Entrepôts: Hubs, Scale, and Trade Costs

Sharat Ganapati  Woan Foong Wong  Oren Ziv
Georgetown University  University of Oregon  Michigan State University

December 2020

Abstract

Entrepôts are hubs that facilitate trade between various origins and destinations. We study these entrepôts, the network they form, and their impact on international trade. We document novel facts about the global trading network: the trade network is a hub-and-spoke system with 80% of trade shipped indirectly via these hubs using larger ships, suggesting that, by concentrating shipments, entrepôts reduce trade costs through scale economies. We estimate direct and indirect trade costs by building a model of endogenous entrepôt formation incorporating route choice by producers into a Ricardian setting and develop a geography-based instrument to estimate a leg-level scale elasticity. Simulating the counterfactual opening of the Arctic Passage and Brexit, we find that network spillovers—impacts on countries unaffected by direct changes—double baseline welfare gains, with scale economies in shipping further tripling them. Counterfactual infrastructure improvements show that entrepôts are globally pivotal nodes and concentrate network spillovers regionally.

Keywords: trade costs, scale, hubs, transport costs, transportation networks, international trade, shipping

JEL Classification: F10, F13, F14

*Contact: sganapati@georgetown.edu, wfwong@uoregon.edu, and orenziv@msu.edu. We thank Treb Allen, Costas Arkolakis, Panle Jia Barwick, Bruce Blonigen, Johannes Boehm, Mark Colas, Kerem Coşar, Anca Cristea, Meredith Crowley, Carsten Eckel, Stefania Garetto, Matthew Grant, Reka Juhasz, Steve Matusz, Eara Oberfield, Nina Pavcnik, Andrés Rodríguez-Clare, Robert Staiger, Meredith Stark, and seminar participants at Stanford University, University of Michigan, University of Virginia, Boston College, Georgetown University, the Federal Reserve Board, Ludwig-Maximilians-Universität Munich, Indiana University, Syracuse University, University of Cambridge, University of Mannheim, and University of Warwick as well as participants at the NBER Conference on Cities, Labor Markets, and the Global Economy, 2019 Mid-Atlantic International Trade Workshop (Duke), 2020 CeMENT Mentoring Workshop, 20th Annual Nordic International Trade Seminars, and other conferences for helpful comments. We apologize for any cancelled presentations due to COVID-19. Andrew Castro, Ray McCormick, Giacomo Romanini, and Philip Valtadoros provided outstanding research assistance.
1 Introduction

Exchanging goods over borders involves more than production and consumption: shipping, transshipping, and distribution can include multiple agents and additional countries beyond producers and consumers. These activities are concentrated at entrepôts, trading hubs which goods travel through—from other origins and bound for other destinations. The notion that entrepôts are integral to the trade network and are engines of national and regional growth is a powerful narrative with a long history.¹

This paper studies entrepôts, the trade network they form, and their impact on international trade. We seek to answer the following questions: (1) What is the network structure underpinning international trade and what role do entrepôts play in this network? (2) What are the costs of direct and indirect trade in this global network? and (3) What are the implications of this network for international trade and welfare?

Our first contribution is to characterize the global containerized trading network by constructing two new datasets that jointly map US-bound shipments’ individual journeys through the global trade network at the shipping container level. Previous work observed origin-destination trade alone or just aggregated ship movements, and was unable to observe indirect trade, which we define as journeys that make stops with the shipment either on-board or transshipped—transferred onto a ship—at additional countries (third-party countries). Our data allows us to observe shipments’ origin and destination, the ships transporting them, as well as their locations and stops, granting us the first comprehensive look at how goods move through the global trading network.

We document three stylized facts that characterize this network. First, the majority of trade—80%—is shipped indirectly. The average shipment stops at two additional countries before its destination, adding 30% in distance travelled.² Figure I plots the average number of stops made by containers by origin. On average, most countries access the US indirectly. We further show that indirect trade significantly adds to the shipping distance of goods and their shipment time. Second, this indirectness is incredibly

¹Governments and local port authorities invest billions of dollars with the specific aim of becoming or maintaining their role as entrepôts. Saudi Arabia has implemented a $7 billion project to expand the container capacity to “be a major east-west marine transshipment location.”(Financial Times, 2015). India spent $4 billion to rival Chinese facilities (Reuters, 2016). Established entrepôt Singapore invested $1.1 billion to boost its capacity to “stay ahead of the curve as a world-class hub port” (Port Technology, 2018) following a $3 billion project to construct an automated container yard (Ship and Bunker, 2012).

²The majority of trade is also transshipped via a third-party country before its destination.
Notes: For each origin country, we calculate the average number of third-party country stops weighted by container volume. The destination country (US) is excluded (in white). Landlocked countries are also excluded (in white), since they would mechanically need to stop at a coastal country. 34 of the shipment origin countries are landlocked accounting for 1.6 percent of total TEUs. The missing remaining countries are excluded either due to lack of overall trade with the US (e.g. Somalia) or due to the merge process (e.g. Namibia). We separate out trade between China, Hong Kong, Macau, and Taiwan.

concentrated, with a large number of shipments channelled through a small number of entrepôts. This establishes that international trade takes place over a hub-and-spoke network. Third, bigger sized ships, which have lower unit costs of shipping, are more likely to serve entrepôt and indirect routes. In sum, indirectness is ubiquitous and concentrated at entrepôts, and the trade network is a hub-and-spoke system—large ships connect global hubs and smaller ships service small local routes. These facts suggest that by centralizing shipments, entrepôts reduce shipping costs through scale economies.\(^3\)

Our second contribution is to estimate link-level and origin-destination trade costs taking into account direct and indirect trade through the network. To achieve this, we build a general equilibrium model of trade with entrepôts and endogenous trade costs. Producers choose shipping routes and compete for consumers in destination countries in a generalized Ricardian setting, flexibly accommodating input-output linkages. Low-cost routes can involve shipments through third-party countries, and entrepôts endogenously arise at locations through which shipping costs are lowest. Crucially, we allow for both scale economies and dis-economies to govern shipping costs on network links as traffic grows.

Using global data on shipment flows (not just US data), we estimate these trade costs for each leg of the network, generating a new set of origin-destination trade costs that is

---

\(^3\)This network structure is not specific to containerized shipping but is also prevalent in air transport and freight services (Rodrique, Comtois and Slack, 2013).
consistent with the global trade network. An advantage of our model is that we need to make very few assumptions on the production and consumption settings; we recover a trade cost matrix that best rationalizes the observed link-level traffic given the observed origin-destination-level trade flows. An important contribution is that we establish the validity of both our estimates and the modeling approach by finding a tight match between our estimated trade costs and external freight rate data, as well as between our predicted network flows and data on container shipment journeys.

Our estimation finds the presence of bilateral scale economies—the causal effect of increasing shipping volumes on decreasing trade cost—using an instrumental variable approach that leverages the geography of the trade network. Embedded in our model is the intuition that some legs have inherently higher traffic (higher demand) because they are geographically closer to the shortest path between large origins and destinations. We use this variation to construct an instrument for shipping quantities: for each leg, we compute the distance to and from the leg relative to the shortest distance between each origin and destination, recovering a weighted average of each leg’s proximity to global trade. We find that a 1% increase in traffic on a given leg reduces costs on the same leg by 0.06%.

Our third contribution is to quantify the impact of the trade network on global trade and welfare, highlighting how changes at nodes operate through the network structure, entrepôts, and scale economies to create widespread impacts. We simulate three sets of counterfactuals using our general equilibrium model. Our first set of counterfactuals considers the ramifications of worsening trade relations between the United Kingdom (UK) and its trading partners—Brexit. Smaller countries like Ireland and Iceland that use the UK as an entrepôt are disproportionately hurt when our analysis accounts for the interaction of network trade costs and scale economies, illustrating how such interactions can lead to different distributional outcomes even when initial changes are unrelated to transport. Our second set of counterfactuals considers the effects of opening up the Arctic Ocean to regular year-round shipping. We find that the network structure of trade distributes gains beyond directly impacted countries with pre-existing shipping routes and doubles the baseline effects. In both counterfactuals, global welfare impacts are further

---

4Our estimation generates four sets of data products: leg-level transport costs, network-consistent origin destination trade costs, country-level networked market access measures, and measures of indirect trade. All four are available for download on our websites.

5The emergence of entrepôts as hubs in geographically advantageous locations is consistent with the findings of Barjamovic et al. (2019).
tripled by the feedback loop imposed by scale economies.

Our last set of counterfactuals quantifies the benefits of transport infrastructure improvements for each country in our sample on both global and local welfare as well as trade levels. We show that (1) entrepôts are pivotal nodes: the global welfare impacts of infrastructure investment are disproportionately high at entrepôts, (2) entrepôts differentially concentrate the benefits of infrastructure improvements regionally compared to non-entrepôts, with scale economies amplifying that concentration, and (3) scale economies allow entrepôts to take better advantage of trade cost reductions elsewhere. These results, that scale economies in transportation concentrate gains locally at and around hubs, highlight scale economies in transportation as a source of agglomeration.

This paper characterizes the nature of the global container shipping network and its implication for international trade. It contributes to the literatures on the presence of networks in trade, on endogenous trade costs, and on trade and transportation technology. Our main contributions are to a growing quantitative literature investigating the role of trade networks (Fajgelbaum and Schaal, 2017; Redding and Turner, 2015; Allen and Arkolakis, 2019). We provide the first and systematic documentation of indirect trade through the containerized shipping network.6 We extend the Allen and Arkolakis (2019) Armington framework where route cost shocks are born by consumers to a general Ricardian setting, where traffic volumes reflect both route choice and head-to-head competition on prices at destinations and show how to estimate the model in a multi-industry setting with missing traffic flows. We affirm the validity of the Allen and Arkolakis (2019) approach by reporting a tight match between our model predictions and data external to our estimation. Heiland et al. (2019) study the impact of the Panama Canal expansion on global containership movements. Relative to this paper, we observe indirect trade and entrepot activities as well as quantify the welfare and trade impacts of the trade network using a general equilibrium model with endogeneous trade costs.7

We also contribute to a growing literature on endogenous transport costs.8 Our trade costs are endogenously determined in equilibrium with trade flows as part of the transportation network. We model transport costs as part of a global network of container

---

6We provide our estimated data on indirect flows which can be used for future empirical work. Previous work have either imputed indirect trade or just used port of call data alone (Wang and Wang, 2011; Kojaku et al., 2019; Lazarou, 2016). They cannot directly observe indirect trade since they do not have information on the loading or unloading of shipments.

7They infer travel routes for country pairs that might use the Canal and find that country pairs whose fastest connection passed through the Panama Canal prior to the expansion traded 9-10% more.

8See Hummels (2007) and Limao and Venables (2001) for reviews of this literature.

4
shipping routes, a setting which accounts for two-thirds of annual trade moved by sea (World Shipping Council). Brancaccio, Kalouptsidi and Papageorgiou (2017) estimate endogenous trade costs arising from search frictions for dry bulk ships carrying homogeneous commodities, where all trade is direct.\(^9\) Using our general equilibrium spatial trade framework, our counterfactuals assess the welfare impacts of these kinds of endogenous transport costs and show how they propagate via the network and through entrepôts.\(^10\)

Finally, we contribute to the literature on transportation technology and trade. Several papers investigate the effects of containerization (Coşar and Demir, 2018; Bernhofen, El-Sahli and Kneller, 2016; Wong, 2020). Ducruet et al. (2019), a complementary study, investigates the impact of infrastructure investment—new port technologies from containerization in the 1970s—on urban and national outcomes. We document the network effects of container shipping on global trade.

One important aspect of transportation technology in our model is the scale economy in shipping, the magnitude of which is in line with findings from the literature on economies of scale in trade.\(^11\) Bartelme et al. (2019) and Lashkaripour and Lugovskyy (2019) both consider the trade consequences of production scale economies, while we consider scale in transportation. Our paper shows how such scale economies interact with the global trade network to concentrate economic activity. In this respect, we are also related to a literature in economic geography which considers the role of localized scale economies in the emergence of agglomerations (Allen and Arkolakis, 2014; Allen and Donaldson, 2018). Scale economies typically generate agglomerations by acting on the volume of economic activity at locations. We show that scale economies can concentrate activity at hubs by acting on transportation costs over a network.

\(^9\)Like taxis, dry bulk ships depart from destinations without cargo and have to search for it. Indirect trade, and the network structure giving rise to trade spillovers due to network linkages, are absent from this setting. Containerships, like buses, travel on fixed schedules between many locations.

\(^10\)Hummels, Lugovskyy and Skiba (2009) and Asturias (2020) study transport costs in the context of market power. While container shipping firms may hold market power, we generalize away from the profits of the shipping companies. We follow Sutton (1991) and allow larger markets to induce entry and competition to lower prices, but do not directly measure market structure. Toy models allowing for monopolistic competition do fit within our framework, however we lack the data to fully incorporate them. We leave the study of how market power works through the hub-and-spoke network for future study.

\(^11\)See especially Alder (2015); Holmes and Singer (2018); Anderson, Vesselovsky and Yotov (2016); Asturias (2020); Skiba (Forthcoming).
2 Data

We compile and combine two proprietary data sets in this project: global ports of call data for containerships, which allows us to reconstruct the routes taken by specific ships, and US bill of lading data for containerized imports, which gives us shipment-level data on United States imports. Independently, these datasets allow us to partially describe the global shipping network. By merging them, we reconstruct the journey of shipments entering the United States, from their origin to their US port of entry. To our knowledge, we provide the most comprehensive reconstruction of the global shipping network and routes undertaken by individual shipments into the US.\textsuperscript{12}

**Figure II:** Map of Global Port of Call Network

![Map of Global Port of Call Network](image)

Notes: Each dot represents a port (total of 1,203 ports). Each line represents a journey between port pairs undertaken by a containership (total of 4,986 ships).

Our proprietary ports of call data from Astra Paging captures vessel movements using ship Automatic Identification System (AIS) transponders.\textsuperscript{13} For each vessel, this captures identifying information, time-stamped ports of call, capacity, and height in the water before and after stopping at each port. The latter two pieces of information jointly indicates the vessel’s load at these ports, allowing us to observe volumes shipped between port pairs.

Our sample covers 4,986 unique container ships with a combined capacity of 18.1 million twenty-foot equivalent shipping units (TEUs)—over 90% of the global container shipping fleet—making 429,868 calls at 1,203 ports from April to October 2014. Figure II shows the coverage of the shipping network in our port of call data. Each line represents

\textsuperscript{12}Data Appendix A.1 explains both data sets and their merge procedure in detail.

\textsuperscript{13}Port receivers collect and share AIS transponder information (including ship name, speed, height in water, latitude and longitude). Using geographic AIS variables, we track global port entry and exit data.
a containership journey. The ports of Singapore, Rotterdam, Tawfiq and Balboa (due to the Suez and Panama Canals respectively)—well-known entrepôts—are particularly connected. We use this global data along with CEPII global trade data when estimating our model in Section 5. However note that with this port of call data alone, shipment journeys along this network remain unobserved.

To remedy this, we zoom in on US trade. We merge our port of call AIS data with US bills of lading data, which captures shipment-level information for all containerized imports. We observe the shipment’s origin country, the port where they are loaded onto containerships (port of lading), and the US port where they are unloaded (port of unlading). In addition, we know the name and identification number of the containership which transported the shipment as well as the shipment’s weight, number of containers (TEUs), and product information. Over the same six months period, we see a total of 14.8 million TEUs weighting 106 million tons were imported into the US from 227 origin countries and laded in 144 countries.

Using details on the containership, unlading port and time of arrival, we match the bills of lading to the journeys of specific containerships, then use the ports of call between lading and unlading to reconstruct each shipment’s path from its foreign origin to US destination. Over 90% of containerized TEUs entering the US can be matched to routes using this method (Appendix Figure A.1 visualizes this merge).\textsuperscript{14} While the shipments’ exact journey between origin and the first stop (port of lading) remain unobserved, this initial portion could only either take place overland (by trucks or rail) or by sea on another ship because it is containerized. This, however, will lead us to under-count the overall level of indirectness. We deal with unobserved transit in Section 5.

3 Stylized Facts

We use our data to explore the nature of the international trade network and the routes taken by goods entering the US along that network. We find indirectness is both ubiquitous and costly—increasing both time and distance travelled for these shipments. The global trade network is a hub and spoke system, concentrating a large number of shipments through a small number of entrepôts, and allowing for the use of supersized cargo ships with lower unit costs.

\textsuperscript{14}See Appendix A.1 for further details on each of these datasets as well as the merge process.
3.1 The Majority of Trade is Indirect

Panel (A) in Figure III reports the distribution of the number of observed country stops made by each shipment, weighted by TEU containers. Only about 20% of containers are exported to the US directly from their origin countries—making no stops in third-party-countries. The average container entering the US stops at around 2 third-party-countries.\(^{15}\) This is also true at the country level: the majority of US trading partners export to it indirectly. This can be roughly gleaned from Figure I. Only shipments from 9 countries typically enter the US directly.\(^{16}\)

**Figure III:** Indirect Trade Distributions, by Container and Country

(A) Country Stops per Container  
(B) US Import Transshipment Share

Notes: Panel (A) shows the distribution of containers by the number of unique third-party countries the containers visited. Panel (B) show the distribution of countries, by the share of country’s shipments that are transshipped en route to the United States. Both plots are at the shipment level and weighted by the aggregate exported containers (TEU). Shipments from landlocked countries are excluded.

Panel (B) in Figure III focuses on an alternative aspect of indirectness, \textit{transshipment}, when shipments are loaded onto containerships in a third-party-country (i.e. not in their countries of origin). The results here echo our findings in the paragraph above: the average shipment from a majority of US trading partners is transshipped in a third-party country—over 60% of US trading partners transship more than 90% of their US-bound goods.\(^{17}\) The high degree of variation in connectivity evident already in Figure

\(^{15}\)Mean of 1.5 and s.d. of 1.3. The average number of port stops is higher (Figure A.3, mean of 4.6 and standard deviation of 3.5). This result is robust for shipment weight and value (Figure A.4). Multiple stops at the same third-party country are not counted. Shipments from landlocked countries are excluded in this analysis.

\(^{16}\)These countries are Canada, Mexico, Panama, Japan, South Korea, Spain, Portugal, South Africa, and New Zealand. China is not included as Hong Kong, Taiwan, and Macau are considered separate countries.

\(^{17}\)Examples include Denmark, Bangladesh, Cambodia, and Ecuador. See Figure A.5 for a map of the
I is explored in Appendix B.3, showing that that variation is reasonably explained by traditional gravity variables, and further explores variation in routes from unique origins into the US.\footnote{The existence of within-origin route variation is an important assumption in our model and is used in our validity checks.}

**Indirect trade increases shipping distances and time.** Are the additional country stops simply incidental stops along the way, or do they constitute a trip that is meaningfully distinct from what a “direct” path would look like? One possibility is that the observed indirectness is optimal but only incidental—perhaps additional stops only have small effects on cost, and therefore may be optimal even if the benefit of indirectness is small. However, the significant additional distance and time incurred by indirect travel, documented here, implies this is unlikely to be the case.

On average, the actual traveled distance between a shipment’s origin and its US destination is 31\% more than its direct ocean distance (Panel (A) in Figure IV). Panel (B) shows the actual traveled distance between a shipment’s last lading location and its final destination. Here the remaining gap is still substantial at 14\%. Table A.1 further evaluates the relationship between indirectness and journey length, finding that doubling the number of stops adds 10\% to distance travelled and 33\% to time travelled, even after controlling for direct journey length or origin-by-destination fixed effects. These distance and time costs do not include pecuniary costs of transshipment. We conclude that indirectness is meaningful in the sense that it is costly. The fact that this organizational structure remains optimal implies that it carries with it a cost reduction over and above these costs. From these results, we can summarize our first stylized fact:

**Stylized Fact 1.** *The majority of containerized trade into the US is indirect and results in a significant increase in shipping distance and time.*

### 3.2 Indirect Trade Is Routed Through Entrepôts

When shipments stop in third-countries, how are they routed? We show that the stops along indirect shipping routes are not arbitrarily distributed throughout the world but instead channelled through well-known entrepôts, which disproportionately service shipments originating in other countries.

Panel (A) of Figure V scatters each country’s share of total third-country stops against its share of total US trade. Some locations are both popular stopping points but also...
Figure IV: Difference Between Traveled Distance and Direct Distance

(A) Shipment Origin to Destination
(B) Place of Lading (Stop 1) to Destination

Notes: These figures show only indirect shipments, with different direct and travelled distances. Dots are shipments, shaded by TEU. Figure (A) compares the direct shipping distance from the country of origin to the US to the actual route travelled. Figure (B) compares the direct shipping distance from the place a container was last loaded on a ship before arrival to the US to the actual route travelled. Sea distances for observed and direct routes are calculated using Dijkstra’s algorithm. The local linear fit line is a locally weighted regression of the observed on direct pair-wise distance.

major countries of origin for goods. A few key countries, especially Korea, Singapore, Panama, and Egypt, are not only high on the Y-axis, denoting they are important third-countries, but they are above the 45° line, indicated that they disproportionately participate in trade as a third-country in US-bound shipments. Panel (B) plots each country’s total volume as a first stop (lading volume) against their origin volume. A similar set of countries lies above the 45° line.19

This leads us to our definition of entrepôt activity:

\[
Entrepôt_i = \frac{\text{Stops}_i}{\sum_{i' \in I} \text{Stops}_{i'}} - \frac{\text{Trade}_i}{\sum_{i' \in I} \text{Trade}_{i'}}
\]  

(1)

where country \(i\)’s entrepôt usage is the difference between its share of all third-country stops minus its share of global trade. Countries long-associated with entrepôt activity score highly. Data in Figure V are US-centric; Figure A.7 repeats the exercise using estimation-consistent traffic and trade. The top 15 measures are our set of entrepôts.20

Third-country country stops (the Y-axes) are significantly more concentrated than trade (the X-axis) at the 99-50, 95-50 and 90-50 levels. Table A.2 reports these concentration ratios for trade, transshipment, and third-country stops, which are high by most

19For top 10 countries by country stops and lading volumes, Figure A.6 tabulates the percent of all goods entering the US stopping in that country, broken into goods originated there and elsewhere.
20Our set of global entrepôt are: Egypt, Singapore, Netherlands, Hong Kong, Belgium, Taiwan, Spain, Saudi Arabia, South Korea, the United Arab Emirates, Morocco, Panama, Malta, Portugal, and the United Kingdom.
Figure V: Direct vs Indirect Shipments

(A) Percent Originated vs Transit Volume

(B) Percent Originated vs Laded

Notes: Figure (A) compares for each country the share of shipments to the United States that originated in a country (x-axis) versus the share that passed through that country, but did not originate there (y-axis). Figure (B) compares the share of shipments to the United States that originated in a country (x-axis) to the share of shipments that were transshipped (last loaded onto a ship) in that country (y-axis). For scale, China is omitted. Shares of shipments are weighted by TEU.

These relationships can be summarized in our second stylized fact:

Stylized Fact 2. Indirect shipping routes are concentrated through well-known entrepôts. International trade occurs over a hub-and-spoke network.

3.3 Indirect Trade and Larger Ships

By revealed preference, shipping through entrepôt appears to generate cost reductions over and above the costs incurred by travelling indirectly. A number of mechanisms may well account for the unobserved cost reduction at entrepôts. Here, we focus on the relationship between indirect shipping and ship size, which we observe and has been found to be inversely related to shipping costs.

Goods travelling through entrepôts ship on larger ships. The regression line in Figure VI shows an unsurprising positive country-level relationship between the volume of US imports and the average size of the incoming ships. The skew of larger circles—representative of countries’ lading share—relative to the regression line implies that countries with larger lading volumes ship on disproportionately larger ships. The dark points

---

21The 99-50 ratio is 400 for third-country stops and transshipment, 96 for trade, while same ratio in employment in the highly concentrated IT sector across US cities is 300 (Moretti, 2019).
22High-traffic routes are served by many carriers, using ships capable of carrying 25,000 containers with automated lading and unlading technologies. Internal and external scale economies in shipping and competition among shippers could all generate a negative relationship between volume and costs, as could factors such as port infrastructure.
23Cullinane and Khanna (2000) find that shipping costs decrease as ship size increases.
indicate countries where all shipments coming into the US are always laded elsewhere. These countries, who disproportionately use entrepôts, are outliers, having shipments arriving on much larger ships than expected given their trading volumes.

**Figure VI: Link Between US Trade and Ship Size**

Notes: Each circle and dot represents an exporting country to the United States. The y-axis shows the total exports from that country to the United States. The x-axis shows the average size of a ship a good from that exporter arrives at the United States. For countries with positive lading volumes, the size of the circle represents the total volume of exports that was last loaded onto a ship in that country. Countries without direct shipments to the United States are denoted with solid black dots.

We pair this visual analysis with Table I, which displays shipment-level regressions. Column 1 regresses, for our sample of shipments, the log of ship size against the log of total origin country volumes shipped (TEUs), confirming a positive relationship. Column 2 adds the log of quantity laded at each shipment’s port of lading. Both coefficients are positive but the coefficient on origin volume is almost halved (0.084 in Column 1 compared to 0.043 in Column 2), indicating that much of the correlation between origin volume and ship size acts through the size of the lading port. Column 3 fully interacts variables in Column 2 with an indicator variable for shipments that are laded in their origin countries. As suggested by the figure, for shipments whose origin country differs from lading country—an indicator value of 0—the correlation between ship size and lading volume is considerably higher (0.130), and shipments’ ship sizes are not strongly correlated with origin country volumes when they lade in third countries (0.009).

Finally, stopping at larger ports matters, even when goods remain on board: goods lading at smaller transshipment points that travel along major routes are also on larger ships. Column 4 of Table I regresses shipments’ log ship size against the log volume laded at their port of lading and the log volume laded at the largest port at which we
Table I: Determinants of Ship Size

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln Volume at Origin</td>
<td>0.0843</td>
<td>0.0432</td>
<td>0.00925</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0163)</td>
<td>(0.0179)</td>
<td>(0.0121)</td>
<td></td>
</tr>
<tr>
<td>ln Volume at Lading</td>
<td>0.0803</td>
<td>0.127</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0202)</td>
<td>(0.0230)</td>
<td>(0.0182)</td>
<td></td>
</tr>
<tr>
<td>1 × (Lading is Origin)</td>
<td>-0.0220</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.300)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 × ln Volume at Lading</td>
<td>-0.0937</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0295)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 × ln Volume at Origin</td>
<td>0.0861</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0220)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln Largest Port Stop</td>
<td>0.121</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0250)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>215,656</td>
<td>215,656</td>
<td>215,656</td>
<td>215,656</td>
</tr>
<tr>
<td>R²</td>
<td>.124</td>
<td>.174</td>
<td>.199</td>
<td>.21</td>
</tr>
<tr>
<td>F-stat</td>
<td>26.82</td>
<td>14.66</td>
<td>13.51</td>
<td>26.73</td>
</tr>
</tbody>
</table>

Notes: Observations are at the shipment level, weighted by TEU, representing all matched imported containers to the United States. ln Ship Size is the natural log of maximum ship capacity in TEU. ln Volume at Origin is the natural log of the sum of all shipments’ TEU by shipment origin country. ln Volume at Lading is the sum of all shipments’ TEU by shipment lading country. The indicator takes a value of 1 if the shipment is laded at the country of origin. ln Largest Port Stop is the maximum of the natural log of the volume of lading at all ports visited between the port of lading and unlading. Standard errors are clustered two ways by lading and destination ports.

We observe the shipment making a port call. The effect of the max-port-size variable is large, positive, and overall stronger than the effect of lading port volumes alone. Additional stops that move through entrepôts allow shipments laded in smaller ports to travel on larger ships. Indirectness facilitates larger ship size beyond transshipment alone.

Results in this section are robust to adding origin fixed effects. We summarize our findings in the following stylized fact:

**Stylized Fact 3.** *Goods from and through entrepôts are loaded onto larger ships.*

These facts outline an inherent trade-off: indirectness increases distance and time costs of trade, but the resulting concentration appears to lower costs. The level of indirectness and concentration we have documented in the data are shaped by this trade-off, and the goal of our empirical estimation is to understand the forces underlying it.

4 Theoretical Framework

We present a model of global trade where shipments are sent indirectly through an endogenously formed transport network. We embed the Allen and Arkolakis (2019) route selection model in a generalized Eaton and Kortum (2002) framework where production
technologies in each industry and country are non-stochastic, but idiosyncratic variation in products’ optimal route generates random variation in product-origin pair prices.

Entrepôts emerge as ports through which goods flow but which are neither the goods’ origin nor their destination. Throughout, we maintain a production and consumption setting that is as general as possible, allowing for any number of goods, industries, and input-output linkages. This model is agnostic to scale economies or dis-economies in transportation costs, which could work to either amplify or attenuate shipments through entrepôts. Restrictions on route cost heterogeneity generate moment conditions that can be matched to the data to yield estimates of leg-specific shipping costs.

4.1 Setup

Consumption and Production  In each country $j$, consumers consume goods $\omega_n \in \Omega_n$ from each of $N$ industries $n$ according to some function $U_j = U_j(C_j)$, where $U_j$ is a continuous, twice differentiable function and $C_j$ is a matrix of quantities of an arbitrarily large number of goods $\omega_n$ in industry $n \in N$ in country $j$. Within each industry and product category, goods are homogeneous and normal. Goods are produced using a variety of traded and non-traded inputs including labor, capital, and traded and non-traded varieties from any industry. The production technology for good $\omega$ is common for all goods in the same industry $n$, and includes a vector of factor inputs $L$, as well as inputs of other goods. Production functions can vary across industries and countries. Cost minimization leads competitive firms in each country within an industry to have identical production costs. Marginal cost of a good $\omega$ is

$$c_{in} \equiv c_{in}(z_{in}, W_i, P_i),$$

where $P_i$ is the matrix of prices of all goods $\omega$ in industries $n$ in country $i$ and $W_i$ is the vector of factor prices in country $i$. Because producers in the same industry and country share the same input prices and production function, costs are shared within country-industries. These costs correspond to the classic Ricardian comparative advantage.

---

24 We allow for the utility function to vary across destinations, and the number of goods in each industry need not be a continuum but can be.

25 The model and empirics can accommodate arbitrarily fine industry classifications in order to ensure this assumption holds.

26 The production function is given by $q_{in}(\omega) = f_{in}(z_{in}, L_{in}, Q_{in})$ where $f_{in}$ is a continuous and twice differentiable country-industry-specific production function, $z_{in}$ is a production technology common to industry $n$ and country $i$, $L_{in}$ is a vector of non-tradable factor inputs, and $Q_{in}$ is a country-industry specific matrix of inputs of other goods $\omega$ from all industries. All inputs are treated as homogeneous.
**Pricing** To sell goods abroad at any destination \( j \in J \), a firm producing product \( \omega \) in industry \( n \) must pay non-transport trade costs \( \kappa_{ijn} \) and iceberg transport costs \( \tau_{ijnr}(\omega) \) after optimally choosing the route \( r \) between \( i \) and \( j \) to minimize the shipping costs incurred. Competitive firms selling from \( i \) to \( j \) price their goods at marginal cost. Observed prices for these products at \( j \) are

\[
p_{ijn}(\omega) = c_{in} \kappa_{ijn} \tau_{ijnr}(\omega),
\]

where purchasers of good \( \omega \) in industry \( n \) at \( j \) source the lowest cost supplier globally.

**Shipping** Producers seek to minimize shipping costs, choosing the lowest cost shipping route available. Shipping route \( r \) is comprised of \( K_r \) legs of a journey with \( K_r - 1 \) stops along the way between the origin, \( i \), (or \( k = 1 \)) and destination \( j \), (or \( k = K_r \)).

Following Allen and Arkolakis (2019), we assume that moving from stop to stop involves iceberg transport costs as well as product- and route-specific idiosyncratic cost shocks \( \epsilon_{ijnr}(\omega) \).\(^{27}\) This shock is drawn from the Fréchet distribution such that \( F_{ijn}(\epsilon) \), the cumulative distribution function of the idiosyncratic draws is the following:\(^{28}\)

\[
F_{ijn}(\epsilon) \equiv \Pr\{\epsilon_{ijnr}(\omega) \leq \epsilon\} = \exp \{-\epsilon^{-\theta}\},
\]

where shape parameter \( \theta > 0 \) captures the randomness or dispersion in the choice of routes from \( i \) to \( j \).\(^{29}\) Higher \( \epsilon_{ijnr}(\omega) \) draws mean industry \( n \) has lower costs for route \( r \).

Accordingly, product \( \omega \)'s shipping cost along route \( r \) from country \( i \) to country \( j \) is:

\[
\tau_{ijnr}(\omega) = \frac{1}{\epsilon_{ijnr}(\omega)} \prod_{k=1}^{K_r} t_{k_{r-1},k_r} = \frac{1}{\epsilon_{ijnr}(\omega)} \tilde{\tau}_{ijr}, \tag{2}
\]

where \( \tilde{\tau}_{ijr} \) is the product of all leg-specific costs \( t_{k_{r-1},k_r} \) and is common to all products taking the same route \( r \). While transport costs are usually a function of distance, we place no structure on these leg-level costs, allowing them to be a flexible function of exogenous and endogenous variables: \( t = f(X_{exog}, X_{end}) \). Each product’s shipping cost from \( i \) to \( j \) is the minimum transport cost route over a set of all other routes for origin \( i \), destination \( j \), and product \( \omega \) in industry \( n \).\(^{30}\)

---

\(^{27}\)Because of the max-stable property of the Frechet distribution, an isomorphic specification would have firm-specific cost shocks with a finite mass of potential competitive firms in each country. This would affect the interpretation of the source of idiosyncratic variation (firm variation or product variation) and of shape parameter \( \theta \).

\(^{28}\)This distribution is the same for each product across industries so product-industry subscripts \( n \) is dropped.

\(^{29}\)This dispersion assumption is reflected in our microdata (Panel (B) in Figure A.9, Appendix B.3) Almost 70 percent of origin countries have fairly low concentration of routes (HHI less than 1500).

\(^{30}\)The price of a product \( \omega \) in industry \( n \) from \( i \) to \( j \) conditional on route \( r \) is \( p_{ijnr}(\omega) = c_{in} \kappa_{ijn} \tau_{ijnr}(\omega) \).
The multiplicative functional form for trade costs in equation (2) is an assumption which allows for an analytical solution to the routing problem. We subsequently establish a tight fit between our estimates and two separate sets of external data in Section 7, which helps alleviate potential misspecification concerns.

This structure is consistent with a host of mechanisms, including but not limited to port-level effects and leg-level scale economies.\(^\text{31}\) In terms of market power, we do not directly model the decision of shipping firms, but rather consider an overall industry equilibrium within a Sutton (1991) framework, where larger markets induce more entrants and lower marginal costs, with profits being absorbed by fixed costs.\(^\text{32}\) Differences between these mechanisms will not impact the model estimation but will manifest in the interpretation of scale economies and for counterfactual predictions.

### 4.2 Equilibrium

We find expressions for two observables: (1) the equilibrium mass of products that will be shipped from any origin to any destination through a specific leg and (2) the total volume of trade between any country pair. We defer imposing market clearing conditions and closing the model until Section 8, which produce gravity and wage and input price expressions.

**Route volume**  Firms from origin \(i\) select the lowest-cost route before consumers in \(j\) select the lowest-cost intermediate good supplier across all the origins countries. We observe \(\omega\) being shipped on route \(r\) from \(i\) to \(j\) only if the final price of \(\omega\), which includes both the marginal cost of production and shipping cost on route \(r\) from \(i\) to \(j\) \((p_{ijnr}(\omega))\), is lower than all other prices of good \(\omega\) from all other origin country-route combinations.

We then consider the probability that a given country and route \(r'\) will be selected as the lowest cost route-supplier combination for good \(\omega\) conditional on price \(p\):

\[
G_{jn\omega}(p) \equiv \Pr \left\{ \min_{i \in I, r \in R_{ij}} p_{ijnr}(\omega) > p \right\} = 1 - \exp \left\{ -p^\theta \cdot \sum_{i} \left( c_{in} \cdot \kappa_{ijn} \right)^{-\theta} \cdot \sum_{r \in R_{ij}} \tilde{\tau}_{ijr}^{-\theta} \right\}.
\]

We can define the joint probability that a route \(r\) is the lowest-cost route from \(i\) to \(j\) for

\(^{31}\)It also allows for spatial correlation in link costs, say between \(t_{kl}\) and \(t_{lm}\).

\(^{32}\)We omit discussion of the optimal shipping network from the perspectives of a firm with market power, and focus on leg-level scale instead.
good $\omega$ and that country $i$ is the lowest-cost supplier of good $\omega$ to $j$ as:

$$\pi_{ijrnw} \equiv \Pr \left\{ p_{ijrnw} \leq \min_{i' \in I, r' \in R_{ij}} p_{ijrnw}' \right\} = \frac{[c_{in}^iK_{ijn}]^{-\theta} \cdot \tilde{\tau}_{ijr}^{-\theta}}{\sum_{i' \in I} (c_{i'n}^{i'}K_{i'jn})^{-\theta} \cdot \sum_{r' \in R_{ij}} \tilde{\tau}_{i'jr}^{-\theta}}.$$  \hspace{1cm} (3)

By the law of large numbers, this is also the share of goods sold in $j$ in industry $n$ coming from $i$ and taking route $r$.\footnote{Recall the number of goods in each industry is set arbitrarily large so that the law of large numbers will hold. The unconditional (pre-selection) average transport cost from $i$ to destination $j$: $\tau_{ijn} = \gamma^{-1/\theta} \left( \sum_{r \in R_{ij}} \tilde{\tau}_{ijr}^{-\theta} \right)^{-1/\theta}$ where $\gamma$ is the function $\Gamma(t) = \int_0^\infty x^{t-1} \exp^{-x} dx$ evaluated at $\left( \frac{1+\theta}{\theta} \right)^{-\theta}$.} Following Allen and Arkolakis (2019), we define the matrix

$$A_n = [a_{ijn} \equiv t_{ijn}^{-\theta}],$$  \hspace{1cm} (4)

where each element is a function of the leg-specific trade cost $a_{ijn} \equiv t_{ijn}^{-\theta}$. Next, we define the matrix $B$, the expected trade cost matrix, as,

$$B_n = [b_{ijn}] \equiv (I - A)^{-1}. $$  \hspace{1cm} (5)

Using these and substituting for the definition of $\tilde{\tau}_{ijr}$ (equation (2)) and summing across routes $r$ that pass between leg $k$ to $l$, we can express the share of imports in industry $n$ in destination $j$ that come from origin $i$ which passes through leg $kl$ as:

$$\pi_{ijrn}^{kl} = \left( c_{in}^iK_{ijn} \right)^{-\theta} \cdot b_{nik}a_{nkli}b_{nlj} \cdot \Phi_{jn}^{-1},$$  \hspace{1cm} (6)

where $\Phi_{jn} = \sum_{i'} (c_{i'n}^{i'}K_{i'jn})^{-\theta} \cdot b_{ni'j}$ is a multilateral resistance term that accounts for the average costs, openness, and connectivity of competitors from all other countries $i'$. This equation is the direct analogue to equation (7) in Allen and Arkolakis (2019) with one key distinction: $\Phi_{jn}$, which accounts for the Ricardian selection of lowest-price sources for each good $\omega$. Intuitively, the traffic flowing on a given leg responds both to that leg’s effectiveness in reducing route costs as well as to competitive forces which make trades increasingly less likely to be pursued for more expensive routes. However, multilateral resistance is at the origin-destination $i,j$-level, and therefore enters proportionately into traffic flows for all link-level $kl$-pairs – a fact that will be crucial for our estimation.

Furthermore from summing across industries, origins, and destinations, we can recover the share of observed global shipping that passed through leg $kl$:

$$\pi^{kl} = \sum_{n} a_{nkl} \cdot \sum_{j} b_{nlj} \Phi_{kn} \Phi_{jn}^{-1}. $$  \hspace{1cm} (7)

Equations (6) and (7) correspond the shares of goods passing through leg $k$ to $l$, including
shipments bound for \( l \) and those continuing onward to other destinations. Because they account both for optimal route selection and competition on price, they correspond to observable volumes after route selection and competition among producers.

The sum of products sold in \( j \) in industry \( n \) from country \( i \) equals the share of products sold in \( j \) in industry \( n \) coming from \( i \) and taking route \( r \), summed across all \( r \) routes:

\[
\pi_{ijn} \equiv \sum_{r} \left[ (c_{in}R_{ij}^{-\theta}) \cdot \tau_{ijr}^{-\theta} \cdot \sum_{r' \in R_{ij}} \phi_{r'}^{-\theta} \right] = \frac{(c_{in}R_{ij}^{-\theta}) \cdot \tau_{ijr}^{-\theta}}{\Phi_{jn}}. \tag{8}
\]

**Closing the model** Factor and goods market clearing and balanced trade conditions close the model. Unnecessary for estimation, we defer them to Section 8 when we conduct counterfactuals.

### 4.3 The Network Effect of Adjustments on Trade

A change in the leg cost between \( k \) and \( l \) (\( t_{kl} \)) can affect trade volumes between an origin \( i \) and destination \( j \) through the trade network. However, Ricardian competition can interact with the trade network to generate unexpected effects. For any change to the cost \( t_{kl} \), trade volumes between \( i \) and \( j \) will adjust according to the following equation:

\[
\frac{dX_{ijn}}{dt_{kl}} = \frac{\partial X_{jn}}{\partial t_{kl}} \cdot \pi_{ijn} + X_{jn} \cdot \left[ \frac{\partial c_{in}}{\partial t_{kl}} \cdot \pi_{ijn} + \frac{\partial \tau_{ijr}}{\partial t_{kl}} \cdot \pi_{ijn} + \frac{\partial \Phi_{jn}}{\partial t_{kl}} \cdot \pi_{ijn} \right].
\]

The first term on the right is the effect of \( t_{kl} \) on trade with \( i \) through a change in the volume consumed at \( j \) in industry \( n \). In square parentheses, the first term is the effect through any changes to the production costs at \( i \), which can happen if the price of inputs changes or through a change in wages. The second term is the effect through trade costs between \( i \) and \( j \) in industry \( n \), and the final term is the effect through multilateral resistance.

What can we say about the signs on these terms? When the trade cost matrix is endogenous to trade volumes, as it would be in the presence of scale economies, these terms are ambiguous, as a change in \( t_{kl} \), by changing trade volumes, changes traffic volumes at each leg, and therefore equilibrium effects on the full matrix of trade costs.

When there is no endogenous scale response, only the final term can be negative. Intuitively, a reduction in trade costs between \( k \) and \( l \) can increase consumption at \( j \), reduce expected trade costs between \( i \) and \( j \), and reduce production costs at \( i \), all of which result in an increase in trade volumes between \( i \) and \( j \). However, a reduction in trade costs between \( k \) and \( l \) also stiffens competitions at \( j \). If this last effect is large
enough, it can overturn the sign of the first three.

In the scale-free case, the total effect is positive if and only if the elasticities of consumption at $j$ ($\epsilon_{X_{jn},t_{kl}}$), production costs at $i$ ($\epsilon_{c_{in},t_{kl}}$), and trade costs between $i$ and $j$ ($\epsilon_{\tau_{in},t_{kl}}$) with respect to $t_{kl}$ are larger than the elasticity of multilateral resistance at $j$ with respect to $t_{kl}$ ($\epsilon_{\Phi_{j},t_{kl}}$). Furthermore, $\frac{\partial X_{ijn}}{\partial t_{kl}} > 0$ if and only if:

$$\epsilon_{X_{jn},t_{kl}} + \left[ \epsilon_{c_{in},t_{kl}} + \epsilon_{\tau_{in},t_{kl}} \right] (1 - \pi_{ijn}) > \sum_{i' \neq i} \left( \epsilon_{c_{i'n},t_{kl}} + \epsilon_{\tau_{i'jn},t_{kl}} \right) \pi_{ijn}. \quad (9)$$

The sum of the effects on production and transport costs between all other countries $i'$ (other than $i$) and $j$ has to be less than a function of the effects on production and transport cost at $i$ and the overall propensity of consumption at $j$ to grow. This last expression shows most clearly that the effect of a decline in trade costs between $k$ and $l$ has the potential to negatively affect trade flows between $i$ and $j$ if it differentially lowers trade and production costs from $i$’s competitors.

5 Estimation

We now show how to link our model to real world data, use the model to recover the trade costs underlying the global shipping network, and estimate a scale elasticity in shipping.

5.1 Estimation Preliminaries

Using equations (6) and (8) we can calculate the probability of any good traveling through leg $kl$ conditional on being sold from origin $i$ to destination $j$. If $X_{ijn}$ is the total value of trade between $i$ and $j$ in industry $n$, we can express the total volume of traffic between $k$ and $l$ in a given industry $n$ as:

$$\Xi_{kl}^{n} \equiv \sum_{i} \sum_{j} X_{ijn} \cdot b_{ikn}a_{kln}b_{ljn}^{-1}b_{ijn}^{-1}. \quad (10)$$

Conditional on the observed trade values $X_{ijn}$, the contribution of trade between $i$ and $j$ to the traffic between legs $k$ and $l$ is invariant to multilateral resistance, tariffs, or technology. This equation is identical to Allen and Arkolakis (2019), despite the significant differences between our frameworks. In particular, expensive trade routes here suffer from Ricardian selection at destination markets, where the route’s impact on prices make them less competitive. Yet, this does not impact the estimation of trade costs. The intuition for this result is that Ricardian selection, non-transportation trade costs such as tariffs, and multilateral resistance all operate by adjusting total trade, but do not differentially favor
one route from an origin $i$ to a destination $j$. Any change to non-transportation costs in one country will affect trade from that country and others proportionally on all routes.

Mapping our model into the data we make one final assumption: for a set of industries $\tilde{N}$, trade costs are identical and all trade ($X_{\tilde{N}} \equiv \sum_{n \in \tilde{N}} X_n$) and traffic ($\Xi_{\tilde{N}} \equiv \sum_{n \in \tilde{N}} \Xi_n^{kl}$) are observable. Summing Equation (10) over industries $n \in \tilde{N}$ yields:

$$\Xi_{\tilde{N}}^{kl} = \sum_i \sum_j X_{i\tilde{N}} \cdot b_{\tilde{N}ik} a_{\tilde{N}kl} b_{\tilde{N}lj} b_{\tilde{N}ij}^{-1}.$$  

Equation (10) tells us that to accurately measure transport costs, we only need data on transportation and traffic for all goods in an industry. Equation (11) tells us that we can use traffic across multiple industries so long as we have the correct trade aggregate, we see all traffic for those industries, and we can assume transport costs are identical in those industries. We implement equation (11) using observed total containerized traffic, where transportation costs are likely similar and isolating trade in containerized industries, applying it in estimation only to legs where all traffic is observed.

### 5.2 Recovering Scale Elasticities

**The cost–scale relationship** The existence of a scale economy in shipping implies that perturbations to the global shipping network that change trade volumes will in turn impact the leg cost matrix estimated in the next section. Such effects must be accounted for in order to correctly estimate counterfactual adjustments.

With perfect data on $\Xi$ and $X$ and the relationship $B = (I - A)^{-1}$, where $a_{kl} = t_{kl}^\theta$, we can use Equation (11) to recover leg level trade costs $A$. As these trade costs are multiplicative, for interpretation we recenter them, and run the following regression:

$$\ln(t_{kl}^\theta - 1) = \alpha_0 + \alpha_1 \cdot \ln \Xi_{kl} + \alpha_2 \cdot \ln d_{kl} + \epsilon_{kl},$$  

where $\alpha_0$ is a constant, $\alpha_1$ is the relationship between price and quantity, $\alpha_2 \cdot \ln d_{kl}$ is the coefficient and measure of log sea-distance from $k$ to $l$ respectively. $t_{kl}^\theta - 1$ allows us to interpret $\alpha_1$ as the elasticity between cost and traffic volumes to a trade elasticity $\theta$. That is, to interpret our results as elasticities, they must be deflated by trade elasticity $\theta$, which we do not recover directly.

Of course, this relationship cannot be taken as causal. Lower cost legs may face larger demand precisely because unobserved cost-reducers induce higher levels of demand on those legs. Essentially, we wish to observe the supply elasticity, but we have only market-clearing prices and quantities. We therefore need a demand shifter.
**Geography-Based Instrument**  We build such a shifter by using the intuition of our model to construct a geography-based instrument for demand. Demand for a given leg will be higher, all else equal, if the leg lies along the most direct route between an origin and a destination. For example, consider routes from origin South Korea to destination the Netherlands. Routes that include the leg China-Singapore are closer to the direct Korea-Netherlands route compared to routes that include the leg China-Australia. As such, more Korea-Netherlands trade should flow through the China-Singapore link than the China-Australia link, which would involve a longer detour. Links that are effectively out of the way for most journeys should, all else equal, face lower demand.

Operationalizing this intuition, we relate the direct sea-distance between an origin and a destination to the distance of two legs as part of a three-leg journey, where the omitted middle leg is the object of interest. We calculate the instrument $z_{kl}$ as:

$$z_{kl} = \sum_{i \neq k,l} \frac{\sum_{j \neq \{k,l\}} \text{Pop}_{i,1960} \text{Pop}_{j,1960} d_{ij}^2 \left(d_{ik} + d_{lj}\right)^2}{d_{jk}}$$

where $d_{ij}$ is the sea distance between origin $i$ and destination $j$, and the square of the relative excess distance between legs $ik$ and $lj$ ($d_{ik} + d_{lj}$) is weighted by the year 1960 population at each origin $i$ and destination $j$, $\text{Pop}_{i,1960}$ and $\text{Pop}_{j,1960}$.

Figure VII plots the first-stage relationship between our instrument and traffic.

For plausible identification, our demand shifter instrument has to be generally uncorrelated with unobserved changes in cost determinants for a particular leg ($\text{corr}(\epsilon_{kl}, \ln z_{kl}) = 0$). As a robustness check, we recalculate our instrument in equation (13) in a simplified setting by omitting the shortest 10 percentile distances for each origin $i$ and destination $j$ respectively. As reported in the results sections, the effects are similar.

An important note is that many mechanisms may be at work generating this observed scale economy, including ship size and market power among others. These mechanisms may generate different out of sample results, and further work should be done to isolate and test for these. Because a multitude of such mechanisms may be at work simultaneously, we choose a model-consistent, agnostic approach in our estimation of scale.

Formally we attempt to construct moments $m_1(\alpha) = Z\epsilon$, based on Equation (12). However, we first need to recover leg-level trade costs $t_{kl}$ for container-ship routes.

---

1960 Population here stands in place of GDP, which may be endogenous to the trade costs in our model. The year is chosen both because immigration and populations prior to 1960 could not plausibly be impacted by 2014 containerized shipping costs.
Figure VII: Residualized Plot of Correlation Between Instrument and Traffic

Notes: The figure shows a binned scatter plot of 1947 link \( kl \) observations with natural log of sea distance between \( k \) and \( l \) is included as a control. The x-axis is the natural log of the instrument \( z_{kl} \). The y-axis is the natural log of traffic on leg \( kl \). The standard error printed is clustered two ways by nodes \( k \) and \( l \).

5.3 Recovering Trade Costs

The recovery of trade costs \( t^\theta_{kl} \) requires two observable objects: trade values and traffic volumes.\(^{35}\) Our traffic data comes from our global port of call AIS shipping data.\(^{36}\) We use aggregate trade data from Centre d’´etudes Prospectives et d’Informations Internationales (CEPII) and their BACI international database for 2014, segregating containerized and non-containerized commodities.\(^{37}\)

In an ideal world, estimation would recover the trade costs that directly rationalize observed bilateral containerized traffic flows—a just identified case. While we directly observe ocean traffic, our data omits land based trade and internal within-country trade. This is an issue for network links between geographically contiguous countries.

We overcome this limitation by assuming a functional form that allows for estimation

---

\(^{35}\)This procedure is agnostic to the exact specification of any particular trade model that generates trade value flows \( X \). By conditioning estimation on these flows \( X \), all origin, destination, and origin-destination factors are controlled for. In particular, items such as all origin-destination tariffs and non-tariff barriers are all accounted for. This does not mean that we can disentangle the two, rather we can directly account for these factors collectively.

\(^{36}\)Units for traffic is in TEU. Recall we estimate ship-by-leg TEUs by combining reported ship draught and maximum TEU. This stage does not rely on merged US Customs data.

\(^{37}\)We use 2014 US Customs data on containerized and non-containerized shipments to construct the share of each HS 4-digit commodity code that is transported by container. All commodities with a containerized share above 80% are labeled as containerized. This procedure shuts down the substitution between containerized and non-containerized transport. In practice we find a bimodal distribution, with some commodities being never containerized (e.g. oil and iron ore) and others always containerized (e.g. washing machines and children’s toys). This process is documented in Appendix A.3.
without requiring the direct observation of overland links. We consider the mapping:\(^{38}\)

\[ a_{ij} = t_{ij}^\theta = \frac{1}{1 + \exp(Y\beta)} \in [0, 1], \]

where \(a_{ij}\) is an element of the matrix \(A\) and the matrix \(Y\) is a vector defined as

\[ Y\beta = \beta_0 + \beta_1 \log \text{sea distance}_{ij} + \beta_2 \log \text{traffic}_{ij} + \beta_3 \log \text{traffic}_i + \beta_4 \log \text{traffic}_j + \beta_5 \mathbb{I}_{\text{backhaul}} + \beta_6 \mathbb{I}\{i, j \in \text{Land Borders}\}, \]

where \(\beta_0\) is an intercept, \(\beta_1\) considers sea distance between the nearest principal port, \(^{39}\) and \(\beta_2\) considers port-to-port traffic. \(\beta_3\) and \(\beta_4\) consider the total incoming and outgoing traffic at ports \(i\) and \(j\) respectively. \(\beta_5\) considers the role of the backhaul problem from Wong (2020), where ship capacity is fixed by the shipping direction with the higher demand. Finally \(\beta_6\) is an indicator for a shared land border.\(^{40}\)

It is crucial to note two things. First, the above equations posit relationships between observables, however, at this stage our objective is not the vector \(\beta\) of coefficients—which may reflect endogenous variables—but the resulting predictions for \(a_{ij}\). We seek to fully saturate the variation in the data in order to generate the closest prediction the data can yield for the matrix \(A\) relative to the just-identified case. This recovers the trade costs while being agnostic to their underlying determinants, including potential market power as well as possible geographic indicators. Secondly, note that parameters for \(\beta\) will yield estimates of every trade cost \(a_{ij}\), but we need not discipline \(\beta\) by comparing traffic on every link. We can omit traffic within-countries as well as between countries that share overland routes and still recover estimates of \(a_{ij}\).

We create a moment \(m_2\) that finds the vector \(\beta\) which minimizes the differences of expected traffic, \(\hat{\Xi}(A(\beta); Y)\), which is constructed from estimates of \(\beta\) as well as trade data \(X\) and observed traffic \(\Xi\) for countries that do not share a land border:

\[ m_2(\beta) = \left( \hat{\Xi}_{kl}(\beta|X, Y, \theta) \right) - \left( \Xi^\text{data}_{kl} \right) \]

As noted, we do not fully observe traffic of containerized goods on geographically contiguous legs, and we do not perform estimation using traffic data from these links. Disciplined by non-land border traffic flows, guesses for \(a_{kl}\) for overland link \(k\) to \(l\) are still generated from any guess of \(\beta\), and these along with observed trade flows \(X\) generate

---

\(^{38}\)This functional form maps from the real numbers to the unit interval, as is required by our theory.

\(^{39}\)For each country pair, we calculate the volume-weighted mean sea distance across all port pairs. These data are available for download from our websites.

\(^{40}\)We do not estimate diagonals \(a_{ii}\). We assume that self-shipping costs do not change in the counterfactual and we estimate our data on international trade data.
predicted traffic flows for sea-only legs where traffic is observed.

5.4 Joint Estimation

We combine our scale estimation and the recovery of the trade costs into a single stage:

\[ m_1(\alpha, \beta) = Z\epsilon(\alpha, t(\beta)) \]
\[ m_2(\beta) = (\hat{\Xi}_{kl}(\beta)) - (\Xi_{data}) \]

We conduct a two-stage GMM procedure, using optimal instrumental variable weights estimation for the first set of moments \( m_1 \), which accounts for our casual estimates of scale, and the inverse of trade volumes on the second set of moments \( m_2 \), which rationalizes leg-level trade costs \( t_{kl} \), conditional on observable world trade \( X \) and container traffic \( \Xi \).

We reiterate that we only conduct inference on the parameters \( \alpha \). We treat \( \beta \) as a set of incidental parameters, important for estimation, but not for inference. The second stage computes an optimal weighting matrix \( W \) using the first stage results.

6 Results

Scale Economy  Table II reports our instrumented scale elasticity from our scale moments (equation (12)). For the widely used value \( \theta = 4 \) (Simonovska and Waugh, 2014), the interpretation of our causal estimate is that increasing volume on a route by 1% would reduce costs by 0.06%.\(^{41}\) These results are consistent with our third stylized fact linking trade and ship-size in Section 3 and lend support to our initial hypothesis that a major role of entrepôts is their facilitation of scale through concentration of shipments.

Link and average bilateral trade costs  Appendix Figure A.10 graphs our resulting matrix of pairwise trade costs. We present the vector \( \beta \) estimates in the Appendix Table A.4 as purely predictive parameters, not fundamentals that we can alter in the counterfactuals. Instead, we simply need to know if our \( \beta \) estimates produce containerized ship traffic that reflects the world. With a full \( A \) matrix, we also can generate a full \( B \) matrix, or average bilateral transport cost between points. We provide our network-consistent link (\( A \)-matrix) and origin-destination (\( B \)-matrix) cost estimates to researchers, and they are available for download on our websites. Appendix Table A.5

\(^{41}\)This leg-level elasticity is more modest, but broadly consistent with the strong scale economies from ship size in Cullinane and Khanna (2000), which measure origin-destination elasticities that would compound, on average, three leg level elasticities. Omitting the shortest 10% routes, produces an elasticity of 0.08—slightly higher but still smaller than others reported in the literature. We search for but do not find evidence of a declining scale elasticity at higher volumes.
Table II: GMM Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln (c_{kl}) )</td>
<td>-0.29</td>
<td>(0.13)</td>
</tr>
<tr>
<td>( \ln (\Xi_{kl}) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \ln (d_{kl}) )</td>
<td>0.57</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.24</td>
<td>(1.45)</td>
</tr>
</tbody>
</table>

Notes: We conduct a two-stage GMM procedure, first using optimal instrumental variable weights estimation the first set of moments and the inverse of trade volumes on the second set of moments. The second stage computes an optimal weighting matrix \( W \) using the first stage results. \( \ln(c_{kl}) \) is the natural log of transportation trade cost on leg \( kl \). \( \ln(\Xi_{kl}) \) is the natural log of traffic volume on leg \( kl \). \( \ln(d_{kl}) \) is the natural log of sea distance between \( k \) and \( l \) computed using Dijkstra’s algorithm.

compares these network bilateral trade costs to distance measures more commonly used.\(^{42}\) Our cost measures have more predictive power than distance and both are significant in a combined specification, implying the measures have distinct predictive power for trade.

Model Fit Figure VIII compares our model-predicted traffic and trade values against their observed counterparts in the data. In Panel (A), we compare actual observed global container traffic shares with the our model-predicted shares using our estimated trade costs. We include both a best fit line and a 45 degree line. We fit the data extremely well, with a correlation between the observed and predicted shares (in logs) of 0.97. Panel (B) compares our estimated trade shares to actual observed trade shares, which we do not target.\(^{43}\) We fit the data well here too with a correlation (in logs) of 0.73.

Alternative Data Definitions Estimates of \( a_{ij} \) are at the country-level. Estimation of a port-level cost matrix is possible. However, that requires a global set of sub-national trade data \( X \), which is not broadly available. Using port traffic and national trade data, we can impute bilateral port trade data and run a version of the above estimation. Results from the port-level estimation are broadly in line with results of main estimation and later scale elasticity estimation, with the correlation between weighted port-pair costs and country-pair costs of 0.6. However, due to the speculative assumptions required to estimate sub-national trade, we view the country-level estimates as more accurate.

\(^{42}\) Using the \( B \)-matrix, we make our bilateral transport cost estimates and country-level measures of market access available for future research.

\(^{43}\) To generate trade flows, we close the model using the full setup in Section 8.
7 Comparison of Model-Predicted Estimates to Data

We now show that our model’s results are highly correlated with two separate sets of data external to our estimation. First, we compare our trade cost estimates with freight rates. Second, we compare model-predicted traffic flows for US-bound shipments to our US microdata. Neither was used in our estimation procedure. In both cases, we find high positive correlations between our results and these external data.

7.1 Cost Estimates with Freight Rates Data

First, we compare our origin-destination expected trade cost estimates $\tau_{ij}$ with port-level container freight rates paid by firms from Wong (2020). These rates are for transporting a standard full container load between port pairs and include the base ocean rate, terminal handling charges at both origin and destination, and fuel surcharge. They are for the largest ports globally which handle more than 1 million containers annually and account for about 73 percent of global container volumes during this time period (World Bank). As such, we are only comparing a subset of the cost estimates from our entire sample with these freight rates (210 observations). We find a high correlation of 0.71 (Figure IX).
Figure IX: Correlation Between Cost Estimates With Actual Freight Rates

![Correlation Between Cost Estimates With Actual Freight Rates](image)

Notes: Data points compare origin-destination predicted costs $\tau_{ij}$ to average freight rates from (Wong, 2020). Circle size are weights for container volumes (TEU). The slope of line is the weighted regression coefficient.

7.2 Traffic Estimates with US Microdata

In order to assess our model’s ability to capture actual indirectness, we compare our model predictions for paths of US-bound shipment traffic to our US microdata. Our estimation, which uses global traffic data rather than the US microdata, delivers predictions for how US-bound shipments travel through the shipping network. Equations (6) and (8) imply

$$\hat{\pi}^{kl}_{iUS} = b_{ik}a_{kl}b_{lj}b_{ij}^{-1}$$

as the ratio of all shipments from $i$ to the US that are observed flowing through leg $k,l$.

We compare the above value from our estimation to the proportion of goods coming into the US from any origin $i$ on leg $kl$, which we call $\pi^{kl}_{iUS,Data}$, by aggregating shipments using link $kl$ in our microdata—described in Section 2 and used in Section 3 but not used in estimation. Column (1) of Table III reports the univariate regression outcome, weighted by total origin TEU. We find that a coefficient of 1 is within the confidence interval. Over half of the variation in the observed distribution can be explained using the predicted probabilities.

Next, summing the predicted probabilities in equation (14) across all origins $i$, the model delivers a prediction for the total amount of US-bound traffic on a given leg $kl$:

$$\hat{\Xi}^{kl} = \sum_{i} X_{iUS} \cdot \hat{\pi}^{kl}_{iUS}$$

where $X_{iUS}$ is the total trade flow from origin $i$ to the US. Again, we can compare this
Table III: Correlation Between Traffic Estimates With Microdata

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\pi}_{iUS}$</td>
<td>0.846</td>
<td>0.827</td>
<td>0.837</td>
<td>0.824</td>
<td>0.844</td>
<td>0.872</td>
</tr>
<tr>
<td>$\hat{\Xi}_{kl}$</td>
<td>1.225</td>
<td>1.241</td>
<td>1.319</td>
<td>1.336</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum_k \hat{\Xi}_{kl}$</td>
<td>652</td>
<td>1.319</td>
<td>2153</td>
<td>108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.513</td>
<td>0.659</td>
<td>0.589</td>
<td>0.513</td>
<td>0.669</td>
<td>0.595</td>
</tr>
<tr>
<td>F</td>
<td>50.54</td>
<td>91.60</td>
<td>22.91</td>
<td>51.75</td>
<td>96.88</td>
<td>22.53</td>
</tr>
</tbody>
</table>

Notes: $\hat{\pi}_{iUS}$ is the model-predicted share of goods from origin $i$ to US destination flowing through leg $k,l$, $\hat{\Xi}_{kl}$ is the model-predicted total US-bound traffic on a given leg $k,l$, and $\sum_k \hat{\Xi}_{kl}$ is the model-predicted total US-bound traffic through node $l$. Their corresponding variables observed in the compiled microdata are indicated with subscript “Data”, $\hat{\pi}_{iUS,Data}$, $\hat{\Xi}_{kl,Data}$, and $\sum_k \hat{\Xi}_{kl,Data}$. Columns (1) to (3) are restricted to nonzero traffic volumes in the US microdata while Columns (4) to (6) includes journeys with zero traffic volumes in the US microdata (All Data). Columns (1) and (4) results are robust to tobit specifications which allow for lower and upper censoring limits. Standard errors clustered by origin and destination countries.

Standard errors clustered by origin and destination countries.

We can then compare this to its counterpart in the microdata, which we call $\sum_k \hat{\Xi}_{kl,Data}$. The coefficient is positive and significant here as well and a coefficient of 1 is within the confidence interval (Column (3), Table III).

In the microdata, there are a number of legs for which there are no observed journeys, or zero traffic volumes. However, our model predicts that there should be some small amount of traffic on every leg. In Columns (4) through (6), we re-run the regressions for each corresponding predicted traffic estimate including these legs with zero observed volumes. Accordingly, there is a big jump in the number of observations with this inclusion. We continue to find a positive and significant correlation between our estimates and the microdata, and that a coefficient of 1 is within the confidence interval. This is because our model predicts extremely low volumes of trade on these legs, and so including these links does not significantly change our results.
Our paper provides a new set of global trade costs which accounts for the trade network. By demonstrating a tight fit between our cost and traffic estimates and two separate sets of observed data external to our estimation, we gain confidence that these estimates reflect actual costs in the trade network. Additionally, these results serve as a check to the validity of the Allen and Arkolakis (2019) approach. Allen and Arkolakis (2019) impute traffic and trade flows within the US highway system for their estimation.\footnote{They assume that the observed traffic for a link is proportional to the underlying value of trade on that link. This assumption is later on verified by comparing their predicted trade flows to actual flows from the Commodity Flow Survey.} Despite the strong structural assumptions made and the limited data requirements, our checks curtail the risk that our estimates are wildly off the mark.

In addition to our leg and origin-destination cost estimates, we provide model-implied indirectness measures for ocean shipping to researchers. These are available for download on our websites.

8 Counterfactuals

In this section we explore three sets of counterfactuals. The first set considers the role of a negative trade shock, the United Kingdom leaving the European Union. The second set studies the trade cost effects of global warming, with the Arctic opening up to trade between the Pacific and Atlantic Oceans, bypassing the Suez and Panama canals. The third set examines the relative impact of infrastructure investments in each country in our data.

To estimate these counterfactuals, we first introduce structural assumptions into our general framework as well as factor and goods market clearing and balanced trade conditions to deliver a quantifiable general equilibrium model.

8.1 Counterfactual Methodology

Closing the model We adopt the Caliendo and Parro (2015) framework. We assume there are three sectors ($N = 3$): containerized tradables $c$, non-containerized tradables $nc$, and nontradables $nt$ ($n \in \{c, nc, nt\}$), all three of which are used as final goods and intermediates in roundabout production. See Appendix D for full details.

Equilibrium in changes Defining the general equilibrium using hat algebra, we consider two sets of changes: (1) link-level transport costs $\hat{t}_{kl} = t'_{kl}/t_{kl}$, which change expected trade costs $\hat{\tau}_{ijn} = \tau'_{ijn}/\tau_{ijn}$, and (2) changes in non-transportation trade costs...
\( \hat{\kappa}_{kl} = \kappa_{kl} / \kappa_{kl} \). Both alter the endogenous costs of production, price indices, wage levels, trade flows, and welfare.\(^{45}\) We solve for how wages and prices change \( \{ \hat{\bar{w}}_{i}, \hat{\bar{P}}_{i} \} \) as a function of changes to model primitives, \( \{ \hat{\tau}_{ijn}, \hat{z}_{in}, \hat{\kappa}_{ijn} \} \), and compute changes in marginal costs \( \hat{c}_{m} \) and trade volumes \( \hat{X}_{ij} \).

**Additional Data** We combine our trade volume data with country level input-output data from the EORA database aggregating over three sectors: non-traded goods, container-shipped traded goods and non-container traded goods and use country level consumption and production data to compute Cobb-Douglas shares \( \eta \) and \( \gamma \).\(^{46}\) This gives us a sample size of 136 countries. We follow the literature and conservatively set \( \theta = 4 \) (Simonovska and Waugh, 2014).

**Procedure** Changes to transport costs are implemented as changes to link costs \( t_{kl} \), which, translated through the model, generate changes in the expected trade cost between every bilateral trading in our data—even those that are not directly connected with each other. Once calculated, these bilateral changes enter isometrically to changes in bilateral non-transportation costs. For analysis which includes the impact of scale, we model a new equilibrium in the short-to-medium run, by following an iterated procedure in Algorithm 1 in Appendix E.1. In this procedure, we start at today’s equilibrium and allow all shippers to optimize their transportation patterns. We then recalculate trade costs at new volumes according to Equation (12). We iterate, allowing re-optimization until a new stable equilibrium is reached. There may be alternative equilibria, however we focus on the unique equilibrium from our current starting point—the world today.\(^{47}\)

### 8.2 Brexit

We begin by modeling Brexit as a 5% increase in trade costs for goods that originate or are destined for the UK.\(^{48}\) We assume that these bilateral non-transportation trade costs will not impact goods that are transship or temporarily stop at British ports. Irish exports destined for Britain will face an increase tariff cost, while Irish exports destined for the United States will not—even if the good stops in Felixstowe en route.

\(^{45}\)As in the literature we assume that trade is balanced up to a constant deficit shifter.

\(^{46}\)We hold trade deficits constant in all our counterfactuals.

\(^{47}\)Kucheryavy, Lyn and Rodriguez-Clare (2019) establishes a common mathematical structure that characterizes the unique equilibrium in multi-industry gravity trade models with industry-level external economies of scale. Their structure requires that the product of the trade and scale elasticities to be not higher than one, which is satisfied in our case.

\(^{48}\)Alternatively, we can model the effect of port infrastructure improvements either between bilateral pairs or for any single nation.
### Table IV: Aggregate Counterfactual Outcomes, Basis Points

<table>
<thead>
<tr>
<th></th>
<th>Welfare</th>
<th>Container Trade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta \kappa_{kl}$ No Scale (1)</td>
<td>$\Delta \kappa_{kl}$ Scale (2)</td>
</tr>
<tr>
<td></td>
<td>$\Delta \kappa_{kl}$ No Scale (3)</td>
<td>$\Delta \kappa_{kl}$ Scale (4)</td>
</tr>
<tr>
<td></td>
<td>$\Delta t_{kl}$ No Scale (5)</td>
<td>$\Delta t_{kl}$ Scale (6)</td>
</tr>
<tr>
<td></td>
<td>$\Delta t_{kl}$ No Scale (7)</td>
<td>$\Delta t_{kl}$ Scale (8)</td>
</tr>
<tr>
<td><strong>Panel (A): Brexit</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Changes</td>
<td>-2.3</td>
<td>-24.5</td>
</tr>
<tr>
<td></td>
<td>-10.0</td>
<td>-112.7</td>
</tr>
<tr>
<td><strong>Panel (B): Arctic Passage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Changes</td>
<td>1.5</td>
<td>38.2</td>
</tr>
<tr>
<td></td>
<td>3.3</td>
<td>101.8</td>
</tr>
<tr>
<td></td>
<td>8.9</td>
<td></td>
</tr>
<tr>
<td><strong>Panel (C): Local Transportation and Non-Transportation Cost Declines</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Changes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.08</td>
<td>2.93</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>(0.20)</td>
<td>(5.31)</td>
</tr>
<tr>
<td>Changes at Entrepôts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.02</td>
<td>0.56</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>(0.05)</td>
<td>(1.10)</td>
</tr>
<tr>
<td>Changes at Non-Entrepôts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.06</td>
<td>1.89</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>(0.16)</td>
<td>(3.23)</td>
</tr>
</tbody>
</table>

Notes: Columns (1)-(4) present aggregate welfare changes. Columns (5)-(8) present changes to aggregate container trade volumes. Columns (1), (2), (5), and (6) present results for cases where non-transportation trade costs are reduced. Columns (3), (4), (7), and (8) present results for cases where transportation costs are reduced. Odd Columns correspond to cases where no scale economy feedback loops are allowed, and even columns present results allowing for scale economy feedback effects. Panel (A) presents results for a 5% increase in non-transportation trade costs $\kappa_{ij}$ between the UK and its trading partners. Case 1 and 2 of the Panel (A) counterfactual correspond to columns (1) and (5) and (2) and (6) respectively. Panel (B) presents results for Arctic Passage counterfactual. Case 1, naive impacts on origin-destination pairs, correspond to Columns (1) and (5). Case 2, allowing for network trade, corresponds to Columns (3) and (7). Case 3, adding in the impact of the scale, correspond to Columns (4) and (8). Panel (C) reports results for our third counterfactual, local transportation and non-transportation cost declines for each of the 136 country in our data. Results for infrastructure improvements (transportation cost reductions) are reported in Columns (3),(4), (7) and (8). Results for non-transportation cost reductions are in columns (1), (2), (5), and (6). Global Changes are broken out into gains accrued at Entrepôts and elsewhere. In each case, we report the mean impact across 136 counterfactual exercises and the standard deviation of the 136 effects in parentheses.

We model two cases: first without, then with the impact of scale on the trade network. In our first case, as in a traditional model, affects are only felt through changes in trade with the UK or through the multilateral resistance term (including the direct effect on global value changes). However, with scale economies, a decrease in British trade will have spillover effects on the trade costs of neighboring countries through the transportation network. Lower trade volumes lead to increased transport costs, not only for the UK, but also countries that currently find the UK to be a preferred entrepôt. Irish exports to the US will be more costly, as they will either have to pay the increased costs of travelling through Britain, find an alternative entrepôt, or take a low-volume and more costly direct trip.

Panel (A) of Table IV reports aggregate effects. The direct effect decreases global
welfare by 0.04%. The introduction of scale economies leads to a decrease of 0.16%. Trade volumes follow a similar pattern. Figure X highlights the distributional effects in terms of welfare (see Appendix Table A.16 for trade volumes). Scale economies amplify effects, significantly impacting the rest of Europe. Notably, the impact of scale is not well-predicted by the non-scale case. Significant effects are seen in Iceland and other Nordic countries, many of which rely on United Kingdom feeder routes to get their goods to large vessels that ply transoceanic trade with Shanghai and New York. As their trade is not easily routed through alternatives, Ireland, Iceland, and these Nordic countries suffer.

**Figure X: Welfare Changes - Brexit**

(A) Tariff Change, No Network Scale Effects

(B) Full Trade Network Effects and Scale Economies

Notes: These two plots show the percent change in welfare (the relative price index) of a simulated 5% increase in trading costs with the United Kingdom for all countries in our dataset. Darker reds reflect a greater increase and blue represents no change. Omitted countries are white. Panel (A) reflects changes if shipping costs remain constant, reflecting only welfare changes due to changes in prices. Panel (B) allows for a scale economy feedback loop on transportation costs for all countries.

8.3 The Arctic Passage

We model the opening of the once-fabled Northeast and Northwest Passages through the Arctic Ocean between North America, Northern Europe and East Asia as a viable shipping route due to global warming. As an example, a ship traveling from South
Korea to Germany would take roughly 34 days via the Suez Canal but only 23 days via the Northeast and Northwest Passages (the Economist, 2018). For every $kl$ pair, we compute the difference in distance using Dijkstra’s algorithm between world maps with and without arctic ice caps (Appendix A.2). Panel (A) of Figure XI compares the top 150 existing shipping routes today and shortest ocean-going distance of these routes after the Arctic sea passage is viable. New routes going through the Arctic passage are in red, non-changing routes are in brown, and abandoned routes are in blue.

**Figure XI:** The Opening of the Arctic Passage

(A) Shipping Routes: Before and After

(B) Welfare Changes for Top 20 Countries

Notes: The red lines in Panel (A) indicate the Arctic sea routes, brown lines indicate routes that do not change, and blue lines indicate existing shipping routes that are eliminated. The width of each route reflects the total number of containers (TEU) on that route. Panel (B) shows the percent change in welfare (the relative price index) of the simulated opening of the Nordic Passage for the top 20 countries with the largest welfare change with network and scale effects. The first bar reflects only the trade cost changes on routes that are directly affected from the opening. The second bar allows for the trade costs to affect indirect trade with network effects while the third bar allows for the endogenous network response to scale economies.

We compare three different cases. First, we consider a network-naive case where we only allow for changes in origin-destination trade costs between country pairs for which the direct bilateral distance decreased.\(^{49}\) Second, we lower $t_{kl}$ for all observed links with positive traffic according to $\alpha_2$ in Equation (12) calculating new distances with the option of traveling through the Arctic Passage. Here, even countries that do not ship directly to each other—e.g. China and Ukraine—experience changes in expected transport trade costs.\(^{50}\) Third, we repeat the second case accounting for the impact of scale: as trade

---

\(^{49}\)We do not allow third-party-countries to take advantage of these reductions.

\(^{50}\)For countries affected in cases one and two, the magnitude of changes are mechanically identical.
costs change, trade volumes change, reducing trade costs further.

Column (1) of Table IV Panel (B) shows that the direct effects of the Arctic Passage are positive, with aggregate welfare increasing 0.015%, and container trade volumes increasing 0.2%. Case two, accounting for the full trade network impact of the passage, including indirect shipping, doubles the aggregate welfare effect to 0.03% and increases worldwide container volumes 0.4%. Finally, allowing for scale economies triples the impact relative to the network results (0.09% welfare gain) and more than doubles the increase in container volumes (a full 1% increase in global traffic).

Panel (B) in Figure XI plots the top 20 most impacted countries, showing gains are particularly pronounced in East Asian entrepôts which disproportionately benefit from the scale economy. Scandinavian countries also gain due to their geography. Denmark and Finland, which in the naive case have zero or a small trade diversion impact, gain due to the ability to leverage the trade network and scale response from the opening.

Figure XII show changes in the relative wage-adjusted price index (interpreted as national welfare, if we omit the costs of climate change) across the three cases. In the baseline scenario in Panel (A), we see increases from trade between countries on the Northeast passage, and spillover impacts at countries not directly impacted—reflecting classic multilateral resistance and cascading effects from value chains. Interestingly, some countries see small trade diversion effects (Panel (B), Figure XI). Panel (B) in Figure XII shows how, through indirect trade, the benefits of the passage pass on to nearby countries not directly impacted. Panel (C) allows for scale economies to amplify effects.

8.4 The Global Impact of Local Infrastructure Improvements

We now turn to examining the relative importance of entrepôts and the shipping network. We run two types of counterfactuals. For all countries, we consider the impact of transportation infrastructure investment in the form of a 1% reduction in transportation costs \((t_{kl})\) to and from the country, which we refer to as the targeted country. We contrast this with a 1% reduction in non-transportation trade costs \((\kappa_{ij})\). For each, we calculate equilibrium changes both accounting for and not accounting for the endogenous impact of changes in scale on shipping costs at all other—impacted—countries. This makes 136 counterfactuals in each of 4 cases. We use these to systematically investigate the relative importance of entrepôts on trade volumes and welfare, examining differences in the ag-

\[51\text{Appendix Figure A.14 shows related changes in country-by-country containerized exports.}\]
Notes: These three plots show the percent change in welfare (the relative price index) for all countries in our dataset. Darker reds reflect a greater increase and blue represents no change. Omitted countries are white. Panel (A) reflects changes if we only allow trade costs to decrease on routes whose distance is directly reduced to the Arctic Passage. Panel (B) reflects changes if we allow all countries to indirectly access the Arctic Passage through the trade network. Panel (C) allows for the network’s endogenous response to scale economies.

We find that entrepôts are pivotal due to their role in the transportation network, and that scale economies heighten their impact. Infrastructure investment at entrepôts generate relatively larger aggregate welfare impacts and network adjustments. Furthermore, entrepôts play a special regional role, concentrating benefits at nearby countries. Finally, scale economies allow entrepôts to take better advantage of changes elsewhere,
as these on average lead to heightened use of entrepôts, further reducing trade costs and leading to out-sized welfare gains for entrepôts.

**Figure XIII: Most Pivotal Nodes: Change in Welfare Excluding Own**

(A) Decrease in Transportation Costs

(B) Decrease in Non-Transport Costs

Notes: Panel (A) shows absolute values for aggregate net change in global welfare after infrastructure investment in the listed country, excluding the country’s own, for the 20 countries with the largest global impact calculated without scale economies. Panel (B) presents the same for reductions to non-transport trade costs. Black bars represent welfare changes for the no-scale case. Overlaid grey bars represent welfare changes allowing for the network’s endogenous response to scale economies.

We begin by identifying the most pivotal nodes in the network. The general equilibrium model points to a convenient metric for how pivotal a node is: the impact of changes at the node on global welfare excluding a country’s own. Locations that are pivotal to global trade will generate the largest adjustments throughout the network. Figure XIII (A) and (B) list the global welfare impact with and without scale responses for the most pivotal nodes in the network. Larger countries dominate the list in Figure XIII (B) reflecting a combination of size and value chain position. By contrast, our 15 entrepôts dominate the list in Figure XIII (A), the most pivotal nodes for infrastructure investment. Scale economies (the overlaid grey bars) augment the differential impact of entrepôts in Figure XIII (A), while in Figure XIII (B), scale economies generate transportation network changes on top of the initial changes, as evidenced by the out-sized impact of Singapore. Only half the non-entrepôt countries in Figure XIII (B) are highly pivotal when accounting for the endogenous scale response.

Next we assess the spatial concentration of both types of cost reductions. Figure XIV is a binned scatter plot of all country pairs across our counterfactuals. It considers the welfare impact of changes relative to the pair’s distance, adjusting for the impacted country’s average change across all 136 counterfactuals within each type of counterfactual.
The blue and green dots in Figure XIV (B) – nearly overlapping – show the impact of non-transportation cost reductions without scale economies by distance when targeted countries are entrepôt and non-entrepôt respectively. The distance gradient, indicative of gravity, is nearly identical and not statistically distinct. The blue and green dots in Figure XIV (A) show the same for infrastructure investments. Investments at entrepôts have larger impacts overall, but this differential impact is localized—decaying at 5 times the rate. Benefits from entrepôts including an endogenous scale response—orange dots in Panel (A)—persist at larger distances, but decay at 7-times the rate compared to the same from other countries, red dots. Entrepôt slopes are statistically distinct. This highlights the regional roles that entrepôts play in concentrating benefits at nearby countries.

**Figure XIV: Spatial Decay of Benefits By Entrepôt status**

(A) Decrease in Transportation Costs

(B) Decrease in Non-Transportation Costs

Notes: Panel (A) shows binned scatter residual for welfare effects on impacted countries of transportation infrastructure in targeted countries vs distance between the targeted and impacted countries. Blue and Red dots are the no-scale and scale cases for counterfactuals where targeted countries are not entrepôts, respectively. Green and orange dots are no-scale and scale cases, respectively, for counterfactuals where the targeted countries are entrepôts. Panel (B) presents the same for reductions in non-transportation trade costs.

Finally, we find that scale economies disproportionately accrue gains to entrepôts. In Panel (C) of Table IV, for both infrastructure and non-transportation cases, scale increases impacts 2.5-3.5 times the baseline. Benefits to entrepôts are broken out. Entrepôts account for an eighth of global GDP but accrue on average a quarter of the welfare effects of both cases without scale, indicating their greater exposure to trade. This ratio grows when scale economies are accounted for, as heightened trade through entrepôts lowers their trade costs disproportionately, generating excess gains. Averaging across counterfactuals attenuates the counterfactual-level differential: in a regression of the ratio of the no-scale to scale effects on entrepôt status, controlling for target country
fixed effects and clustering by targeted country, we find entrepôts have systematically higher ratios. The coefficient on an entrepôt dummy is 1.212 (SE of 0.595) and 0.755 (SE of 0.379) for infrastructure and non-transportation counterfactuals, respectively. These results, that scale economies in transportation concentrate gains locally at and around hubs, highlight scale economies in transportation as a source of agglomeration.

9 Conclusion

World trade takes meandering routes, aided by scale economies that consolidates trade in entrepôts. We characterize this global container shipping network and its implications for international trade. Guided by a series of novel and salient facts, we model world trade with endogenous trade costs, estimating both the underlying trade costs on all containership routes, as well as scale economies. We then quantify the impact of the trade network on global trade and welfare, highlighting how changes at nodes operate through the network structure, entrepôts, and scale economies to create widespread impacts.

While the focus of this paper is in the recovery of trade costs and the general equilibrium effects, there are two aspects that lend themselves to further study. First, is the extension of entrepôts and hub-and-spoke networks from just container shipping to other markets. We are singularly focused on containerized shipping in our setting as containerized trade accounts for the majority of global seaborne trade. The hub and spoke network and its implications for trade is not specific to just containerized trade (Rodrigue, Comtois and Slack, 2013).

Second is our treatment of market power. Our framework considers the endogenous adjustment of marginal cost but assumes shipping costs and the marginal cost of shipping are the same. Our treatment of scale economies in this paper is intentionally agnostic to the multitude of potential underlying mechanisms that are likely at work. Future work should especially be done to consider mechanisms the roles of fixed costs in enabling the scale economies in containerized shipping, such as the costs incurred by potential oligopolies in setting shipping networks and the endogenous creation of firm-specific hub-and-spoke networks.\textsuperscript{52} While sector-specific research has been done on these networks, fruitful research should take into consideration a tractable general equilibrium framework to be able to quantify welfare effects.

\textsuperscript{52}In particular, we can account for leg-level monopolies and variable markups, but we cannot account for within-firm spillovers in sea route selection.
References


Appendix (For Online Publication Only)

Appendix A  Data Construction

A.1 Shipment Microdata

We compile and combine two proprietary micro-data sets in this project: global ports of call data for all containerships, which allows us to reconstruct the routes taken by specific ships, and United States bill of lading data for containerized imports, which gives us shipment-level data on imports into the United States. Independently, each of these datasets allow us to partially describe the global shipping network. By merging them, we are able to reconstruct nearly the entire journey most shipments entering the United States take, from their initial origin point or place of receipt, to the port of entry into the United States. To our knowledge, we provide the most comprehensive reconstruction of the global shipping network and routes undertaken by individual shipments into the United States.

**Port of call data**  We partner with Astra Paging, which provides us with port of call data for containerships. Astra Paging’s data captures vessel movements using the transponders on these ships (known as the automatic identification system, AIS). A network of receivers at ports collects and shares AIS transponder information (including ship name, speed, height in water, latitude, and longitude). Using the geographic variables in the AIS data, Astra Paging marks entry and exit into a number of ports all over the world and provides us with a dataset of ships’ entry and exit from ports of call, timestamps, and ships’ height in the water, or draft. Using these data elements, we are able to calculate an estimated shipment volume between each port pair by taking the observed draft relative to maximum observed draft and multiplying by total ship capacity.

Our sample covers the a six months period, from April to October 2014. Over this period, we have information on 4,986 unique container ships with a combined capacity of 18.13 million TEU. This represents over 90% of the global container shipping fleet. These ships make 429,868 calls at 1,203 ports. Ports with no AIS receivers or where information is not shared do not show up in our data. In addition, if transponders are turned off or transmissions are not recorded, ports of call can be missed. However, transponders are required to be operational by the International Maritime Organization on ships engaging in international voyages 300 gross tons, applying to all containerships in our sample.
Bill of lading data We partner with Panjiva Inc. (Now a division of Standard and Poor’s) to acquire bill of lading information for all seaborne US imports from April to October 2014. Panjiva cleans this data to standardize the names of the ports, ships, companies, and container volumes. We subset this data to only consider goods that arrive on seaborne container ships.

We put together proprietary bills of lading data, which captures shipment-level information for all containerized imports into the United States. International shipping relies on an industry-standardized system of bills of lading, which act as receipts of shipment, recording all information on the shipment, all the parties involved in the shipping process. The US Customs and Border Patrol (CBP) agency collects these bills in addition to customs information at all ports of entry into the US and this data is obtained from the agency by Panjiva.\(^1\)

Our data captures the foreign location where the shipment originated from, the foreign port where it was loaded on the containership which brings it into the US, and the US port where it was unloaded from the containership. In addition, we know the name and International Maritime Organization identification number of the containership (IMOs) which transported the shipment as well as the shipment’s weight, number of containers (TEUs), and product information. For a subset of the shipments, we observe value information.

Over a six months of US imports from April to October 2014, we see a total of 14.8 million TEUs weighting 106 million tons were imported into the US from 227 shipment origin countries, 225 place of receipt countries, and 144 countries with ports of lading. This accounts for about three quarters of the 2014 TEU and tonnage imports, 77 percent and 74 percent respectively (Maritime Administration, US Department of Transportation).\(^2\) Non-containerized goods, including goods on roll-ons (vehicle carriers), bulk cargo liners (for commodities), and non-containerized cargo ships are not observed in our data.

Specifically, for the purposes of this study, our data captures the following location information for each shipment into the US: the foreign location where the shipment originated from (*shipment origin*), the foreign port where it was loaded on the containership

---

\(^1\)US Bill of Lading data is immediately available for direct purchase from the Department of Homeland Security or though a lag using a Freedom of Information Act. However, this raw data requires substantial computing resources for processing and needs to be standardized over time.

\(^2\)In particular, we miss containers that arrive on trucks and trains from either Mexico or Canada. Our estimation strategy explicitly accounts for this unobserved data.
which brings it into the US (port of lading), and the US port where it was unloaded from the containership (port of unlading). In addition, we know the name and identification number of the containership which transported the shipment as well as the shipment’s weight, number of containers (TEUs), and product information. For a smaller subset of the shipments, we observe value information.

This data set allows us to start tracing the journey of a shipment from its origin to its destination US port, in particular we can determine whether this shipment was loaded at its origin location onto the vessel that brings it directly to its final US destination, or if it went through at least one other location during its journey. When matched with the port of call data, we can reconstruct most of the remaining journey after its port of lading.

**Reconstructing shipment routes** Using the containership information, port of arrival information, timing of unlading and ports of call at US ports, and port of lading information, we are able to match the bills of lading to the journeys of specific containerships, then use the ports of call between lading and unlading to reconstruct each shipment’s path from its foreign origin to US destination.

We use Vessel IMOs which are identifiers that are unique to containership vessels and stay with vessel hulls for the lifetime of their operation. Only about 4000 ships are identified in Bills of lading by IMO. An additional (roughly) 2,000 ships are matched to IMOs using a fuzzy string match, after which matches are made with the help of excellent undergraduate research assistants.

Ports of arrival are recording using UNLOCODEs in the AIS data and US Census Schedule D codes in the Bill of Lading data. We construct a crosswalk with the excellent help of undergraduate research assistants.

Port of lading are recorded using CBP’s Schedule K Foreign ports on Bill of Lading and UNLOCODES in the port of call data. We construct a crosswalk between these with the help of our excellent undergraduate research assistants.

What remains unobserved is the shipment’s journey between its Origin and the first stop (port of lading location) we observe in our data. In particular, this initial portion of the shipment’s journey could take place overland (by trucks or rail) or by sea on another ship. This information is not recorded by both our datasets and therefore is impossible for us to observe. This will under-count overall indirectness overall, but will not affect our model estimation.
For each bill of lading, we match ship, date of unlading and port of unlading to the AIS data on ships’ port of call. Once we match shipments to ships, we record each port of call in the AIS data before the port of unlading as a stop the shipment makes, then remove all stops observed before the ship stopped at the port of lading. If the port of lading is not observed, the route is discarded and the shipment remains unmatched. Furthermore, any routes that include the port of unlading before the date of unlading are discarded, as they represent loops where the port of call for the port of lading is missing.

Over 90% of containerized TEU entering the US are on Bills of Lading can be matched to routes using this method. Appendix Figure A.1 visualizes this merge. Unmatched shipments may have missing and unrecoverable ship information, or ports of call that do not match lading and unlading records on bills of lading. In addition, a small number of reconstructed routes have implicit voyage speeds above 50KPH, and are discarded.

Figure A.1: Combined Dataset: Routes Undertaken by Shipments into the US

As an example, Figure A.2 plots for all containerized trade from the United Arab Emirates (UAE), the proportion that stops in each country. This illustrates the paths shipments take when being transported from the UAE on to the US. Shipments from the UAE collectively stop in many countries before continuing onto the US. Many of the most popular are regional neighbor hubs, including Egypt, Pakistan, but Spain and China also facilitate UAE-US trade.

A.2 Geographic Distance Data

Geographic distance data is computed using two rasterized (with pixels) world maps. One map consists of the all the navigable oceans and large seas, with a polar ice cap, as well as the Suez and Panama canals. The second map, assumes that the Arctic ice sheet melts away due to anthropogenic climate change. In both maps we compute the sea distance between ports of call, and aggregate to the national level using using port-to-port container flows. We do this computation in R using using Dijkstra’s algorithm.
A.3 Aggregate Economic and Trade Statistics

For our main estimation, we also require data on the value of containerized trade between countries. We use aggregate trade data from Centre d’études Prospectives et d’Informations Internationales (CEPII) and their BACI international database for 2014. This database aggregate data from the UN Comtrade Database, aligning data from origin and destination countries. This provides us data on trade volumes from origin to destinations by industry using Harmonized System (HS) codes.

To aggregate industry trade to industries that use container shipments versus trade that does not, we use aggregate data from 2014 from the United State Customs, as disseminated by Schott (2008). This data reports the share of shipments by HS Codes that arrive by containerships. We consider 4-digit HS Codes as a consistent level of aggregation. The distribution of containership share by HS code is bi-modal, with one peak around 0% and another around 100%. We use a cutoff of 80%. So HS codes that are shipped by containership to the US over 80% of the time are classified as “containerizable” trade.

For aggregate trade and economic statistics for using in the counterfactual, we use the Eora global supply chain database with a multi-region input-output table (EORA-
MRIO).\(^5\) We collapse all world trade into three categories; those that are non-tradable, those that typically traded over oceans by containerized vessels, and those that are not typically traded over oceans by containerized vessels.\(^6\) We again classify industries using the methods of Schott (2008).

**Appendix B Additional Descriptive Results**

In this section we report additional results and robustness checks related to the analysis in Section 3.

**B.1 Additional Indirectness Results**

Figure A.3 reports the histogram of number of port stops minus the port of lading if the port of lading is in the country of origin, and the port of unlading. We exclude landlocked countries. The mean number of third-port stops is 4.6 and fewer than 5% of shipments do not stop at additional ports.

*Figure A.3: Distribution of Port Stops per Container (TEU)*

Notes: Figure reports the distribution at the shipment level of the number of unique port stops minus the port of lading if the port of lading is in the country of origin, and the port of unlading, weighted by shipment TEU. Shipments from landlocked countries are excluded.

Next, in Figure A.4, we rerun the analysis in Panel (A) of Figure III weighting by Tons in Panel (A) and USD in Panel (B). For the latter, a minority of shipments data report dollar values. Overall, results are similar to our main results using TEU.

Figure A.5 reports the percent of shipments laded in a third-party country by country of origin. Countries that are closer and trade more with the US are less likely to transship goods at third-party countries—a fact we explore in more detail in Appendix B.3.

Finally, we further explore the result that additional stops increase distance and time

\(^5\) Freely available for academic use from https://worldmrio.com/.

\(^6\) This includes bulk shipping, roll-on roll-off ships, as well as air freight.
Figure A.4: Distribution of Third-Party Countries Involved in Bilateral Trade by Weight and Value

Panel (A) Number of Countries per Ton       Panel (B) Number of Countries per USD Value

Notes: Panel (A) reports the distribution of the total number of unique third-party country stops made by shipments entering the US, weighted by shipment tons (kg). Panel (B) reports the same but weights by value for the portion of shipments for which value measures are reported.

Figure A.5: Transshipped Trade Share between Origin and US Destination

Notes: Figure plots for each country the share of country’s originated shipments transshipped in a third-party country, weighted by TEU. Lighter colors indicated lower levels of transshipped trade share (ie. more direct trade). The US is not included since it is the destination country. Landlocked countries are also excluded. 34 of the shipment origin countries are landlocked accounting for 1.6 percent of total TEUs. The missing remaining countries are either due to lack of overall trade with the US (e.g. Somalia) or due to the merge process (e.g. Namibia).

costs of trade. In Table A.1, we regress, at the shipment level, log of observed distance (Columns 1 through 4) and time (Column 5), on the number of country stops made by a shipment. All port distances are computed using Dijkstra’s algorithm, and time is computed by the difference in AIS logs for port of lading and unlading. Results are clustered two ways by port of unlading and port of lading.

Column (1) reports the baseline relationship: an elasticity of 0.112 (SE 0.022) on stops, controlling for the computed direct sea distance between the port of lading and
Adding port of lading fixed effects (Column 2) or port of unlading fixed effects (Column 3) does not appreciably change the result. In Column 4, we add port of lading-by-unlading fixed effects. Here identification comes from variation between routes where goods come on and off boats at exactly the same ports, but where different ships take different routes (the existence of this variation is explored further in Appendix B.3). Remarkably, the elasticity here remains stable as 0.104 (SE 0.0300). Column 5 repeats our most heavily controlled-for exercise in Column 4 but with time travelled as the variable of interest. We find a striking elasticity of 0.333 (SE 0.0819) which implies that after accounting for the port of lading and unlading fully, doubling stops along the way increases journey time by 33%.

Table A.1: The Relationship Between Indirectness, Distance, and Time

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln Observed Dist</td>
<td>0.112</td>
<td>0.109</td>
<td>0.101</td>
<td>0.104</td>
<td>0.333</td>
</tr>
<tr>
<td>ln Country Stops</td>
<td>0.0223</td>
<td>0.0237</td>
<td>0.0270</td>
<td>0.0300</td>
<td>0.0819</td>
</tr>
<tr>
<td>ln Direct Dist</td>
<td>0.881</td>
<td>0.918</td>
<td>0.896</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln Direct Dist</td>
<td>0.0276</td>
<td>0.0347</td>
<td>0.0282</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lading Port FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Unlading Port FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Lading-Unlading Ports FE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>215,655</td>
<td>215,655</td>
<td>215,655</td>
<td>215,656</td>
<td>215,656</td>
</tr>
<tr>
<td>R²</td>
<td>.942</td>
<td>.954</td>
<td>.945</td>
<td>.966</td>
<td>.774</td>
</tr>
<tr>
<td>F-stat</td>
<td>1360.62</td>
<td>1818.20</td>
<td>1242.46</td>
<td>12.11</td>
<td>16.49</td>
</tr>
</tbody>
</table>

Notes: Table presents regression coefficients for regression of ln Observed distance, the natural log of sea distance traveled between all reported ports of call, or ln Time Travelled, the natural log of time between port of call at port of land and port of unlading, and ln Country stops, the natural log of unique third-country stops, as well as ln Direct Distance, the natural log of the sea distance between the port of lading and unlading. Distances are calculated using Dijkstra’s algorithm and measured in kilometers while time is measured in hours. Observations are shipment level and weighted by TEU. Shipments originating in landlocked countries are omitted. Standard errors in parentheses are clustered two ways by the port of lading and port of unlading.

B.2 Additional Concentration Results

Panel (A) of Figure A.6 tabulates, for each of the top ten countries, the percent of all goods entering the US stopping in that country. The share of shipments accounted for by shipment origination is in blue while shipments observed stopping in the country but not originated in the country is in red. Unsurprisingly, many recognizable entrepôts are listed, including Korea, Panama, Singapore, and Egypt. Perhaps more surprisingly, more than 50% of the containers entering into the US stop in China. While this panel sums to over 1, since each container stops in more than one country, over 80% of shipments to
Table A.2: Concentration Ratios

<table>
<thead>
<tr>
<th></th>
<th>Third-Party Stops</th>
<th>Transshipment</th>
<th>Trade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max/50</td>
<td>426</td>
<td>476</td>
<td>400</td>
</tr>
<tr>
<td>99/50</td>
<td>398</td>
<td>476</td>
<td>96</td>
</tr>
<tr>
<td>95/50</td>
<td>213</td>
<td>135</td>
<td>27</td>
</tr>
<tr>
<td>90/50</td>
<td>112</td>
<td>91</td>
<td>15</td>
</tr>
</tbody>
</table>

Notes: Data present concentration ratios across countries in our data. Third-party Stops are the sum total TEU-weighted shipments that use a country as a third-party country. Transhipments are the TEU-weighted sum total of shipments transshipped at a country, and Trade is the total volume of trade from a country. Countries are ranked and percentile ratios are presented. For example, the country with used by the most shipments (by TEU) as a third-country stop acts as such for 426 times the number of shipments stopping at the median (50th-percentile) country.

the US stop in at least 1 of 5 countries: China, Panama Singapore, Korea, or Egypt.\(^7\)

Panel (B) replicates Panel (A) but for country of lading. Here the total of all bars (including those not graphed) sum to 1, and China again dominates as a source of lading. A few of these top countries, like Germany in (A) and Italy in (B) are majority blue, implying they are important to the US because of their role as an origination country. Other countries, like Singapore, are differentially red, and appear important as entrepôt rather than as countries of origin.

Figure A.6: Roles of Countries in Bilateral Trade: Origin vs Entrepôts

(A) Share of Shipments Stopping in Country, for Top Ten Countries  
(B) Share of Shipments Laded in Country, for Top Ten Countries

Notes: The blue portion in Panel (A) highlights share of all incoming US shipments that originate in the indicated country while the red accounts for percent of all incoming US shipments stopping in that country (not originated), weighted by TEU. Panel (B) replicates Panel (A) but for country of lading. \(^\)
B.3 Variation in Connectivity

There is a high degree of variance in indirectness across countries, as shown in Figures I and A.5. This variation is reasonable explained by traditional gravity variables. In Panel (A) of Figure A.8, we plot number of stops against country GDP and find that countries with higher GDPs are more likely to have less stops on their journeys to the US. In Panel (B), we plot number of stops against distance instead and find that countries which are closer have more direct trade with the US. These results are robust to using port stops instead of country stops (Table A.3) as well as to weighting by containers, tons, and value. One natural interpretation of this would be the endogenous response of shippers to the scale of shipments from these countries. Of course, the availability of direct trade to the US could in principle reverse the causality.

Do shipments from a given origin follow a unique path to the US? Panel (A) in Figure A.9 shows the distribution in the number of unique routes to the US by origin country. With an average of about 397 routes with wide variation (sd 681), observed routes from a single origin are indeed varied. The countries with the highest number of unique routes are big trading partners like China, the United Kingdom, Germany, and well-established entrepôts like Hong Kong. Countries with the lowest unique routes are smaller trading partners like American Samoa, Nauru, Tonga, and Montserrat. The existence of this within-origin route variation will be a particularly important assumption in our model and external validity checks.
**Figure A.8:** Larger and Closer Countries Have Lower Number of Average Stops

**(A) Stops vs Country Size**

![Graph A](image)

**Note:** Figure presents Binned scatter plot with observation at the origin level weighted by total TEU. Landlocked countries are excluded.

**(B) Stops vs Distance**

![Graph B](image)

Notes: Table presents coefficients from country-level regression of ln Cty Stops, the natural log of the TEU-weighted average number of third-party country stops made for shipments from country against ln GDP, the natural log of the country’s GDP, and ln Distance, the natural log of the sea distance between the countries. Observations are weighted by total TEU. Landlocked countries are excluded. Robust standard errors in parentheses.

We can measure the concentration of these unique routes by constructing a Herfindahl-Hirschman Index (HHI) for each origin country using the container shares of each route. Panel (B) in Figure A.9 shows that almost 70 percent of origin countries have fairly low concentration of routes (HHI less than 1500). The average HHI overall is 1475 (sd 1974). Examples of countries with high levels of concentration are like Vanuatu, Cuba, and Liberia while countries with low levels of concentration are Macau, Hong Kong, and Belgium-Luxembourg.

**Appendix C  Estimation**

This section reports additional details, results, and robustness checks from our trade cost estimation.
Figure A.9: Variation in Trade Indirectness

(A) Number of Unique Routes by Origin-Destination Pair

(B) Distribution of Route Concentration

Notes: Panel (A) plots the kernel density plot for the total number of unique routes from a given origin country. Panel (B) plots the distribution of the HHI index for routes from each country.

Table A.4: Predictive Trade Cost Estimates

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$ (intercept)</td>
<td>7.95</td>
</tr>
<tr>
<td>$\beta_1$ (log distance)</td>
<td>0.17</td>
</tr>
<tr>
<td>$\beta_2$ (log route traffic)</td>
<td>-1.04</td>
</tr>
<tr>
<td>$\beta_3$ (log outgoing port traffic)</td>
<td>0.28</td>
</tr>
<tr>
<td>$\beta_4$ (log incoming port traffic)</td>
<td>0.28</td>
</tr>
<tr>
<td>$\beta_5$ (land borders)</td>
<td>-0.39</td>
</tr>
<tr>
<td>$\beta_6$ (back-haul)</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

Notes: Results presented here are moment from GMM estimation in Section 5. These results are not causal, and cannot be used for either inference or counterfactuals. They represent the predictive power of various (possibly endogenous) variables in predicting a trade cost matrix that rationalizes leg-level containerized traffic flow.

C.1 Recovery of Predicted Trade Costs

Table A.4 shows the results of our estimation that predicts leg-level trade costs. Positive values for $\beta$ indicate increases in trade costs and negative values indicate decreases in trade cost. We find that distance increases trade cost and increased shipping traffic decreases trade costs, after fully accounting for the role of total trade values in $X$.

However, these estimates are not casual, and cannot be used for either inference or counterfactuals. They represent the power of various (including highly endogenous) variables in predicting a trade cost matrix that rationalizes leg-level containerized traffic flow. The fit of our predictions is highlighted in Figure VIII.\(^8\)

This analysis reflects the spirit of pure prediction and cannot satisfy the “Lucas Cri-

\(^8\)If we had more possible useful predictive variables, we would use a machine learning technique to tease out the best basis of variables to predict model-consistent trade costs.
Figure A.10: Trade Cost Estimates, All Legs

Notes: This map displays the recovered trade cost between all origins and destinations for containership legs in the AIS data. Lighter colors indicate lower trade costs.
Source: Authors’ calculations of AIS and Bill of Lading Data.

C.2 Additional Estimation Results

Below, estimated route costs are drawn. Thicker and lighter colors are lower-cost routes. Shorter and more heavily trafficked routes are the cheapest. The effect of scale is observable here: Syria to France is one of the highest cost legs, significantly higher than Singapore to Gibraltar, a much longer distance. Even among the subset of bilateral pairs for which we observe traffic, the triangle inequality is violated 280 times.

Figure A.11 plots bilateral incoming and outgoing trade costs for Singapore and Lebanon separately. Singapore is not only well-connected but both as an origin and destination has some of the cheapest legs. Singapore ships to Lebanon which has both

---

9This estimation follows Allen and Arkolakis (2019) and abstracts away from endogeneity and model mis-specification concerns.
fewer and shorter connections.

### C.3 Analysis of Trade costs

Figure A.12 plots country-level averages of the expected trade cost (from the B-matrix). Entrepôts such as Egypt, Panama, and (not visible) Singapore and Gibraltar have generally cheaper trade costs, as does China, due to the scale of shipping as well as access to nearby low-cost entrepôt (Korea, Singapore, and Japan).

Table A.5 reflects the log-linear relationship between our estimated trade cost $\tau$, aggregate bilateral trade values, and distance. These results highlight the reduced form relationships between these three variables, as well as the predictive power of our computed trade costs. Without origin or destination fixed effects, our trade costs alone can explain 29% variation of global trade. The logarithm of distance can account for less than 3%. We do not take this as a horse race, but rather indicative that these two measures are distinct: Our cost measures, $\tau$, measure network proximity and real shipping network relationships. Distance is a proxy for other, orthogonal variables which impact trade volumes as well.

#### Table A.5: The Relationship between Trade Volumes and Network-Consistent Trade Costs and Distance

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Log $\tau_{ij}^{-\theta}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log trade volumes</td>
<td>0.462 (0.0296)</td>
<td>0.449 (0.0308)</td>
<td>0.758 (0.0313)</td>
<td>0.524 (0.0287)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log dist</td>
<td>-0.752 (0.0980)</td>
<td>-0.406 (0.0909)</td>
<td>-1.325 (0.0621)</td>
<td>-0.652 (0.0599)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>12.67 (0.352)</td>
<td>14.76 (0.892)</td>
<td>16.16 (0.746)</td>
<td>15.63 (0.312)</td>
<td>19.87 (0.554)</td>
<td>19.11 (0.435)</td>
</tr>
<tr>
<td>Orig. Dest FE\s</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>22,985</td>
<td>22,985</td>
<td>22,985</td>
<td>22,985</td>
<td>22,985</td>
<td>22,985</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.286</td>
<td>0.029</td>
<td>0.294</td>
<td>0.760</td>
<td>0.751</td>
<td>0.770</td>
</tr>
</tbody>
</table>

Notes: Table presents regression coefficients from the regression of the natural log of trade volumes on the natural log of $\tau_{ij}^{-\theta}$, the natural log of model-estimated origin-destination trade costs raised to the trade elasticity, and the natural log of distance, the sea distance between the origin and destination measured in kilometers. Column (1)-(3) report results for cost and distance independently, then combined. Columns (4)-(6) rerun regressions in (1)-(3), respectively, adding origin and destination fixed effects.

### Appendix D General Equilibrium Model in Changes

To close our model, we adopt the Caliendo and Parro (2015) framework. A continuum of intermediate goods $\omega_n$ are used in the production of composite goods that are in turn used
Figure A.11: Trade Costs by Country

Notes: This map plots estimated link costs from Singapore in Panel (A) to Singapore in Panel (B), from Lebanon in Panel (C), and to Lebanon in Panel (D). Lighter colors indicate lower trade costs.
Figure A.12: Expected Trade Costs, Country Average

Notes: This figure plots average estimated origin-destination pair trade costs across all destinations for each country. Lighter colors indicate lower expected trade costs.
domestically both as final goods and as materials for intermediate production by firms in each industry \( n \). We assume there are three sectors \( (N = 3) \): containerized tradables \( c \), non-containerized tradables \( nc \), and nontradables \( nt \) \( (n \in \{c, nc, nt\}) \). Intermediates in the \( nt \) sector are only sourced domestically while \( \omega_{nc} \) and \( \omega_c \) goods are sourced internationally. Trade routes are modeled for all three sectors but we only consider transportation cost changes for the containerized sector \( \omega_c \).

**Consumption** In each country \( i \), consumers consume composite goods \( m_{in} \) from each sector \( n \), maximizing Cobb-Douglas utility.

\[
U_i = \prod_{n} m_{in}^{\eta_n}, \text{ where } \sum \eta_n = 1,
\]

where \( \eta_n \) is the Cobb-Douglas industry share, \( \sum_n \eta_n = 1 \).

**Intermediate goods production** The traded goods are intermediates, which are used in each country as building blocks for industry composite goods. In each country \( i \) and industry \( n \), firms produce a continuum of intermediate goods, indexed in each industry by \( \omega_n \in \Omega_n \). There are two types of input required for the production of \( \omega \): labor and composite goods. The production of intermediate goods across countries differs in their efficiency by a country-industry specific constant \( z_{in} \), a Ricardian technology. The production technology for intermediate \( \omega \) is

\[
q_{in}(\omega) = z_{in} [l_{in}]^{\gamma_{in}} \prod_{n'} m_{in'}^{\gamma_{in'}},
\]

where \( l_{in} \) is labor. \( \gamma_{in}^{n'} \) is share of materials from sector \( n' \) used in production of intermediate good \( \omega \), \( \gamma_{in} \) is share of value added, with \( \sum_{n'} \gamma_{in}^{n'} = 1 - \gamma_{in} \). The marginal cost of production for firms is

\[
c_{in} \equiv \frac{\Upsilon_{in} w_{i}^{\gamma_{in}} \prod_{n'} P_{in'}^{\gamma_{in'}}}{z_{in}}, \tag{15}
\]

where \( w_i \) is the wage in country \( i \), \( P_{in'} \) is the price of a composite good from sector \( n' \), and constant \( \Upsilon_{in} = \prod_{n'} \left( \frac{\gamma_{in}^{n'}}{\gamma_{in}} \right)^{\gamma_{in'}} (\gamma_{in})^{\gamma_{in}} \).

**Composite goods production** In each country \( i \), composite goods in industry \( n \) are produced using a CES aggregate of intermediates \( \Omega_n \), purchased and sold domestically at marginal cost. In traded industries, intermediates are sourced internationally from lowest-cost suppliers. Using the standard aggregation, the resulting price at \( j \) of the
composite in industry $n$ is expected to be the following (where $A_n$ is a constant):

$$P_{jn} = A_n \left[ \sum_{i=1}^{I} c_i^{-\theta_n} p_{ij}^{-\theta_n} \tau_{ijn}^{-\theta_n} \right].$$

(16)

The production costs in country $i$ and industry $n$ respond to a shock to a given $t_{kl}$ according to the equation:

$$\hat{c}_{in} = \hat{w}_{i}^{\gamma_{in}} \prod_{k=1}^{N} \hat{P}_{nk}^{\gamma_{ijn}}.$$

(17)

The change in the price of the composite intermediate good in country $i$ and industry $n$ relative to shock to $t_{kl}$ is:

$$\hat{P}_{in} = \left[ \sum_{i=1}^{J} \pi_{ijn} [\hat{r}_{ijn} \hat{c}_{in}]^{-\theta_n} \right]^{-1/\theta_n}.$$

(18)

Bilateral trade shares between $i$ and $j$ in industry $n$ will change according to standard changes through production and transport costs:

$$\hat{\pi}_{ijn} = \left[ \frac{\hat{c}_{in} \hat{r}_{ijn}}{\hat{P}_{in}} \right]^{-\theta_n}.$$

(19)

Trade volumes similarly adjust:

$$X_{in}' = \sum_{k=1}^{N} \gamma_{ink} \sum_{j=1}^{I} \frac{\hat{\pi}_{ijn}'}{1 + \kappa_{ijn}} X_{jk}' + \alpha_{in} I_{i}'.$$

(20)

Lastly, trade is balanced to a deficit shifter such that:

$$\sum_{n=1}^{N} \sum_{i=1}^{I} \frac{\hat{\pi}_{ijn}'}{1 + \kappa_{ijn}} X_{in}' - D_i = \sum_{n=1}^{N} \sum_{i=1}^{I} \frac{\hat{\pi}_{ijn}}{1 + \kappa_{ijn}} X_{jn},$$

(21)

where $I_{i}' = \hat{w}_{i} w_{i} L_i + \sum_{n=1}^{N} \sum_{i=1}^{I} \frac{\tau_{ijn}'}{1 + \kappa_{ijn}} X_{in}' + D_i.$

**Appendix E Counterfactual Results**

**E.1 Counterfactual Procedure**

Algorithm 1 describes the algorithm for finding the new equilibrium after an adjustment that induces an endogenous scale response in the network.

**E.2 Counterfactuals: Additional Figures**

Since some of the Asian entrepôts are smaller and harder to see on a global map, Figure A.13 zooms in on the welfare changes of Singapore, Hong Kong, and Taiwan as well
Algorithm 1 Scale Counterfactual Algorithm

1: procedure Welfare Change($X_0, \Xi_0, \hat{t}$) \hspace{1cm} \triangleright \text{Find a new equilibrium}
2: Initialize current trade flows $X_0$ and traffic $\Xi_0$
3: Initialize changes in cost fundamentals $\hat{\tau}$ \hspace{.5cm} \triangleright \text{Example: shipping distances changes}
4: Compute $A_0 = A(\Xi_0; \hat{\tau})$ \hspace{1cm} \triangleright \text{Following equation 12}
5: Compute $B_0 = (I - A_0)^{-1}$
6: Initialize difference $= \infty$, tolerance $= \epsilon$
7: while difference $< \text{tolerance}$ do
8: Update trade flows $X_1 = X(B_0)$ \hspace{1cm} \triangleright \text{Solving 8.1}
9: Update traffic $\Xi_1 = \Xi(X_1, A_0, B_0)$ \hspace{1cm} \triangleright \text{Following equation 11}
10: Update leg costs $A_1 = A(\Xi_1)$
11: Update trade costs $B_1 = (I - A_1)^{-1}$
12: Compute $\text{difference} = \Sigma_{ij}(B_1 - B_0)^2$
13: Update $A_0 = A_1$ and $B_0 = B_1$
14: Return final trade flows $X_1$
15: Compare welfare and price index changes between $X_1$ and $X_0$ \hspace{1cm} \triangleright \text{Solving 8.1}

as their surrounding countries as a result of the opening of the Arctic Passage. In the baseline scenario in Panel (A), we see that these entrepôts have a direct welfare increase from the passage opening since they have direct routes to Northern European countries and North America. When allowing for indirect trade in Panel (B), the neighboring countries of these entrepôts see an increase in welfare because they are now able to benefit from using these entrepôts to trade with the Northern European countries and North America. When allowing for scale economies to amplify effects in Panel (C), the entrepôts and their neighboring countries are going to benefit even further as a results of this indirect trade.

The concentration of welfare gains in entrepôts from this counterfactual highlights a novel source of agglomeration—scale economies in transportation and transport networks can help contribute to and shape entrepôts. This is further explored in our third counterfactual.

Figure A.15 shows the impact of our two counterfactual cases on the UK’s 20 largest trading partners in welfare percent changes. Black bars show the impact of increased non-transportation trade frictions with the UK. Grey bars show the impact with scale effects changing transportation costs through the UK. All partners experience outsized losses due to scale economies. Most of these losses come through increase trade costs in the Netherlands and Belgium, which far from benefiting from our counterfactual, lose because of decreased volumes as well. Ireland in particular, which our microdata tells us sends 50% of goods to the US through the UK, experiences large additional losses.
Figure A.13: Welfare Changes on Asian Entrepôts - Arctic Passage

(A) Only Directly Affected Routes
(B) Full Trade Network Effects
(C) Full Trade Network Effects and Scale Economies

Notes: These three plots are a magnified part of figure XII to show the percent change in welfare (the relative price index) for a subset of Asian Entrepôts in our dataset. Darker reds reflect a greater increase and blue represents no change. White represents omitted countries. Panel (A) reflects changes if we only allow trade costs to decrease on routes whose distance is directly reduced to the Arctic Passage. Panel (B) reflects changes if we allow all countries to indirectly access the Arctic Passage through the trade network. Panel (C) allows for the endogenous network response to scale economies.

Finally, figure A.16 reports global trade volume changes under our two cases. These results largely mirror our welfare results in the main text.
Figure A.14: Export Volume Changes - Arctic Passage

(A) Only Directly Affected Routes

(B) Full Trade Network Effects

(C) Full Trade Network Effects and Scale Economies

Notes: These three plots show the percent change in exports from all countries in our dataset. Darker reds reflects a greater increase in exports. White represents omitted countries. Panel (A) reflects changes if we only allow trade costs to decrease on routes whose distance is directly reduced to the Arctic Passage. Panel (B) reflects changes if we allow all countries to indirectly access the Arctic Passage through the trade network. Panel (C) allows for the endogenous network response to scale economies.
Figure A.15: Welfare Changes - Brexit - Largest Trading Partners

Notes: Bars show the percent change in welfare (the relative price index) of a simulated 5% increase in trading costs with the United Kingdom the largest 15 trading partners. The first bar reflects changes if shipping costs remain constant, reflecting only welfare changes due to changes in prices. The second bar allows for the endogenous network adjustment to scale economies.

Figure A.16: Export Volume Changes - Brexit

(A) Trade Cost Change, No Network Scale Effects

(B) Full Trade Network Effects and Scale Economies

Notes: These two plots show the percent change in exports of a simulated 5% increase in trading costs with the United Kingdom for all countries in our dataset. Darker reds reflects a greater increase. White represents omitted countries. Panel (A) reflects changes if shipping costs remain constant, reflecting only trade changes due to changes in prices. Panel (B) allows for the endogenous network adjustment to scale economies.