is associated with a higher fixed cost of

$$\log F = \alpha_1 \log \Phi + \alpha_0. \tag{14}$$

If $\alpha_0 = \log f - \alpha_1 \log \phi$, then the condition in equation 14 ensures that the original technology of fixed cost f and marginal cost w_1/ϕ is consistent with the menu of available technologies.

A firm takes the technological choice of all other firms as given, and chooses the technology that maximizes its profitability. In order to characterize the solution, suppose that all firms choose a technology of Φ and F except for a firm that deviates and instead chooses productivity Φ_d and fixed cost F_d. Let the profitability of the deviating firm in the third stage be π_d, characterized in equation 13. The following condition must hold so that no firm finds it profitable to deviate from the symmetric equilibrium

$$\frac{d\pi_d}{d\Phi_d} = \frac{dF_d}{d\Phi_d}. \tag{15}$$

4.3.4 First Stage for Shipper's Problem

The number of shippers that enter the market, N, is determined by the free entry condition

$$\pi = w_i F, \tag{16}$$

where π is characterized in equation 13 in stage 3 and F is determined in stage 2. Note that there are no aggregate profits from the transportation industry.

4.4 Government

The government gives consumers a lump-sum rebate on all tariff revenue. The tariff revenue in country j is

$$R_j = (\tau_{ij} - 1)\, p_{ij} c_{ij}. \tag{17}$$

There is a similar condition for country i.

4.5 Labor Clearing Conditions

Finally, the labor clearing conditions if only labor from country i is

$$\underbrace{\frac{c_{jj}}{\Phi_j} + \sum_{i'=1}^{\Omega_i}\sum_{j'=1}^{\Omega_j} \mathbb{I}_{j'i'}\frac{c_{j'i'}}{\Phi_j}}_{\text{Labor used by production firm}} \underbrace{+\sum_{i'=1}^{\Omega_i}\sum_{j'=1}^{\Omega_j} \mathbb{I}_{i'j'}\frac{c_{i'j'}}{\Phi^T} + \sum_{i'=1}^{\Omega_i}\sum_{j'=1}^{\Omega_j} \mathbb{I}_{j'i'}\frac{c_{j'i'}}{\Phi^T} + N_{i'j'}F^T_{i'j'}}_{\text{Labor used by transportation firms}} = \underbrace{L_i}_{\text{Labor endowment}}, \tag{18}$$

where $N_{i'j'}$ is the number of shippers that operate between i' and j', and $F^T_{i'j'}$ is the fixed cost that they choose in the second stage. In country j, the condition is

$$\underbrace{\frac{c_{jj}}{\phi_j} + \sum_{i'=1}^{\Omega_i}\sum_{j'=1}^{\Omega_j} \mathbb{I}_{j'i'}\frac{c_{j'i'}}{\phi_j}}_{\text{Labor used by production firm}} = \underbrace{L_j}_{\text{Labor endowment}}. \tag{19}$$

In equilibrium, the balanced trade condition between the two countries is redundant due to Walras's Law.

4.6 Characterizing Solution to Transportation Industry

I now analytically characterize the solution for the transportation industry. These characterizations will allow a better understanding of the determinants of the variables, such as markups, in the model. I also use these expressions when calibrating the model parameters.

Proposition 1. *Consider the market to ship consumption goods from i' to j'. Equilibrium transportation costs must satisfy*

$$T_{i'j'} = \underbrace{\frac{\epsilon_{i'j'}}{\epsilon_{i'j'}-1}}_{\textit{Markup}} \underbrace{\frac{w_i}{\Phi_{i'j'}}}_{\textit{Marginal Cost}}, \tag{20}$$

where $\epsilon_{i'j'}$ is the perceived price elasticity of demand of individual shippers. The perceived price elasticity of a shipper is

$$\epsilon_{i'j'} = N_{i'j'}\frac{T_{i'j'}}{p_{i'j'}}\kappa_{i'j'}, \tag{21}$$

where $\kappa_{i'j'}$ is determined as follows

$$\kappa_{i'j'} = \gamma - [\gamma-\sigma]\frac{p_{i'j'}c_{i'j'}}{p_{ij}c_{ij}} - (\sigma-1)\frac{\tau_{ij}p_{i'j'}c_{i'j'}}{w_jL_j+R_j}. \tag{22}$$

We have a similar set of conditions characterizing transportation costs from j' to i'.

Proof. See the Appendix for the full proof. □

Equation 20 shows that transportation costs can be expressed as a markup over marginal cost, where the markup is a function of the perceived price elasticity. The first term of the perceived price elasticity in equation 21 is the total number of shippers. As this number increases, the perceived price elasticity of shippers increases. In the case that the number of shippers approaches infinity, the market structure becomes perfectly competitive, and markups approach 1. The second term, $T_{i'j'}/p_{i'j'}$, is the relative size of transportation costs in the final delivery price of the consumption good. As this number declines, transportation costs account for a smaller fraction of the delivery price, and the perceived price elasticity declines.

The last term, $\kappa_{i'j'}$, is determined by the importance of trade flows between ports. To understand this term, consider the following three cases:

$$\kappa_{i'j'} = \begin{cases} \gamma & \text{if } \frac{p_{i'j'}c_{i'j'}}{p_{ij}c_{ij}} = 0 \text{ and } \frac{p_{i'j'}c_{i'j'}}{w_jL_j+T_j} = 0 \\ \sigma & \text{if } \frac{p_{i'j'}c_{i'j'}}{p_{ij}c_{ij}} = 1 \text{ and } \frac{p_{i'j'}c_{i'j'}}{w_jL_j+T_j} = 0 \\ 1 & \text{if } \frac{p_{i'j'}c_{i'j'}}{p_{ij}c_{ij}} = 1 \text{ and } \frac{p_{i'j'}c_{i'j'}}{w_jL_j+T_j} = 1 \end{cases}$$

The first case is one in which trade from port i' to j' accounts for a small portion of trade from country i to j. In this case, I find that $\kappa_{i'j'} = \gamma$. On the other hand, suppose that trade from i' to j' accounts for a large fraction of trade from i to j, but accounts for a small fraction of j's total expenditure share. Then $\kappa_{i'j'} = \sigma$, which is related to the elasticity of substitution across the country-specific goods. In the quantitative section, I find that $\gamma > \sigma$. Thus, as trade from i' to j' becomes more important, shippers are able to exercise more market power. This result is due to the fact that shippers internalize their ability to influence the p_{ij} price index. The last case is if trade from i' to j' accounts for a large fraction of country j's spending, then $\kappa_{i'j'} = 1$. This result is due to the fact that shippers understand that they can influence the country's aggregate price index, P_j.

If competition arising across ports was not explicitly modeled, then there would be two stark choices when taking the model to the data. The first would be to define the market as all the shippers that operate between the same port pair. Thus, a shipper that operates between Oakland and Callo would not consider the potential competition from shippers operating between Los Angeles and Callo. The second choice would be to define the market as all the shippers that operate between two countries. In that case, a shipper would consider competitors operating from Los Angeles to Callo as exactly identical to those operating from Oakland to Callo.

In this methodology, port pairs are imperfectly substitutable, and the parameter that governs the substitutability can be estimated in the data. The results are consistent with other models of oligopolistic competition with CES industry demand in which the market share of a firm is sufficient information to infer markups, as in Atkeson and Burstein (2008). In my methodology, three different market shares are needed to determine the markup. First, the market share of a shipper among the firms that operate between two ports, which in this case is the same as knowing the number of firms. Second, for total bilateral trade, the market share of trade accounted for by two ports. Third, for a country's total expenditure, the market share accounted for by trade between two ports.

In Proposition 2, I derive a condition that characterizes the profitability of shipping firms as a function of the value of bilateral trade and the number of shippers.

Proposition 2. *The equilibrium profits for shippers in equation 13 can be rewritten as*

$$\pi = \frac{c_{i'j'}p_{i'j'}}{N^2\kappa_{i'j'}} + \frac{c_{j'i'}p_{j'i'}}{N^2\kappa_{j'i'}}. \tag{23}$$

Proof. See the Appendix for the full proof. □

Equation 23 shows that profitability is increasing in the equilibrium trade flows between ports and decreasing in the number of shippers. The profitability of shippers is also decreasing in $\kappa_{i'j'}$ and $\kappa_{j'i'}$, which are determined by the relative importance of trade between two ports in accounting for overall trade between countries.

5 Calibration

I now calibrate the model presented in Section 4. In Section 6, I use the calibrated model to conduct quantitative experiments in which an increase in tariffs is simulated. The simulations allow studying both changes in the transportation industry and the implications on welfare.

I calibrate a two-country model for each foreign country in the dataset. Thus, the focus is on a bilateral reduction in tariffs as opposed to a multilateral reduction. In each calibration, the following parameters are normalized to 1 without loss of generality: the labor endowment of the United States, the productivity of the production firms, x_i, in each country, and wages in the United States. I also use the labor of the United States as the input for the shippers, which implies that the United States is country i.

It is useful to make a point about how the number of shippers is mapped from the data to the model. A shipper in this model corresponds to the number of firms that transport goods between a U.S. and a foreign port.

5.1 Elasticity of Substitution Across Ports Pairs (γ)

To estimate γ, I derive a gravity equation from the model of port-to-port trade flows, which is a result of the CES demand structure for the final good. I combine equations 7 and 10 to find the following condition that characterizes the port-to-port value of trade

$$\log p_{i'j'}c_{i'j'} = (1-\gamma)\log p_{i'j'} + \log \beta_{i'} + \log \beta_{j'} + \log\left(p_{ij}^{\gamma-\sigma}\tau_{ij}^{-\sigma}P_j^{\sigma-1}\zeta\left(w_jL_j + R_j\right)\right) \tag{24}$$

I then estimate the following regression using all of the observations in the dataset

$$\log TradeFlows_{i'j'} = \eta_0 + \eta_1 \log Price_{i'j'} + I_{i'} + I_{j'} + \epsilon_{i'j'}, \tag{25}$$

where $TradeFlows_{i'j'}$ is the value of trade flows from port i' in the United States to j' abroad, $Price_{i'j'}$ is the delivery price of the good at the destination, $I_{i'}$ is a U.S.-port fixed effect, and $I_{j'}$ is a foreign-port fixed effect. The intuition behind the identification strategy is that, controlling for origin and destination ports, I want to compare the changes in trade flows with differences in the delivery price of the good. If small changes in prices are associated with large changes in trade flows, then I infer that γ is relatively high. In the estimation, an origin U.S. port dummy is included to control for $\beta_{i'}$ as well as a foreign destination port dummy to control for the last two terms in equation 24.

As discussed in Section 4.2, the delivery price is equal to the factory-gate price, p_{ii}, plus transportation costs, $p^T_{i'j'}$. To determine the importance of transportation costs in the final delivery price, I use the fact that the ad-valorem transportation cost is 10.7 percent for the average commodity entering the United States, as documented by Anderson and van Wincoop (2004). The ad-valorem transportation cost from i' to j', $\tau^T_{i'j'}$, is defined as

$$\tau^T_{i'j'} = \frac{Price_{ii} + Freight_{i'j'}}{Price_{ii}}, \tag{26}$$

where $Price_{ii}$ is the value in dollars of the factory-gate good in country i. I find $Price_{ii}$ such that it satisfies

$$\sum_{i'=1}^{\Omega_i} \sum_{j'=1}^{\Omega_j} Share_{i'j'} \tau^T_{i'j'} = 1.107, \tag{27}$$

where $Share_{i'j'}$ is the fraction of exports from country i which is accounted for by exports from i' to j' in the data. To solve for this value, I first guess $Price_{ii}$ and find the ad-valorem transportation cost for each observation using equation 26 along with $Freight_{i'j'}$ from the pricing data. I then use equation 27 to find the implied aggregate ad-valorem transportation cost. The condition in equation 27 is satisfied when the factory-gate price in the U.S. is $Price_{ii} = \$12,265$. Thus, the destination price used in estimating equation 25 is $Price_{i'j'} = \$12,265 + Freight_{i'j'}$.

The relative importance of transportation costs in the overall price of the good will affect the estimation of γ. For example, suppose that the transportation costs are a relatively small part of the delivery price of the good. Then the delivery price will not vary significantly across destinations. Consequently, γ would need to be large in order to rationalize the observed variation in trade flows.

Table 9 reports the results of the estimation of equation 25. Across all specifications, there is a negative correlation between trade flows and the delivery price of the good. In the first specification, fixed effects are not included. In the second and third specifications, only U.S. port or destination port dummies are included. The fourth specification includes both U.S. and destination port dummies, and is the preferred specification since it is consistent with the model. The coefficient on the delivery price of the good is $\eta_1 = -15.11$, indicating that $\gamma = 16.11$. These results suggest that there is a high degree of substitutability across port pairs when transporting goods between countries.

5.2 Transportation Technology

In the next step, I determine the marginal cost of shipping goods from the United States to foreign ports, and the fixed cost that rationalizes the number of shippers in each market. With this information, I find evidence that firms face a trade-off between marginal cost and fixed cost. I also find that high prices are driven by the high marginal costs of serving the market and not high markups.

5.2.1 Marginal Cost

The next step is to use the transportation costs from the data in order to back out the marginal cost of transporting goods from U.S. to foreign ports. As can be seen in equation 20, transportation costs depend both on the marginal cost and the markup. The markup depends on factors related to competition, such as the number of shippers, as well as on the importance of the trading route in accounting for bilateral trade between the two countries. Thus, I adjust for these factors when using transportation costs to determine marginal cost. To understand the identification strategy, suppose that there are two markets with the same perceived price elasticity. However, if one market has a higher transportation cost, then the model interprets this fact as that market having a higher marginal cost.

I first calculate the perceived price elasticity characterized in equation 21. For the calculation, I use $\sigma = 5$, which in Section 6 generates similar trade elasticities relative to those reported in the literature (e.g. see Head and Mayer (2014)). This value of σ is also similar to the value estimated in Simonovska and Waugh (2014) for the canonical Armington (1969) model. Once the perceived price elasticity is calculated, I use equation 20 to find the marginal cost.

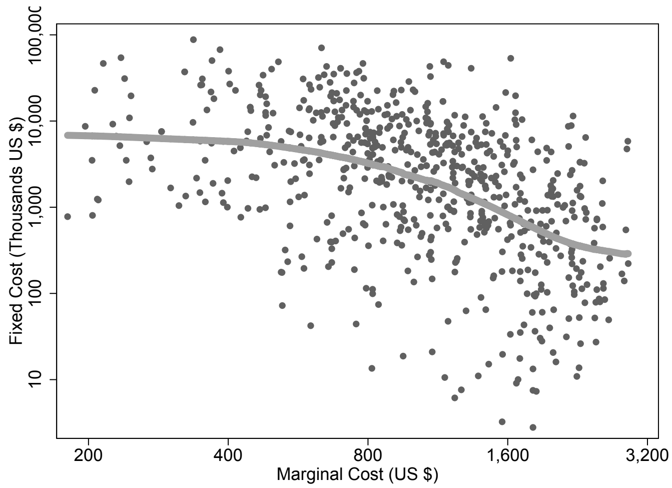
5.2.2 Fixed Cost

I use the condition in equation 23, along with the zero-profit condition, to determine the fixed cost of entry for each market. The zero profit condition implies that the fixed cost can be determined using information about the size of trade flows, number of shippers, and $\kappa_{i'j'}$. The intuition behind the identification strategy is that profits are an increasing function of the trade flows between two ports. Thus, if two markets are the same except one has larger trade flows, I interpret the latter market as having a higher fixed cost.

Since the number of shippers is continuous in the model and not in the data, I make an adjustment to the number of shippers used to find the fixed cost. For example, if there are 2 shippers that operate between two ports, then there is information about the bounds on the fixed costs: the fixed costs are low enough for two shippers to enter, but high enough to prevent a third shipper from entering. In order to capture a middle range of the possible fixed costs, if there are 2 shippers in the data, I use 2.5 shippers to find the fixed cost.

FIGURE 1

FIXED COST VERSUS MARGINAL COST



5.2.3 Description of Transportation Technology

Figure 1 shows a scatterplot of the fixed costs versus marginal costs found. The data was fitted with a LOWESS regression function, which is non-parametric in order to flexibly characterize the data. I find a robust pattern in which higher fixed costs are associated with lower marginal costs. In order to better understand how these findings relate to market size, Panel A of Figure 2 shows the scatterplot of fixed costs versus port-to-port trade flows; Panel B shows a scatterplot of the marginal cost versus port-to-port trade flows. Larger markets have both a higher fixed cost and a lower marginal cost.

5.2.4 Are High Transportation Costs Driven by Markups or Marginal Costs?

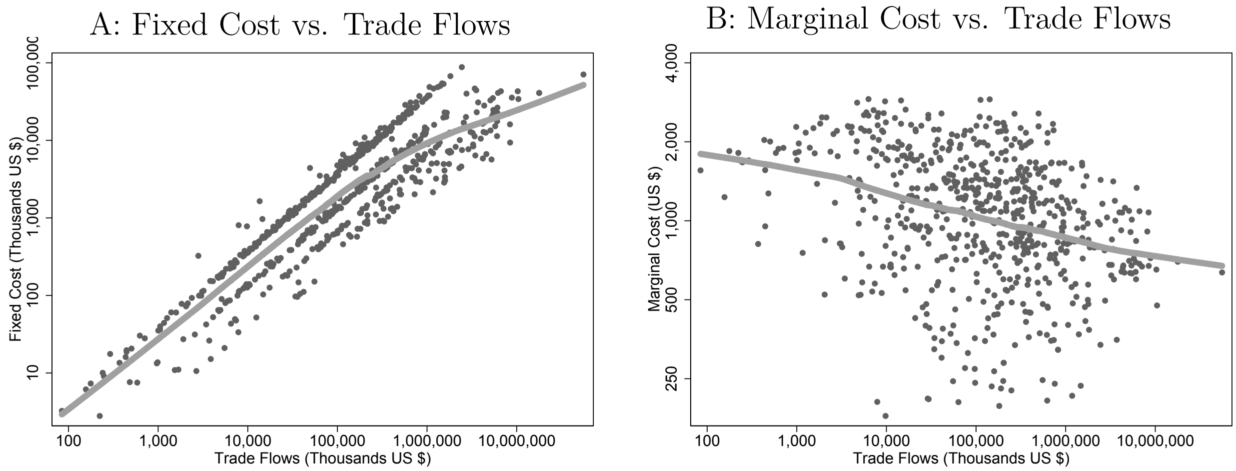
As can be seen in equation 20, high transportation costs can be driven by either high marginal costs or high markups. Thus, there are two potential forces in the model that can explain why large markets have lower transportation costs. The first is related to the technological choice of shippers. As market size grows, shippers find it worthwhile to adopt a technology with a lower marginal cost and a higher fixed cost. The second is related to the fact that larger markets tend to also have more shippers, which lowers markups.

In order to better understand these two forces, I calculate the correlation between transportation costs and marginal costs, and find that it is 0.85, which indicates that differences in marginal cost are the main driver of differences in transportation costs. As a second step, I calculate the correlation between transportation costs and markups, and find that it is -0.21. Thus, differences in markups tend to mitigate the effects of a high marginal cost on transportation costs.

These relationships are based on cross-sectional variation in the data. As discussed below, this finding is consistent with the quantitative exercise. When tariffs are raised in the quantitative exercise,

Figure 2

Fixed Cost and Marginal Cost vs. Trade Flows



changes in transportation costs in the model are driven entirely by changes marginal cost. Also, changes in markups mitigate the effect that increases in marginal cost have on transportation costs. This latter fact is consistent with models in which there is imperfect pass-through of changes in marginal cost: increases in marginal cost are not fully reflected in transportation costs and as a result markups decline.

5.2.5 Discussion of Fixed and Marginal Costs

I hypothesize that the technological factors of the transportation industry are the main driver of the relationship observed in Figures 1-2. There is evidence from the transportation literature that firms invest in larger ships in order to lower their average cost per container. For example, Cullinane and Khanna (2000) calculate this relationship and find that the total shipping cost per container declines with ship size (see Figure 6 for example). Sys, Blauwens, Omey, Voorde, and Witlox (2008) find similar results. There is also anecdotal evidence that is consistent with these findings. For example, an interview with Nils Anderson, the CEO of shipping firm AP Møller-Maersk, reveals that the company has invested in larger ships, such as the Triple E-class container ship, as a way of lowering the average cost per container. He cites that fuel costs on these larger ships are \$300-\$400 lower per 40-foot container for a round trip between Asia and Europe (see Milne (2013)).

If shippers use larger ships, through the lens of the model, there is an increase in the fixed cost. At the same time, the shipping firms would be expected to receive a lower marginal cost in order to compensate for the higher fixed cost. Suppose that firms face the following technological trade-off between fixed and marginal cost

$$\log FC^T = \alpha_0^T - \alpha_1^T \log MC^T, \tag{28}$$

where FC^T is the technological component of the fixed cost, MC^T is the technological component of the marginal cost, and α_0^T and $\alpha_1^T > 0$ are technological in nature. For example, if two ports are very far away, then α_0^T is higher.

In reality, shippers also face non-technological costs. For example, there may be poor regulatory environments that increase costs for firms (e.g. red tape, delays at ports, or strong unions that attempt to extract rents). I suppose that there are additional costs that affect the total fixed cost, FC, and the total marginal cost, MC, in the following way

$$FC = \xi_{FC}^R FC^T, \tag{29}$$

$$MC = \xi_{MC}^R MC^T, \tag{30}$$

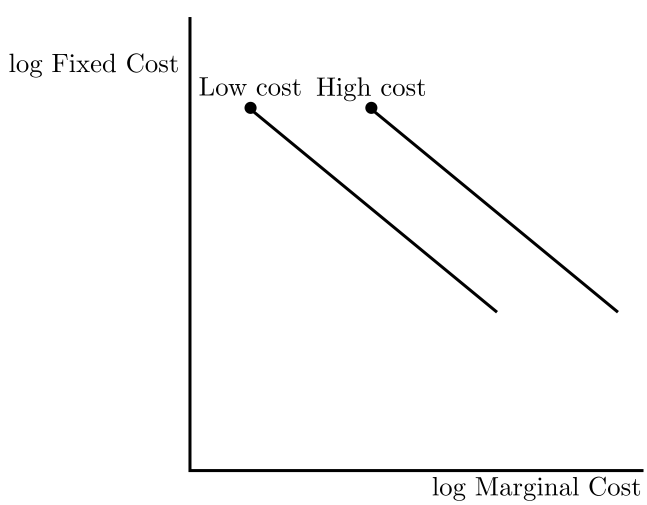
where $\xi_{FC}^R \geq 1$ and $\xi_{MC}^R \geq 1$. Parameters ξ_{FC}^R and ξ_{MC}^R represent the non-technological costs that firms face. I substitute equations 29 and 30 into equation 28 and re-arrange to find

$$\log FC = -\alpha_1^T \log MC + \left[\alpha_0^T + \alpha_1^T \log \xi_{MC}^R + \log \xi_{FC}^R\right]. \tag{31}$$

Thus, the trade-off between the total fixed cost and the total marginal cost is determined by α_1^T, which is technological in nature. For a given marginal cost, the level of fixed costs is determined by a mixture of technological and non-technological costs. Figure 3 shows an example of the choice that shippers face in a high- versus low-cost scenario. Notice that the formulation in equation 31 is consistent with equation 14 in the model when $\alpha_1 = \alpha_1^T$ and $\alpha_0 = \alpha_0^T + \alpha_1^T \log \xi_{MC}^R + \log \xi_{FC}^R$.

FIGURE 3

SIMPLE EXAMPLE OF FIXED COST VERSUS MARGINAL COST



5.3 Port-level Expenditure Weights

To determine port-level expenditure weights $\beta_{i'}$ and $\beta_{j'}$, 4the expenditure weight for one U.S. port and one foreign port is normalized to 1. This is a normalization because the expenditure weights determine the relative importance of each port. Given the normalization, I re-estimate equation 25 for each individual country to find the port-level expenditure weights. The omitted variable for the U.S. and foreign port is the port that was initially normalized to 1. When estimating equation 25 in Section 24 to find γ , all of the observations are pooled in order to maximize the amount of variation in the data to identify the substitutability across ports. In order to find the expenditure weights of ports, I only use the data for a particular destination country. The reason is that this allows better matching the trade flows across ports due to factors not related to transportation costs. For example, the distribution of port-level activity for United States-Peru trade may be different from that of United States-Germany. This gives more flexibility in matching the data and the market shares of ports in accounting for bilateral trade. Finally, for the trade flows, I use the average of imports and exports for each U.S.-foreign port pair.

5.4 Labor Endowments, L_i, and Home Bias Parameters, ζ

I now calibrate the labor endowments and home bias parameters. To do so, I feed into the model the number of shippers and transportation costs observed in the data. I also use the parameters derived from Sections 5.1-5.3. Thus, in addition to shippers and transportation costs, I also use the marginal and fixed costs of shippers found previously. I then simulate the model and find the labor endowment of the foreign country so that the model matches the relative GDP of the two countries. The home bias parameter is also set so that the model matches the ratio of imports to GDP for both countries. Note that since the GDP of each country is a calibration target, once the home bias parameter is calibrated, the model matches the import penetration of both countries.

Notice that when the model is calibrated, the shippers are not allowed to re-optimize in terms of number of shippers, technological choice, and transportation costs charged. This calibration, however, is a model-consistent way of identifying the labor endowments and the home bias parameters. For example, suppose that a set of labor endowments and home bias parameters are chosen, and data is simulated from the model, allowing for the decision of shippers to be endogenous. These model parameters can be recovered using this methodology on the model-simulated data.

In order to calibrate the labor endowment and the home bias parameters, it is necessary to map the transportation costs observed in the data, which are in dollars, to transportation costs in the model. The ad-valorem transportation costs are set to be equal in the model and data. Goods shipped from a U.S. port to foreign port must satisfy

$$\frac{Freight_{i'j'} + \$12,265}{\$12,265} = \frac{T_{i'j'} + p_{ii}}{p_{ii}}, \tag{32}$$

where $Freight_{i'j'}$ is the transportation costs in the data as mentioned in Section 3. In equation 32, I use the formulation for ad-valorem transportation costs from equation 26. Notice that the factory-gate price for the U.S. good, $p_{ii} = w_i/x_i$, is 1 since I normalize $w_i = x_i = 1$. I re-arrange to arrive at

$$T_{i'j'} = \frac{Freight_{i'j'}}{\$12,265}. \tag{33}$$

Once $T_{i'j'}$ has been determined, the parameter that governs the efficiency of the shippers, $\phi_{i'j'}$, can be backed out. This parameter is directly related to the marginal cost of shippers as can be seen in equation 20. I re-arrange this equation to find

$$\Phi = \frac{\epsilon_{i'j'}}{\epsilon_{i'j'} - 1} \frac{w_i}{T_{i'j'}}, \tag{34}$$

where I use the perceived price elasticities previously found in Section 5.2. Notice that I have normalized $w_i = 1$, which allows determining Φ without the need of simulating the model.

There is a similar issue of mapping the backed out fixed costs , which are in dollars, to the fixed costs in the model. The fixed costs are set in the model, F^T, to match the following condition

$$\frac{w_i F^T}{GDP_i^{Model}} = \frac{FC^{Data}}{GDP_i^{Data}}, \tag{35}$$

where GDP_i^{Model} and GDP_i^{Data} are the GDP of country i in the model and data respectively, and FC^{Data} is the fixed cost inferred from the data. I re-arrange equation 35 and solve for F^T.

5.5 Calibrating the Relationship Between FC versus MC

The parameters that govern the relationship between fixed cost and marginal cost can now be calibrated. As seen in Figure 3, the transportation technology is characterized by the slope of the line, governed by α_1 in equation 14, and the level, governed by α_0. In order to pin down the level, I use the fixed and marginal cost backed out in Section 5.2 as an option for shippers in each market. This aims at capturing the heterogeneous costs across pairs, which equation 31 shows is a mixture of technological and non–technological costs. Notice that this condition only pins down the level of the menu of technological choices that firms face and that the technological choice that firms make is endogenous. The backed out marginal and fixed costs in the data, which are in dollars, are mapped to those in the model as described in Section 5.4.

Parameter α_1 is calibrated in the following manner. First, the model is solved holding fixed the technological choice of firms in the case that all firms choose a very low fixed cost and high marginal cost. Here, I choose a marginal cost of \$2,000. Given the low fixed cost, there is positive entry in all markets. However, the condition that characterizes the optimal technology, characterized in equation 15, is not satisfied. Thus, I raise the fixed costs across markets until the optimal technology condition is satisfied. Furthermore, I drop any markets for which the number of shippers drops below 1 as the fixed cost is raised. In the final calibration, I drop 16 percent of the markets in the original sample. However, these markets tend to be small, and for the average country constitute 2 percent of total bilateral trade flows in the data.

Increasing α_1, which makes it more costly to lower the marginal cost, results in shippers electing to have a higher marginal cost and lower fixed cost technology. Thus, in equilibrium, a higher α_1 is associated with a higher average number of shippers. If $\alpha_1 = 2.60$, then there are 2.5 shippers in the average market. Table 1 shows that this is the average that can be seen in the data.

5.6 Distribution of Markups in Calibrated Model

I now discuss the distribution of markups implied by the model. I focus on the markups charged by shippers from U.S. ports to foreign ports. The markups are weighted by the number of firms in each market, in order to calculate the average markup charged by shippers. First, I find that the distribution is normally distributed and with a mean of 1.49. The distribution has a standard deviation of 0.06 and the 90th and 10th percentiles of the distribution are 1.55 and 1.40 respectively. The distribution is similar for the markups from foreign ports to U.S. ports.

I have not found other studies that calculate the markups of firms in the containerized maritime transportation industry that I could compare my results with. The only study that I could find was De Loecker and Eeckhout (2017), which reports weighted markups for broad two-digit NAICS industry codes in the United States. The closest industry is the transportation industry, which is code 48. This broad industry category includes diverse modes of transportation including air, rail, trucking, taxis, and pipelines (transporting oil and natural gas); this industry also includes the transport of both passengers and goods. For this industry, they find a weighted markup of a little under 1.25 (see Figure B.6 therein). Thus, the markups in my model are higher than those found in the transportation industry in the United States. Obviously, the transportation industry includes many modes of transportation that are not related to containerized maritime transportation. Thus, it is also useful to compare my markups with the weighted markups for the entire economy, which the authors find to be 1.67. Overall, the markups from the model are within a plausible range relative to those found by De Loecker and Eeckhout (2017).

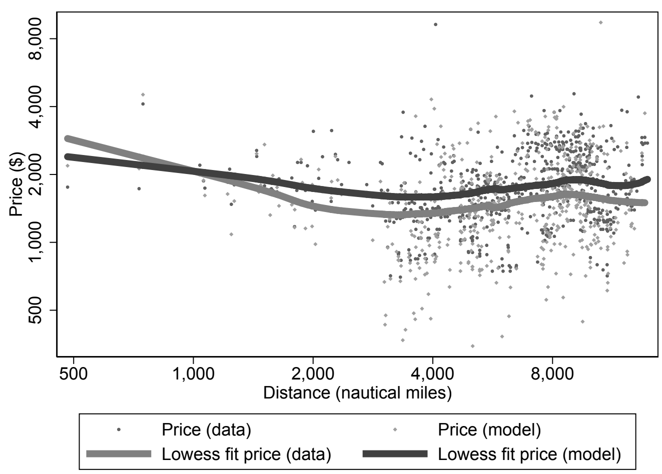
5.7 Comparing the Calibrated Model and Data

In Tables 10-14, I re-generate Tables 1-5 and include the same statistics for the model-generated data. In order to compare model output and data, I convert the transportation costs and value of trade flows from the model into dollars using the strategy outlined in Section 5.4. The model captures the key features of the data summarized in Section 3 reasonably well. Table 10 reports the unconditional distribution of transportation costs for shipping from the U.S. to foreign ports, the number of shippers, and trade flows. The distribution of transportation costs and number of shippers is of special interest given the goal of the quantitative exercise. The model matches the unconditional distribution well for both of these variables. Notice that the median number of shippers in the model is 2.5, which is a calibrated target. The other distribution moments, however, were not targeted in the calibration. Table 12 shows that the model generates a similar relationship between transportation costs and the GDP of countries as in the data, and similarly for Table 14 with average shipper size. Table 13 shows that the model does not generate as strong of a positive relationship between the number of shippers and GDP as observed in the data.

Figure 4 shows the relationship between transportation costs and distance in both the data and model. A LOWESS regression function is also included for both cases. In both cases, there is not a strong relationship between transportation costs and distance. Figure 5 shows the relationship between transportation costs and the GDP of the destination country in both the data and model. There is a negative correlation between transportation costs and the GDP of the destination country. Also, in both figures, the model reasonably captures the relationships found in the data.

As an additional robustness exercise, in Sections 6.2-6.3, I compare how transportation costs, number of shippers, and average shipper size change as trade flows change in the model and compare those with the results from the reduced form approach presented in Section 3.6.

FIGURE 4
PRICE VS. DISTANCE

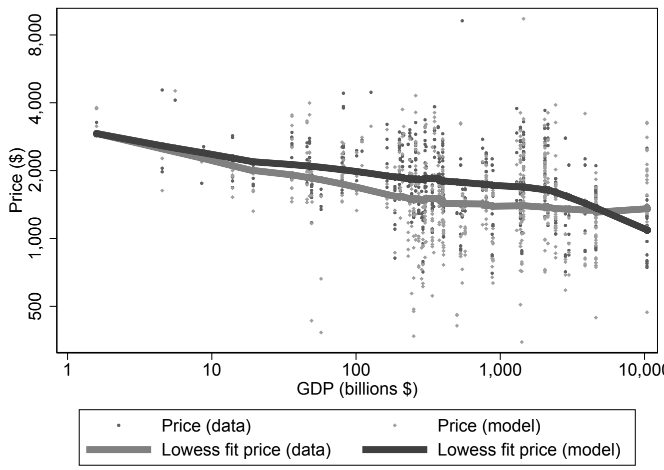


6 Quantitative Exercise

In this section, I raise tariffs by 1 percent for both the United States and the foreign country. As mentioned before, I calibrate a two-country model for each U.S.-foreign country pair. I decompose the changes in real income into what can be accounted for by changes in tariffs, transportation costs,

Figure 5

Price vs. GDP



and wages. I also study changes in the transportation industry as a result of the lower trade flows. Finally, I compare these results with those of other empirical methods that use IVs.

6.1 Changes in Real Income

The real income of a country before the change in tariffs, Y, is defined as

$$Y = \frac{wL + R}{P}, \tag{36}$$

where w is wage, L is the labor endowment, R is the tariff revenue, and P is the aggregate price index. The real income after the change in tariffs, Y', is similarly defined. The percentage change in real income, ΔY, is

$$\Delta Y = \frac{Y' - Y}{Y} * 100. \tag{37}$$

In all cases, there is a decline in real income for both the United States and the foreign country. On average, the United States loses -0.000037 percent and the foreign country loses -0.002637 percent. The average transportation costs increased by 1.49 percent, with the 90th and 10th percentiles at 1.99 and 1.17 respectively.

One concern when conducting the welfare analysis could be that the full cost of new transportation infrastructure is not taken into account. For example, suppose that two countries liberalize trade and as a result trade flows increase. In the quantitative exercise, more shippers enter, and shippers will choose a technology with a higher fixed cost and a lower marginal cost. These changes may require new or improved transportation infrastructure. I argue, however, that the cost associated with the new transportation infrastructure is considered in the welfare analysis. The reason is that shippers are usually charged fees to use the port, which in most cases is owned by the government or a related entity. If the shippers are charged for both the costs of operating the port and the capital costs, then the cost of infrastructure is captured by the marginal and fixed costs inferred.

The aim is to understand the importance of changing transportation costs in accounting for changes in real income. As mentioned before, quantitative trade models do not typically account for changes in transportation costs when evaluating the effects of changing policy. Thus, I decompose ΔY into the changes that can be accounted for by transportation costs, tariffs, and wages. To decompose changes in real income, first define $Y'_{Tariffs}$ as

$$Y'_{Tariffs} = \frac{wL + R'}{P'_{Tariffs}}, \tag{38}$$

where $P'_{Tariffs}$ is the price index if all equilibrium variables are left the same when computing the price index and only tariffs are changed. To compute $Y'_{Tariffs}$, I change the tariff revenue and adjust the price index to reflect the new tariff rates. In the case of an increase in tariffs, tariff revenue increases but the price index also increases. Notice that this is not an equilibrium outcome. For example, changes in tariffs would imply that wages and transportation costs would adjust, but if these variables are adjusted to the equilibrium outcome, the relative importance of the changes in tariffs versus transportation costs and wages cannot be understood. Similarly, let Y'_{Transp} be

$$Y'_{Transp} = \frac{wL + R}{P'_{Transp}}, \tag{39}$$

where P'_{Transp} is the price index if transportation costs are changed and everything else is left constant. Finally, let Y'_{Wages} be

$$Y'_{Wages} = \frac{w'L + R}{P'_{Wages}}, \tag{40}$$

where P'_{Wages} is the price index if wages are changed and all other equilibrium variables are left the same. The percentage change in real income accounted for by the change in tariffs is

$$\Delta Y_{Tariffs} = \left(\frac{Y'_{Tariffs} - Y}{Y}\right) * 100, \tag{41}$$

and similarly for ΔY_{Transp} and ΔY_{Wages}, which are the changes accounted for by transportation costs and wages, respectively. These changes do not necessarily sum to the percentage change in real income. Thus, a residual term is defined as follows

$$\Delta Y_{Residual} = \Delta Y - \Delta Y_{Tariffs} - \Delta Y_{Transp} - \Delta Y_{Wages}. \tag{42}$$

Qquantitatively, this residual term is small in all the simulations.

Transportation costs account for a significant portion of the changes in real income. On average, changing transportation costs account for 46 percent of the losses in real income for the United States. The 90th and 10th percentiles are 77 and 21 respectively. Tariffs account for 53 of the losses in real income for the United States. Thus, transportation costs are almost as important as tariffs in accounting for changes in real income. The changes in wages, on the other hand, are quantitatively less important. Wages account for 2 percent of the real income losses on average. The 90th and 10th percentiles are 11 and -3 respectively. This last result indicates that changes in wages are not the main driving factor of the results. It also confirms that using labor from the United States for the transportation industry does not drive the welfare results. We may be concerned, for example, that changes in the demand for transportation services may affect the wages of the United States and consequently impact the estimated welfare effects. Finally, the distribution of these components for the foreign country is similar to that of the United States. For example, on average, 43 percent of the losses in the foreign country are accounted for by changes in transportation costs, which is similar to the losses in the United States.

6.2 Changes in Transportation Costs

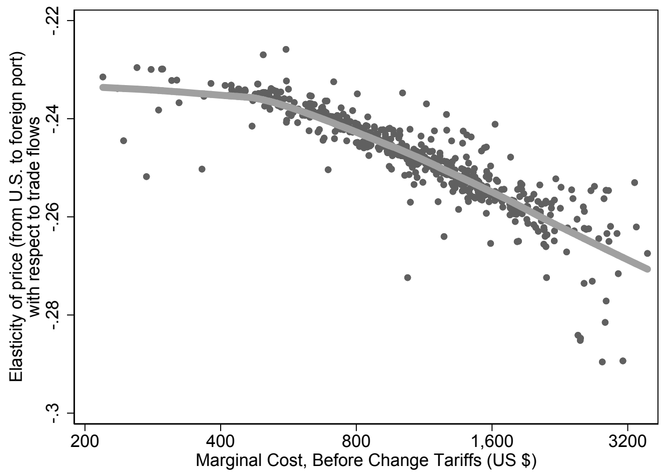
I now discuss the changes in transportation costs as a result of the reduction in tariffs. On average, the increase in tariffs results in a 1.49 percent increase in transportation costs from the United States to foreign ports. In equation 20, the percentage changes in transportation costs can be decomposed into those accounted for by the marginal cost and markups. Thus, there are two effects that could explain the increase in transportation costs. The first one is that firms change their technology due to the smaller trade flows by raising their marginal cost and lowering their fixed cost. The other effect is that firms exit due to the smaller trade flows, and as a result markups increase.

The average percentage change in markups is -0.14 and the average percentage change in marginal cost is 1.64. Thus, the changes in transportation costs are driven entirely by the changes in marginal cost. Notice that there are two contradictory forces that affect markups. On the one hand, the decrease in the number of shippers raises markups. On the other hand, the increase in marginal cost lowers markups because of imperfect pass-through: the full increase in marginal cost is not fully reflected in the transportation costs. The latter force dominates, which explains why markups decline.

Using the information from the model, I calculate the ratio of the percentage change in transportation costs and the percentage change in trade flows for each observation. Doing so yields an elasticity of the responsiveness of transportation costs to changes in trade flows through an increase in tariffs. In the mean case, a 1 percent decrease in trade flows is associated with an increase of 0.25 percent in transportation costs. This elasticity varies most consistently with the initial level of the marginal cost. Figure 6 shows the elasticity of transportation costs with respect to trade flows compared to the marginal cost before the change in tariffs. There is a negative relationship between these two variables. Thus, the small markets that have a high marginal cost also tend to be those with the most responsive transportation costs. Next, the markets with high marginal costs also have the most responsive changes in technology. Figure 7 compares the percentage change in the marginal cost with the marginal cost before the change in tariffs.

Figure 6

Elasticity of price (from U.S. to foreign port) with respect to trade flows vs. initial MC



As a robustness check, I compare the elasticity of transportation costs with respect to trade flows implied by the model and by the IV approach in Section 3.6. In the model, a 1 percent decrease in trade flows is associated with a 0.25 percent increase in transportation costs, while using the IV approach, there is an increase of 0.24 percent. Thus, the model and the IV approach imply similar elasticities for the responsiveness of transportation costs with respect to trade flows.

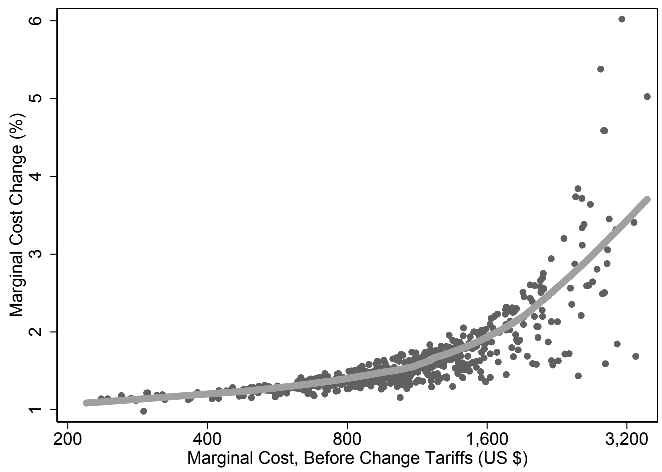
6.3 Does the Decline in Trade Flows Lead to Fewer Shippers or Lower Average Shipper Size?

Equation 4 shows that a decline in trade flows can lead to either fewer shippers or a decline in the average shipper size. On average, there is a decline of 0.88 percent in the number of shippers, and the average shipper size declines by 4.97 percent. These findings indicate that changes in trade flows lead primarily to changes in the average shipper size. Table 15 reports the mean elasticity for these statistics: on average, a 1 percent decrease in trade flows is associated with a 0.16 percent decrease in the number of shippers and a 0.85 percent decrease in the average shipper size. Thus, the average shipper size, and not the number of shippers, is more reactive to changes in trade flows.

As a robustness check, I compare these results with the elasticities implied by the IV approach in Section 3.6. The IV approach implies that a 1 percent decline in trade flows is associated with a decline of 0.10 percent in the number of shippers and 0.90 percent in the average shipper size. Thus,

Figure 7

Percent change MC after increase in tariffs vs. initial MC



the model and IV approach have similar predictions in terms of how responsive the number of shippers and average shipper size are to changes in trade flows.

Table 15 also compares the implied elasticities from the model as well as the empirical methods previously used in the literature. The predictions of the two methods are similar. There are two points to make about this comparison. First, notice that endogenous fixed costs are necessary in the model in order to match the entry patterns of shippers: a model without endogenous fixed costs would overpredict the entry of shippers and underpredict the average shipper size. Second, the reduced form strategy that makes use of the IV approach provides useful insights as to how the transportation industry responds to changes in trade flows. The question that I aim to answer, however, is related to the implications of considering the transportation industry when conducting a welfare analysis. A model is needed for the welfare analysis for two reasons. First, it is important to consider the fixed and marginal costs in the analysis. These costs, however, are typically not observed in the data and require an underlying model to recover them. Second, there are general equilibrium effects that the quantitative model incorporates when considering a change in tariffs.

7 Conclusion

In this paper, I have studied the effects of a change in tariffs on the transportation industry and the implications on the estimated welfare effects of the policy change. The mechanisms analyzed in this paper, however, apply more broadly to any situation in which there is a change in market size. For example, there has recently been emphasis by policymakers on reducing the procedural delays that shipments face in customs. Lowering these types of trade costs will increase trade flows and lower transportation costs in the same manner as tariffs. These effects may even be present in situations that are not directly related to international trade. For example, if a country improves its road infrastructure leading to more international trade, then this country may also enjoy lower international transportation costs.

The study of the transportation industry and of how transportation costs are determined is a relatively underexplored area in the field of international trade. This paper points to two areas for future research. First, the analysis has not focused on how transportation firms decide to structure their network. As mentioned before, there are strong complementarities across ports: a ship that travels from the United States to Chile can also make stops in ports in Peru. Thus, the transportation costs of a country are determined by the size of its own trade flows in addition to the trade flows of its neighbors. These spillovers across countries can have implications for trade policies that have not been explored in the literature. For example, if there is a U.S.-Chile free trade agreement, it could lower the transportation costs for neighboring countries like Peru. One exception is the work of Brancaccio, Kalouptsidi, and Papageorgiou (2017), who study these types of spillovers across countries.

Second, it would be useful to expand the analysis to include other modes of transportation. I have focused on the containerized maritime transportation industry, which allows collecting detailed data about this one mode of transportation and conducting a careful analysis. There are, however, alternative ways in which goods can be transported including trucks, rail, and plane. It would be useful to understand how sensitive changes in transportation costs are in other modes of transportation. Research suggests that there are declines in trade costs when total bilateral trade, which includes all modes of transportation, increases. For example, Anderson, Vesselovsky, and Yotov (2016) use a gravity framework and find that U.S.-Canada trade costs decline with increased trade flows. Breaking down these effects into different modes of transportation would be informative for policymakers in order to fully understand the effects of changes in trade policy.

References

Alchian, A. A., and W. R. Allen (1964): *University Economics*. Wadsworth.

Alvarez, F., and R. Lucas (2007): "General Equilibrium Analysis of the Eaton–Kortum Model of International Trade," *Journal of Monetary Economics*, 54(6), 1726–1768.

Anderson, J. E. (2011): "The Gravity Model," *Annual Review of Economics*, 3(1), 133–160.

Anderson, J. E., and E. van Wincoop (2004): "Trade Costs," *Journal of Economic Literature*, 42(3), 691–751.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

Armington, P. S. (1969): "A Theory of Demand for Products Distinguished by Place of Production," *Staff Papers (International Monetary Fund)*, 16(1), pp. 159–178.

Asturias, J., M. García-Santana, and R. Ramos (2016): "Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India," *Centre for Economic Policy Research Discussion Paper 11283*.

Atkeson, A., and A. Burstein (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98(5), 1998–2031.

Atkin, D., and D. Donaldson (2015): "Who's Getting Globalized? The Size and Implications of Intranational Trade Costs," *Working Paper*.

Behrens, K., C. Gaigné, G. I. Ottaviano, and J.-F. Thisse (2006): "How Density Economies in International Transportation Link the Internal Geography of Trading Partners," *Journal of Urban Economics*, 60(2), 248 – 263.

Behrens, K., C. Gaigne, and J.-F. Thisse (2009): "Industry Location and Welfare when Transport Costs are Endogenous," *Journal of Urban Economics*, 65(2), 195 – 208.

Behrens, K., and P. M. Picard (2011): "Transportation, Freight Rates, and Economic Geography," *Journal of International Economics*, 85(2), 280 – 291.

Bougheas, S., P. O. Demetriades, and E. L. Morgenroth (1999): "Infrastructure, Transport Costs and Trade," *Journal of International Economics*, 47(1), 169 – 189.

Brancaccio, G., M. Kalouptsidi, and T. Papageorgiou (2017): "Geography, Search Frictions, and Endogenous Trade Costs," *NBER Working Paper No. 23581.*

Clark, X., D. Dollar, and A. Micco (2004): "Port Efficiency, Maritime Transport Costs, and Bilateral Trade," *Journal of Development Economics*, 75(2), 417–450, 15th Inter American Seminar on Economics.

Cullinane, K., and M. Khanna (2000): "Economies of Scale in Large Containerships: Optimal Size and Geographical Implications," *Journal of Transport Geography*, 8(3), 181–195.

De Loecker, J., and J. Eeckhout (2017): "The Rise of Market Power and the Macroeconomic Implications," *Working Paper*.

Donaldson, D. (Forthcoming): "Railroads of the Raj: the Economic Impact of Transportation Infrastructure," *American Economic Review*.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Fink, C., A. Mattoo, and I. C. Neagu (2002): "Trade in International Maritime Services: How Much Does Policy Matter?," *The World Bank Economic Review*, 16(1), 81–108.

Francois, J. F., and I. Wooton (2001): "Trade in International Transport Services: The Role of Competition," *Review of International Economics*, 9(2), 249–261.

Head, K., and T. Mayer (2014): "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by E. H. Gita Gopinath, and K. Rogoff, vol. 4 of *Handbook of International Economics*, pp. 131–195. Elsevier.

Holmes, T. J., and E. Singer (2017): "Indivisibilities in Distribution," *Working Paper*.

Hummels, D., and V. Lugovskyy (2006): "Are Matched Partner Trade Statistics a Usable Measure of Transportation Costs?," *Review of International Economics*, 14(1), 69–86.

Hummels, D., V. Lugovskyy, and A. Skiba (2009): "The Trade Reducing Effects of Market Power in International Shipping," *Journal of Development Economics*, 89(1), 84–97.

Hummels, D., and A. Skiba (2004a): "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture," *Journal of Political Economy*, 112(6), 1384–1402.

——— (2004b): "A Virtuous Circle? Regional Tariff Liberalization and Scale Economies in Transport," in *FTAA and Beyond: Prospects for Integration in the America*, ed. by A. Estevadeordal, D. Rodrik, and A. V. Alan M. Taylor. Harvard University Press.

Jonkeren, O., E. Demirel, J. van Ommeren, and P. Rietveld (2011): "Endogenous Transport Prices and Trade Imbalances," *Journal of Economic Geography*, 11(3), 509–527.

Levinson, M. (2008): *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*. Princeton University Press.

Limao, N., and A. Venables (2001): “Infrastructure, Geographical Disadvantage, Transport Costs, and Trade,” *The World Bank Economic Review*, 15(3), 451–479.

Milne, R. (2013): “Nils Andersen, CEO, AP Møller-Maersk,” *Financial Times*.

Moreira, M., C. Volpe, and J. Blyde (2008): *Unclogging the Arteries: the Impact of Transport Costs on Latin American and Caribbean Trade*. Inter-American Development Bank.

Mori, T., and K. Nishikimi (2002): “Economies of Transport Density and Industrial Agglomeration,” *Regional Science and Urban Economics*, 32(2), 167–200.

Samuelson, P. A. (1954): “The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments,” *The Economic Journal*, 64(254), 264–289.

Simonovska, I., and M. E. Waugh (2014): “Trade Models, Trade Elasticities, and the Gains from Trade,” *NBER Working Paper No. 20495*.

Skiba, A. (Forthcoming): “Economies of Scale in Transportation and Regional Welfare,” *Review of International Economics*.

Sutton, J. (1991): *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. MIT Press.

Sys, C., G. Blauwens, E. Omey, E. V. D. Voorde, and F. Witlox (2008): “In Search of the Link between Ship Size and Operations,” *Transportation Planning and Technology*, 31(4), 435–463.

U.S. Department of Transportation (2011): *America’s Container Ports: Linking Markets at Home and Abroad*. U.S. Department of Transportation, Bureau of Transportation Statistics, Washington, DC.

Wilmsmeier, G., J. Hoffmann, and R. J. Sanchez (2006): “The Impact of Port Characteristics on International Maritime Transport Costs,” *Research in Transportation Economics*, 16, 117 – 140, Port Economics.

Wong, W. F. (2018): “The Round Trip Effect: Endogenous Transport Costs and International Trade,” *Working Paper*.

Tables

Table 1
Summary Statistics

			Percentile		
	St. dev.	Mean	10	50	90
Freight ($)	713	1,853	1,120	1,708	2,749
Number of shippers	1.4	2.1	1.0	2.0	4.0
Trade flows (million $)	3,862	906	6	131	1,708
Average shipper size (million $)	658	288	4	79	756
Distance (nautical miles)	2,931	6,970	3,328	7,059	10,716

Table 1 reports descriptive statistics for the assembled dataset, which is described in Section 3. It reports the standard deviation, mean, 10 percentile, 50th percentile, and 90th percentile for the following: freight prices ($), number of shippers, total port-to-port containerized trade flows (million $), average shipper size (million $), and distance (nautical miles). The average shipper size is calculated as the total port-to-port containerized trade flows divided by the number of shippers.

Table 2
Transportation Costs and Distance

	(1)	(2)
Distance	0.102***	-0.0162
	(0.0260)	(0.0373)
FE	No	Yes
Observations	664	664
R-Squared	0.0229	0.531

Table 2 reports the results of the estimation of equation 1. The dependent variable is the log transportation cost from a U.S. to a foreign port, which is described in Section 3.1. The independent variable is the distance, measured in nautical miles, between the two ports described in Section 3.4. Column 1 reports the results without fixed effects. Column 2 includes an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 3
Transportation Costs and Market Size

	(1)	(2)	(3)	(4)
GDP	-0.0880***	-0.101***	-0.0725***	-0.0755***
	(0.00816)	(0.00809)	(0.00698)	(0.00714)
Distance		0.171***		0.0665*
		(0.0240)		(0.0352)
FE	No	No	Yes	Yes
Observations	664	664	664	664
R-Squared	0.149	0.210	0.601	0.603

Table 3 reports the results of the estimation of equation 1 with the addition of the GDP of the destination country as an independent variable. The dependent variable is the log transportation cost from a U.S. to a foreign port, which is described in Section 3.1. The independent variables are the log GDP (U.S. dollars) of the destination country where the foreign port is located, and the log distance, measured in nautical miles, between the two ports described in Section 3.4. Columns 1-2 report the results without fixed effects. Columns 3-4 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 4
Number of Shippers and Market Size

	(1)	(2)	(3)	(4)
GDP	0.0909***	0.0985***	0.0772***	0.0757***
	(0.0141)	(0.0145)	(0.0156)	(0.0160)
Distance		-0.0984**		0.0331
		(0.0429)		(0.0788)
FE	No	No	Yes	Yes
Observations	664	664	664	664
R-Squared	0.0590	0.0664	0.266	0.266

Table 4 reports the results of the estimation of equation 2. The dependent variable is the log number of shippers between a U.S.-foreign port pair, which is described in Section 3.2. The independent variables are the log GDP (U.S. dollars) of the destination country where the foreign port is located, and the log distance, measured in nautical miles, between the two ports described in Section 3.4. Columns 1-2 report the results without fixed effects. Columns 3-4 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 5
Average Shipper Size and Market Size

	(1)	(2)	(3)	(4)
GDP	0.469***	0.522***	0.393***	0.397***
	(0.0495)	(0.0501)	(0.0452)	(0.0464)
Distance		-0.687***		-0.0737
		(0.149)		(0.229)
FE	No	No	Yes	Yes
Observations	664	664	664	664
R-Squared	0.119	0.147	0.529	0.529

Table 5 reports the results of the estimation of equation 2, except that the dependent variable is the log average shipper size between a U.S.-foreign port pair. To construct average shipper size, the number of shippers, which is described in Section 3.2, is divided by the total containerized trade flows in U.S. dollars between two ports, which is described in Section 3.3. The independent variables are the log GDP (U.S. dollars) of the destination country where the foreign port is located, and the log distance, measured in nautical miles, between the two ports described in Section 3.4. Columns 1-2 report the results without fixed effects. Columns 3-4 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 6
Transportation Costs and Trade Flows

	(1)	(2)	(3)	(4)	(5)	(6)
Trade flows	-0.0927***	-0.0914***	-0.0766***	-0.152***	-0.160***	-0.239***
	(0.00483)	(0.00481)	(0.00492)	(0.0183)	(0.0171)	(0.0359)
Distance		0.0690***	0.0203		0.0440*	0.0979*
		(0.0210)	(0.0317)		(0.0247)	(0.0536)
FE	No	No	Yes	No	No	Yes
IV (Population)	No	No	No	Yes	Yes	Yes
Observations	664	664	664	664	664	664
R-Squared	0.358	0.368	0.663	0.212	0.174	0.0679

Table 6 reports the results of the estimation of equation 3. The dependent variable is the log transportation cost from a U.S. to a foreign port, which is described in Section 3.1. The independent variables are the total value of containerized trade flows, described in Section 3.3, and the log distance, measured in nautical miles, between the two ports described in Section 3.4. The results in columns 4-6 use the population of the country where the foreign port is located as an IV for the value of containerized trade flows (U.S. dollars). Columns 1-2 and 4-5 report the results without fixed effects. Columns 3 and 6 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 7
Number of Shippers and Trade Flows

	(1)	(2)	(3)	(4)	(5)	(6)
Trade flows	0.106***	0.106***	0.130***	0.0975***	0.0968***	0.0980**
	(0.00902)	(0.00906)	(0.0110)	(0.0309)	(0.0282)	(0.0484)
Distance		0.00701	0.0539		0.00374	0.0693
		(0.0395)	(0.0708)		(0.0406)	(0.0723)
FE	No	No	Yes	No	No	Yes
IV (Population)	No	No	No	Yes	Yes	Yes
Observations	664	664	664	664	664	664
R-Squared	0.172	0.172	0.381	0.171	0.170	0.372

Table 7 reports the results of the estimation of equation 3, except that the dependent variable is the log number of shippers between a U.S.-foreign port pair, which is described in Section 3.2. The independent variables are the total value of containerized trade flows (U.S. dollars), described in Section 3.3, and the log distance, measured in nautical miles, between the two ports described in Section 3.4. The results in columns 4-6 use the population of the country where the foreign port is located as an IV for the value of containerized trade flows. Columns 1-2 and 4-5 report the results without fixed effects. Columns 3 and 6 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 8
Average Shipper Size and Trade Flows

	(1)	(2)	(3)	(4)	(5)	(6)
Trade flows	0.894***	0.894***	0.870***	0.903***	0.903***	0.902***
	(0.00902)	(0.00906)	(0.0110)	(0.0309)	(0.0282)	(0.0484)
Distance		-0.00701	-0.0539		-0.00374	-0.0693
		(0.0395)	(0.0708)		(0.0406)	(0.0723)
FE	No	No	Yes	No	No	Yes
IV (Population)	No	No	No	Yes	Yes	Yes
Observations	664	664	664	664	664	664
R-Squared	0.937	0.937	0.953	0.937	0.937	0.952

Table 8 reports the results of the estimation of equation 3, except that the dependent variable is average shipper size. The dependent variable is the log average shipper size between a U.S.-foreign port pair. To construct average shipper size, the total value of containerized trade flows in U.S. dollars between two ports, which is described in Section 3.3, is divided by the number of shippers, which is described in Section 3.2. The independent variables are the total value of containerized trade flows (U.S. dollars), described in Section 3.3, and the log distance, measured in nautical miles, between the two ports described in Section 3.4. The results in columns 4-6 use the population of the country where the foreign port is located as an IV for the value of containerized trade flows. Columns 1-2 and 4-5 report the results without fixed effects. Columns 3 and 6 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 9
Estimation of Elasticity of Substitution Across Port Pairs, γ

	(1)	(2)	(3)	(4)
Price at destination	-21.79***	-19.94***	-22.99***	-15.11***
	(1.541)	(1.547)	(1.812)	(1.791)
Origin Port FE	No	Yes	No	Yes
Destintion Port FE	No	No	Yes	Yes
Observations	664	664	664	664
R-Squared	0.232	0.306	0.641	0.727

Table 9 reports the results of the estimation of equation 25. The dependent variable is the log value of containerized trade flows from the U.S. to the foreign port (U.S. dollars), which is described in Section 3.3. The independent variable is the log delivery price of the U.S. good (U.S. dollars) at a destination port; the construction of this variable is described in Section 5.1. Column 1 reports the results without any fixed effects. Column 2 reports the results using only origin port fixed effects. Column 3 reports the results with only destination port fixed effects. Column 4 reports the results with both origin and destination port fixed effects. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 10
Summary Statistics

				Percentile		
		St. dev.	Mean	10	50	90
Freight ($)	Data	713	1,853	1,120	1,708	2,749
Freight ($)	Model	801	1,640	814	1,508	2,662
Number of shippers	Data	1.4	2.1	1.0	2.0	4.0
Number of shippers	Model	1.2	2.8	1.8	2.5	3.9
Trade flows (million $)	Data	3,862	906	6	131	1,708
Trade flows (million $)	Model	2,446	698	5	115	1,493
Average shipper size (million $)	Data	658	288	4	79	756
Average shipper size (million $)	Model	479	188	2	47	450

Table 10 reports descriptive statistics for the assembled dataset and the model-generated data. The summary statistics of the data are identical to those in Table 1. The assembled data is described in Section 3. The model and quantitative work are presented in Sections 4-6. The standard deviation, mean, 10 percentile, 50th percentile, and 90th percentile are reported for the following: freight prices ($), number of shippers, total port-to-port containerized trade flows (million $), average shipper size (million $), and distance (nautical miles). The average shipper size is calculated as the total port-to-port containerized trade flows divided by the number of shippers. The freight, trade flows, and average shipper size from the model have been converted into dollars using the steps described in Section 5.4.

Table 11

Transportation Costs and Distance (Model vs. Data)

	(1)	(2)	(3)	(4)
Distance	0.102***	0.0987***	-0.0162	-0.0684
	(0.0260)	(0.0360)	(0.0373)	(0.0573)
FE	No	No	Yes	Yes
Data/Model	Data	Model	Data	Model
Observations	664	556	664	556
R-Squared	0.0229	0.0134	0.531	0.371

Table 11 reports the results of the estimation of equation 1 using both data and output from the model. The results using the data are identical to those in Table 2. The dependent variable is the log transportation cost from a U.S. to a foreign port. The construction of transportation costs in the data is described in Section 3.1. The model and quantitative work are presented in Sections 4-6. The independent variable is the distance, measured in nautical miles, between the two ports described in Section 3.4. Columns 1-2 report the results without fixed effects. Columns 3-4 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 12

Transportation Costs and Market Size (Model vs. Data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP	-0.0880***	-0.0675***	-0.101***	-0.0814***	-0.0725***	-0.0479***	-0.0755***	-0.0474***
	(0.00816)	(0.0124)	(0.00809)	(0.0126)	(0.00698)	(0.0126)	(0.00714)	(0.0131)
Distance			0.171***	0.157***			0.0665*	-0.00714
			(0.0240)	(0.0359)			(0.0352)	(0.0591)
FE	No	No	No	No	Yes	Yes	Yes	Yes
Data/Model	Data	Model	Data	Model	Data	Model	Data	Model
Observations	664	556	664	556	664	556	664	556
R-Squared	0.149	0.0510	0.210	0.0829	0.601	0.387	0.603	0.387

Table 12 reports the results of the estimation of equation 1 with the addition of the GDP of the destination country as an independent variable using both data and output from the model. The results using the data are identical to those in Table 3. The dependent variable is the log transportation cost from a U.S. to a foreign port. The construction of transportation costs in the data is described in Section 3.1. The model and quantitative work are presented in Sections 4-6. The independent variables are the log GDP of the destination country where the foreign port is located, which in the case of the data is in U.S. dollars; and the log distance, measured in nautical miles, between the two ports described in Section 3.4. Columns 1-4 report the results without fixed effects. Columns 5-8 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 13

Number of Shippers and Market Size (Model vs. Data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP	0.0909***	-0.00930	0.0985***	0.00384	0.0772***	-0.0218**	0.0757***	-0.0176*
	(0.0141)	(0.00896)	(0.0145)	(0.00899)	(0.0156)	(0.00928)	(0.0160)	(0.00967)
Distance			-0.0984**	-0.149***			0.0331	-0.0655
			(0.0429)	(0.0257)			(0.0788)	(0.0435)
FE	No	No	No	No	Yes	Yes	Yes	Yes
Data/Model	Data	Model	Data	Model	Data	Model	Data	Model
Observations	664	556	664	556	664	556	664	556
R-Squared	0.0590	0.00194	0.0664	0.0591	0.266	0.331	0.266	0.334

Table 13 reports the results of the estimation of equation 2 using both data and output from the model. The results using the data are identical to those in Table 4. The dependent variable is the log number of shippers between a U.S.-foreign port pair. In the case of the data, its construction is described in Section 3.2. The model and quantitative work are presented in Sections 4-6. The independent variables are the log GDP of the destination country where the foreign port is located, which in the case of the data is in U.S. dollars; and the log distance, measured in nautical miles, between the two ports described in Section 3.4. Columns 1-4 report the results without fixed effects. Columns 5-8 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 14

Average Shipper Size and Market Size (Model vs. Data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP	0.469***	0.423***	0.522***	0.456***	0.393***	0.402***	0.397***	0.402***
	(0.0495)	(0.0551)	(0.0501)	(0.0567)	(0.0452)	(0.0568)	(0.0464)	(0.0594)
Distance			-0.687***	-0.378**			-0.0737	0.00692
			(0.149)	(0.162)			(0.229)	(0.267)
FE	No	No	No	No	Yes	Yes	Yes	Yes
Data/Model	Data	Model	Data	Model	Data	Model	Data	Model
Observations	664	556	664	556	664	556	664	556
R-Squared	0.119	0.0962	0.147	0.105	0.529	0.399	0.529	0.399

Table 14 reports the results of the estimation of equation 2 except that the dependent variable is average shipper size using both data and output from the model. The results using the data are identical to those in Table 4. The dependent variable is the log average shipper size between a U.S.-foreign port pair. To construct average shipper size, the total value of containerized trade flows between two ports is divided by the number of shippers. In the case of the data, the number of shippers is described in Section 3.2 and the value of containerized trade flows in U.S. dollars is described in Section 3.3. The model and quantitative work are presented in Sections 4-6. The independent variables are the log GDP of the destination country where the foreign port is located, which in the case of the data is in U.S. dollars; and the log distance, measured in nautical miles, between the two ports described in Section 3.4. Columns 1-4 report the results without fixed effects. Columns 5-8 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table 15
Elasticities with Respect to Trade Flows (Model vs. IV)

Variable	Model (mean)	Regression with IV
Price	-0.25	-0.24
Shippers	0.16	0.10
Average shiper size	0.85	0.90

Table 15 reports the elasticities with respect to trade flows for prices, shippers, and average shipper size implied by the model and the regression with the IV. In the case of the model, first the percentage change in the variable of interest is calculated and it is divided by the percentage change in trade flows to find the elasticity for each U.S.-foreign port pair. The mean elasticity over all U.S.-foreign port pairs is then obtained. In the case of the regression with the IV, the price elasticity results come from the estimation of equation 3 using the population of the destination country as an IV for containerized trade flows. Similar equations are estimated for shippers and average shipper size as dependent variables to find the corresponding elasticities with respect to trade flows for those variables. The results from the estimation are reported in Tables 6-8. The preferred specification, which controls for distance and origin port-destination region fixed effects, is in column 6 of these tables.

A Appendix: Data

In this section, we discuss additional details in the preparation and usage of the assembled data.

A.1 Transportation costs

Transportation costs are collected from the freight forward APX. APX has a shipping rate "calculator" intended for users to receive immediate shipping quotes (http://www.apx-ocean-freight.com/get-quote.html). We use quotes for full container load (FCL) in which the exporter is purchasing the use of the entire container. APX has an option for less than container load (LCL) in which they combine shipments of various clients. We receive quotes to ship a 20 foot container of textiles from all available US origins to all available destinations. Although we request the quotes to ship textiles, we tried various commodities and found that there were no differences in transportation costs to ship a 20 foot container depending on the product we intended to ship.

A.2 Number of shippers

The data on the number of shippers comes from the Journal of Commerce (JOC) Global Sailings Schedule. This information is used by exporters to determine voyage dates for ships. We download all of the data for October 15, 2014 to November 14, 2014. Within this date range, we use the "carrier" field to determine the shipping firm. Furthermore, we combine shippers that form part of the various alliances: CKYH (Hanjin, Yang Ming, K-Line, COSCO Container Lines), G6 (Hapag-Lloyd, APL, OOCL, NYK Container Line, Hyundai, Mitsui O.S.K. Lines), and Maresk-Safmarine (Maersk Line, Safmarine).

A.3 Port-to-Port Containerized Trade Flows

The port-to-port containerized trade flows are derived from the Waterborne Databanks issued by the US Maritime Administration for the years 2000-2005. The data contains information on US waterborne international trade. For exports, it contains information about the the US port of origin, the foreign port where the shipment is headed next, the final destination country. Notice that the destination port where the shipment is headed next does not need to be located in the final destination country because a container can be transhipped at an intermediary port before reaching a final destination. For the same reason, we also have information about the country of origin and whether the shipment was transshipped in the United States on its way to a foreign destination. The data also contains HS 6 product code, SITC revision 3 industry code, value of shipment in US dollars, weight of the shipment in kilograms, percent of shipment value that is containerized, and the percent of weight that is containerized. For imports, the data is similar except that the data also contains CIF charges (insurance and freight). We consider trade flows that are US exports (the origin is a US port) or US imports (the final destination port is a US port). Thus, we do not consider trans-shipments through the US.

Because the Waterborne Databanks are the only source of port-to-port containerized trade flow data, we scale up the value of both imports and exports by the percent increase in total bilateral trade between the United States and the foreign country using trade data from the World Integrated Trade Solutions (WITS) by the World Bank.

We list here some additional notes regarding the trade flows used in Section 5:

- Section 5.1: Notice that $TradeFlows_{i'j'}$ in equation 25 corresponds to $p_{i'j'}c_{i'j'}$ in the model, where $p_{i'j'} = w_i/x_i + T_{i'j'}$ includes transportation costs. To account for the transportation costs, we multiply the value of exports in the data by $\tau^T_{i'j'}$ that we found in equation 26 to find $TradeFlows_{i'j'}$.

- Section 5.2.1: From the United States to the foreign port, we use $TradeFlows_{i'j'}$ for the value of trade, just as we did in Section 5.1. This give us enough information to calculate $\epsilon_{i'j'}$, in equation 21, and $\kappa_{i'j'}$, in equation 22.

- Section 5.2.2: We have the information needed to calculate the profitability of shipping goods from the US to the foreign port, found in equation 23. To calculate $\kappa_{i'j'}$ from the foreign port to the US port, we need to use the value of trade flows that incorporate the transportation costs (as in the same manner as Section 5.1 and Section 5.2.1). To do so, we use the reported import values plus the CIF charges. With that information, we can calculate the profitability of shippers.

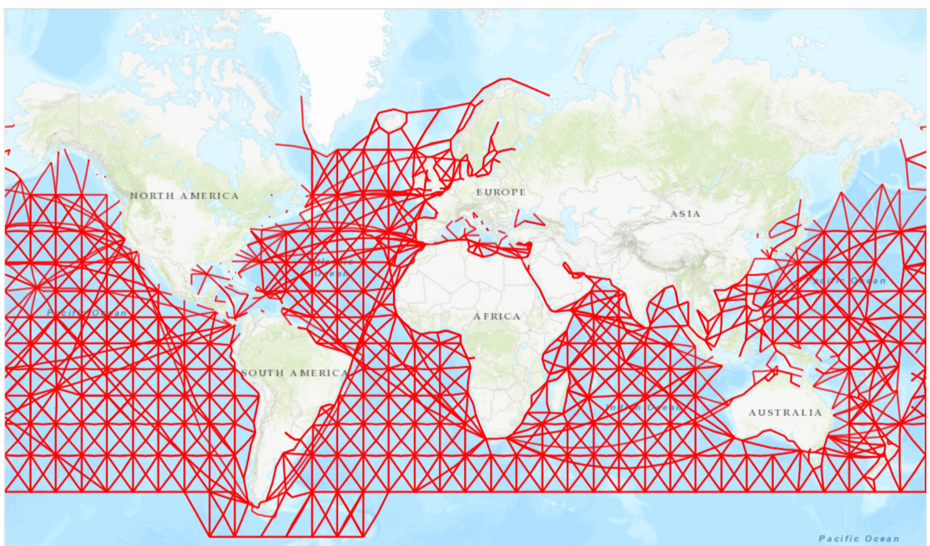
A.4 Shortest Navegable Distance Between Ports

First, we use the Global Shipping Lane Network shapefile that is published by the Oak Ridge National Labs CTA Transportation Network Group. The shapefile contains geospatial information about the network of trading lanes used by maritime shippers. Figure A1 shows a map of this network. We also use the World Port Index provided by the National Geospatial-Intelligence Agency, which provides the location of each port. We then assign each port to the closest point on the network of shipping routes. Finally, there are many possible routes that a shipper can take. In order to solve for the least costly path, we use Dijkstra's algorithm to find the shortest path between an origin and destination. We implement this algorithm using the Network Analyst tool in ArcGIS.

A.5 Combing Continguous Ports

We combine the data for Los Angeles port (Census port codes 2704 and 2791) and Long Beach port (Census port code 2709). The reason is that these two ports, while technically independent, are adjacent to each other and are considered to be the same port in practice. We similarly combine data for Newark port (Census port code 4601) and New York port (Census port code 1001).

Figure A1
Global Shipping Lane Network



B Appendix: Proofs

B.1 Proof of Proposition 1

Proof. We write the maximization problem in the third stage of a shipper operating from i' to j' that takes as given the quantity supplied by its competitors, $c^{-n}_{i'j'}$ as

$$\max_{c^n_{i'j'}} \; c^n_{i'j'} \left(T_{i'j'} - \frac{w_1}{\Phi} \right), \tag{A1}$$

where $c^n_{i'j'} + c^{-n}_{i'j'} = c_{i'j'}$. We can similarly write the maximization problem from j' to i'. Notice that we can separately solve the problem for each direction because the shippers have a constant returns to scale technology. The first order condition becomes

$$T_{i'j'} - \frac{w_1}{\Phi} + c^n_{i'j'} \frac{dT_{i'j'}}{dc^n_{i'j'}} = 0. \tag{A2}$$

We can show that this implies that

$$T_{i'j'} = \frac{\epsilon_{i'j'}}{\epsilon_{i'j'} - 1} \frac{w_1}{\Phi}, \tag{A3}$$

where

$$\epsilon_{i'j'} = -\frac{T_{i'j'}}{c^n_{i'j'}} \frac{dc^n_{i'j'}}{dT_{i'j'}} \tag{A4}$$

Notice that equation A3 is the same as equation 20.

We now derive the condition in equation 21 for $\epsilon_{i'j'}$. We use equations 7 and 10, along with the fact that $c^n_{i'j'} + c^{-n}_{i'j'} = c_{i'j'}$, to find

$$c^n_{i'j'} + c^{-n}_{i'j'} = \beta_{i'} \beta_{j'} p^{-\gamma}_{i'j'} p^{\gamma-\sigma}_{ij} P^{\sigma-1}_j \zeta \left(w_j L_j + R_j \right) \tau^{-\sigma}_{ij}. \tag{A5}$$

We differentiate this expression with respect to $c^n_{i'j'}$ and subsequently divide both sides by $c_{i'j'}$ to find

$$\frac{1}{c_{i'j'}} = -\gamma \frac{\frac{dp_{i'j'}}{dc^n_{i'j'}}}{p_{i'j'}} + (\gamma - \sigma) \frac{\frac{dp_{ij}}{dc^n_{i'j'}}}{p_{ij}} + (\sigma - 1) \frac{\frac{dP_j}{dc^n_{i'j'}}}{P_j}. \tag{A6}$$

To derive the first term on the right hand side of equation A6, we know that $p_{i'j'} = w_i/x_i + T_{i'j'}$, which implies that

$$\frac{dp_{i'j'}}{dc^n_{i'j'}} = \frac{dT_{i'j'}}{dc^n_{i'j'}}. \tag{A7}$$

For the second term, we re-write equation 11 so that shippers take as given the price index of other port-pairs

$$p_{ij} = \left(\beta_{i'} \beta_{,j'} p^{1-\gamma}_{i'j'} + \bar{p}_{i'j'} \right)^{\frac{1}{1-\gamma}}, \tag{A8}$$

where $\bar{p}_{i'j'}$ is taken as given by the shipper. We differentiate this expression with respect to $c^n_{i'j'}$ and subsequently divide both sides by p_{ij} to find

$$\frac{\frac{dp_{ij}}{dc^n_{i'j'}}}{p_{ij}} = \frac{\beta_{i'}\beta_{,j'}p_{i'j'}^{1-\gamma}}{p_{ij}^{1-\gamma}} p_{i'j'}^{-1} \frac{dT_{i'j'}}{dc^n_{i'j'}}. \tag{A9}$$

We know that

$$\frac{p_{i'j'}c_{i'j'}}{p_{ij}c_{ij}} = \frac{\beta_{i'}\beta_{,j'}p_{i'j'}^{1-\gamma}}{p_{ij}^{1-\gamma}}, \tag{A10}$$

which we substitute into equation A9 to arrive at

$$\frac{\frac{dp_{ij}}{dc^n_{i'j'}}}{p_{ij}} = \frac{p_{i'j'}c_{i'j'}}{p_{ij}c_{ij}} p_{i'j'}^{-1} \frac{dT_{i'j'}}{dc^n_{i'j'}}. \tag{A11}$$

For the third term, we differentiate equation 8 with respect to $c^n_{i'j'}$ and then divide by P_j to find

$$\frac{\frac{dP_j}{dc^n_{i'j'}}}{P_j} = \frac{\zeta\tau_{ij}^{1-\sigma}p_{ij}^{1-\sigma}}{P_j^{1-\sigma}} p_{ij}^{-1} \frac{dp_{ij}}{dc^n_{i'j'}}. \tag{A12}$$

We can show that

$$\frac{\tau_{ij}p_{ij}c_{ij}}{w_jL_j + R_j} = \frac{\zeta\tau_{ij}^{1-\sigma}p_{ij}^{1-\sigma}}{P_j^{1-\sigma}}, \tag{A13}$$

which we substitute into equation A12 to arrive at

$$\frac{\frac{dP_j}{dc^n_{i'j'}}}{P_j} = \frac{\tau_{ij}p_{ij}c_{ij}}{w_jL_j + R_j} p_{ij}^{-1} \frac{dp_{ij}}{dc^n_{i'j'}}. \tag{A14}$$

We plug in the condition in equation A11 to find

$$\frac{\frac{dP_j}{dc^n_{i'j'}}}{P_j} = \frac{\tau_{ij}p_{i'j'}c_{i'j'}}{w_jL_j + R_j} p_{i'j'}^{-1} \frac{dT_{i'j'}}{dc^n_{i'j'}}. \tag{A15}$$

We use the conditions in equations A7, A11, and A15 in equation A6 and apply symmetry across shippers to arrive at equations 21 and 22. □

B.2 Proof of Proposition 2

Proof. We use the condition in equation 20 to find that

$$\frac{w_1}{\Phi} = \frac{\epsilon_{i'j'} - 1}{\epsilon_{i'j'}} T_{i'j'} \tag{A16}$$

We substitute into the profitability condition in equation 13 the solutions from Proposition 1 to find

$$\pi = c^n_{i'j'} T_{i'j'} \left(\frac{1}{\epsilon_{i'j'}} \right) + c^n_{j'i'} T_{j'i'} \left(\frac{1}{\epsilon_{j'i'}} \right). \tag{A17}$$

We then plug in for the expression of $\epsilon_{i'j'}$ and $\epsilon_{j'i'}$ characterized in equation 21 to find

$$\pi = \frac{p_{i'j'} c^n_{i'j'}}{N_{i'j'} \kappa_{i'j'},}. \tag{A18}$$

We use the fact that $c_{i'j'} = N c^n_{i'j'}$ to find the condition in equation 23. □

C Appendix: Additional Comparisons of Model and Data

Table A1
Transportation Costs and Trade Flows (Model vs. Data)

	(1)	(2)	(3)	(4)
Trade flows	-0.0766***	-0.110***	-0.239***	-0.182***
	(0.00492)	(0.00747)	(0.0359)	(0.0644)
Distance	0.0203	-0.0201	0.0979*	0.0114
	(0.0317)	(0.0481)	(0.0536)	(0.0574)
FE	Yes	Yes	Yes	Yes
IV (Population)	No	No	Yes	Yes
Data/Model	Data	Model	Data	Model
Observations	664	556	664	556
R-Squared	0.663	0.559	0.0679	0.479

Table A1 reports the results of the estimation of equation 3 using both data and output from the model. The results using the data are identical to those in Table 6. The dependent variable is the log transportation cost from a U.S. to a foreign port. The transportation costs in the data are described in Section 3.1. The model and quantitative work are presented in Section 4-6. The independent variables are log total value of containerized trade flows and log distance, measured in nautical miles between the two ports as described in Section 3.4. The total value of containerized trade flows in the data is described in Section 3.3. The results in columns 4-6 use the population in which the foreign port is located as an IV for the value of containerized trade flows. Columns 1-2 and 4-5 report the results without fixed effects. Columns 3 and 6 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table A2
Number of Shippers and Trade Flows (Model vs. Data)

	(1)	(2)	(3)	(4)
Trade flows	0.130***	0.0708***	0.0980**	-0.0872
	(0.0110)	(0.00570)	(0.0484)	(0.0715)
Distance	0.0539	-0.119***	0.0693	-0.0501
	(0.0708)	(0.0367)	(0.0723)	(0.0637)
FE	Yes	Yes	Yes	Yes
IV (Population)	No	No	Yes	Yes
Data/Model	Data	Model	Data	Model
Observations	664	556	664	556
R-Squared	0.381	0.486	0.372	.

Table A2 reports the results of the estimation of equation 3, except that the dependent variable is the log number of shippers, using both data and output from the model. The results using the data are identical to those in Table 7. The dependent variable is the log number of shippers between a U.S.-foreign port pair. The number of shippers in the data is described in Section 3.2. The model and quantitative work are presented in Section 4-6. The independent variables are log total value of containerized trade flows and the log distance, measured in nautical miles, between the two ports described in Section 3.4. The value of total containerized trade flows in the data is described in Section 3.3. The results in columns 4-6 use the population in which the foreign port is located as an IV for the value of containerized trade flows. Columns 1-2 and 4-5 report the results without fixed effects. Columns 3 and 6 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.

Table A3

Average Shipper Size and Trade Flows (Model vs. Data)

	(1)	(2)	(3)	(4)
Trade flows	0.870***	0.929***	0.902***	1.087***
	(0.0110)	(0.00570)	(0.0484)	(0.0715)
Distance	-0.0539	0.119***	-0.0693	0.0501
	(0.0708)	(0.0367)	(0.0723)	(0.0637)
FE	Yes	Yes	Yes	Yes
IV (Population)	No	No	Yes	Yes
Data/Model	Data	Model	Data	Model
Observations	664	556	664	556
R-Squared	0.953	0.988	0.952	0.969

Table A3 reports the results of the estimation of equation 3, except that the dependent variable is log average shipper size using both data and output from the model. The results using the data are identical to those in Table 8. To construct average shipper size, we divide the total value of containerized trade flows in U.S. dollars between two ports by the number of shippers. In the data, the number of shippers is described in Section 3.2 and the total value of containerized trade flows in U.S. dollars is described in Section 3.3. The model and quantitative work are presented in Section 4-6. The independent variables are the total value of containerized trade flows and the log distance, measured in nautical miles, between the two ports described in Section 3.4. The results in columns 4-6 use the population in which the foreign port is located as an IV for the value of containerized trade flows. Columns 1-2 and 4-5 report the results without fixed effects. Columns 3 and 6 include an origin port-destination region fixed effect. Regional fixed effects are defined by the World Bank and are described in footnote 12. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively.