Geography, Search Frictions and Trade Costs

Giulia Brancaccio[*], Myrto Kalouptsidi[†], Theodore Papageorgiou[‡§]

Princeton University, Harvard University, CEPR and NBER, McGill University

March 2017

PRELIMINARY

We build a framework that models the behavior of both exporters and transportation agents (ships); its spatial equilibrium determines world trade costs and flows. Our framework has two novel features: (i) trade costs are endogenous and determined jointly with trade flows and as a result they depend on the entire network of trade linkages across countries; (ii) search frictions between exporters and ships limit trade. The model features geography, search frictions, and forward-looking optimizing ships and exporters. We collect a unique dataset of shipping contracts, global vessel movements from satellites and sea weather conditions. The data reveal large trade imbalances and asymmetric trade costs. We provide an empirical strategy to flexibly obtain the matching process between ships and exporters in a setup where searching exporters are unobserved and the researcher takes no stance on the presence of search frictions. Our estimated framework is then used to address a number of questions: What are the world trade elasticities with respect to transport costs? How do shocks propagate through the network of countries? We consider the impact of a slow-down in China as well as the opening of the Northwest Passage. Finally, we quantify the trade lost due to search frictions.

[*]Princeton University, Department of Economics, Fisher Hall, Princeton, NJ 08544, USA, giulia@princeton.edu

[†]Harvard University, Department of Economics, Littauer Center, Cambridge MA 02138, USA, myrto@fas.harvard.edu

[‡]McGill University, Department of Economics, 855 Sherbrooke St West, Room 519, Montreal, QC, H3A 2T7, Canada, theodore.papageorgiou@mcgill.ca

[§]We gratefully acknowledge financial assistance from the Griswold Center for Economic Policy Studies at Princeton University. We are thankful to Costas Arkolakis, Aris Christou, Manolis Galenianos, Ken Hendricks, Fabian Lange, Marc Melitz, Ariel Pakes for their many helpful comments.

1 Introduction

About 90% of international trade is carried by the global shipping industry.[1] To export, an exporter has to find an available vessel and contract for a voyage and a price.[2] In turn, the ship is optimally choosing its travels in search of cargo, thinking about its future options. This spatial equilibrium between exporters and ships determines the trade costs that different countries face, as well as the trade flows between different regions of the world. What is the role of geography (i.e. country locations, natural inheritance of goods) in determining trade costs and flows? Is the matching process between exporters and ships efficient? How does ship behavior affect the behavior of exporters?

In this paper, we build a framework that models the behavior of both exporters and transportation agents (ships). The spatial equilibrium of this model determines world trade. Our framework has two novel features. First, trade costs are endogenous and determined jointly with trade flows. As such, trade costs depend on the entire network of trade linkages across countries, rather than just the bilateral (distance between) trading partners. We show that equilibrium trade costs provide a novel link to understand trade patterns. Second, search frictions between exporters and ships can limit trade. We estimate this model using unique data on shipping contracts and global vessel movements; our empirical strategy allows us to flexibly recover the matching process between ships and exporters, as well as obtain the main model primitives of interest (ship costs, exporter valuations and costs). We use our framework to tackle a number of questions of interest: What is the size of world trade elasticities with respect to transportation costs? How do shocks propagate through the world; for instance, how would a Chinese slow-down trickle through the network of countries, or how would the opening of the Northwest Passage affect trade costs and trade flows? Finally, what is the loss due to search frictions between exporters and ships?

We focus on dry bulk ships, which carry mostly commodities (grain, ore, coal, etc.) and whereby an exporter hires the entire vessel for a specific voyage. We construct a unique database that allows us to study international trade through the lens of the shipping industry. In particular, we combine (i) a dataset of shipping spot contracts involving the ship, dates, trip origin/destination and price; (ii) satellite data reporting every 6 minutes ships' positions, speed, and level of draft; the draft allows us to distinguish loaded from empty (ballast) trips; (iii) weather conditions at world oceans.

We first use our data to uncover some novel facts about (i) world trade flows; (ii) trade costs; (iii) search frictions between ships and exporters. First, the satellite data reveals that most countries are either

[1] Source: International Chamber of Shipping.

[2] Different segments of shipping function differently. In this paper we focus on bulk shipping, where exporters of bulk commodities fill up an entire vessel and hire it for a single trip (much like a taxi driver or a rental car); see Section 2.2.

net importers or net exporters of commodities transferred by dry bulk vessels. Related to this, at any point in time a staggering 40% of ships are traveling without a cargo (ballast). We show that this natural trade imbalance, often overlooked in the trade literature, is a key driver of trade costs, which in turn feedbacks into the trade patterns. Second, we find substantial asymmetry of shipping prices across locations: as an example, shipping from China to Australia costs about 7,500 US dollars per day, while shipping from Australia to China costs substantially more, at 10,000 US dollars per day. Part of this asymmetry can be explained by the ships' matching opportunities in the destination: all else equal, the prospect of having to ballast after the destination port leads to higher prices. To illustrate using the example above, Australia is an exporter to China, whose imports in raw materials have grown dramatically in recent decades (to build infrastructure), whereas its exports are considerably lower. As a result, prices to ship goods to China are considerably higher, as matching opportunities are limited there. Third, we find evidence of search frictions between ships and exporters. More specifically, using our satellite data we find that at a given time, in most countries there are simultaneously ships arriving empty, while other ships are leaving empty, even though ships are homogeneous. This suggests social wastefulness: the cargo picked up by the ship that arrived empty, could have been transferred by the ship that left empty instead. In addition, again consistent with the presence of search frictions, shipping prices exhibit substantial dispersion, within a time/origin/destination triplet, indicating that the law of one price does not hold in this market.

We build a dynamic spatial search model of the global shipping industry. The model features three key ingredients: (i) geography; (ii) search frictions; (iii) forward-looking ships and exporters that optimally choose their travels and exporting destinations respectively. Geography enters the model through different trip durations across different ports. In addition to this natural geography, locations differ in their economic geography, namely their natural inheritance in commodities of different value. Exporters and ships match randomly, and search frictions prevent the matching of all possible pairs. Prices are determined by Nash bargaining. Ships are homogeneous and forward looking: when negotiating a trip they also take into account matching opportunities at the destination. If unmatched, ships decide whether to wait at their location or travel empty (ballast) someplace else, taking into account their expected discounted stream of profits at each location.

We estimate our model using the collected data. There are two sets of core model primitives: (i) the matching function between ships and exporters, as well as the global distribution of searching exporters; and (ii) exporter valuations and exporting costs, as well as ship sailing and port waiting costs. Using data on the number of ships and matches, as well as the weather, we obtain the former, while using data on shipping prices, as well as ship ballast choices and exporter destination choices, we estimate the latter.

In particular, we adopt a novel approach to flexibly recover both the matching function, as well the searching exporters. A sizable literature has estimated matching functions in several different contexts (e.g. labor markets, taxicabs).[3] For instance, in labor markets, one can use data on unemployed workers, vacancies and matches to recover the underlying market matching function. In the market for taxi rides, one observes taxis and their rides, but not hailing passengers; in recent work, Frechette, Lizzeri and Salz (2015) and Buchholz (2015) have used such data, coupled with a parametric assumption on the matching function, to recover the passengers. In our case, similar to the taxi industry, we observe ships and matches, but not searching exporters. Our approach draws from the literature on nonparametric identification (Matzkin (2003)). In particular, we show how to recover nonparametrically the matching function, as well as the global distribution of exporters relying on the joint density of matches and ships, as well as sea weather as an instrument that exogenously changes arriving ships. To provide intuition, consider the following test for search frictions: weather shocks exogenously shift ship arrivals at port; in markets with more exporters than ships, this should not affect matches unless there are search frictions. We show that here, matches are indeed affected by weather shocks, which both suggests again that search frictions are present and allows us to obtain the curvature of the matching function. In summary, we make two contributions. First, unlike the existing literature, we do not take a stance on the presence of search frictions. When one side of the market (in this case exporters) is unobserved or mismeasured, it is difficult to discern whether search frictions are present. Second, we avoid parametric restrictions on the matching function; this is important, as here parametric restrictions are directly linked to welfare implications.

The remaining primitives are obtained from ship and exporter choices and prices. In particular, first we recover ship sailing and port wait costs via Maximum Likelihood, based on their optimal ballast choice probabilities. As ships are forward looking this is a dynamic discrete choice problem and we solve for ship value functions inside the likelihood (similar to Rust (1987)). Then, we obtain exporter valuations directly from prices: once ship primitives are known, we employ the surplus sharing condition derived from Nash bargaining to back out each valuation corresponding to each individual contract price. We can thus obtain the distribution of valuations (conditional on an origin and a destination) nonparametrically. Finally, we use trade flows (loaded trips) to recover exporter costs by destination.

Our model provides a unique framework to study how international trade costs and flows are determined. The novel feature here is that trade costs are endogenous and our counterfactuals reveal that

[3] See Petrongolo and Pissarides (2001) for a survey.

taking this into account is important. For instance, a decline in the cost of sailing, leads to an increase in exports in some regions, but a decline in others: the decline in the sailing cost not only increases the value of a match between a ship and an exporter, but also makes ballasting less costly for ships. As a result ships are now less likely to wait in port after unloading a cargo and more likely to travel empty towards large exporters. On net this pushes up the exports of net exporters like Brazil and reduces the exports of net importers like China.

We also illustrate the importance of trade networks and market conditions in neighboring countries, by considering a slow-down in China. In particular, the reallocation of ships over space amplifies the effect of the slow-down in neighboring markets. Besides the direct effect to countries whose exports relied heavily on the growth of the Chinese economy, our model points out that there is a secondary effect driven by the reduced supply of ships in that region of the world: the many ships that ended up in China are no longer around. This impacts negatively both China's own exports (import-export complementarity), but also neighboring countries' toward which these ships would ballast. Large exporters further away such as Brazil and North America benefit from this increased ship availability and increase their exports to other destinations.

We also show how the opening of the Northwest Passage affects trade not only for countries whose routes are directly affected, but also other countries through network linkages. Finally, we also quantify the trade lost due to search frictions and showcase how in counterfactual world with no frictions, there is a shift to trading more valuable commodities.

Related Literature We relate equally to three broad strands of literature: (i) trade and geography; (ii) search and matching; (iii) industry dynamics.

First, our paper endogenizes trade costs and so naturally it relates to the large literature in international trade studying the importance of trade costs in explaining trade flows between countries (e.g. Anderson and Van Wincoop (2003)). In much of this literature, trade costs are treated as "residuals" that explain the gap between actual bilateral trade flows between countries and predicted trade flows conditional on variables such as size, distance, common border/language, importer/exporter fixed effect, etc. Here, we consider what happens to trade flows when transport prices (a substantial component of trade costs) are determined in equilibrium, jointly with trade flows. In addition, we document important features of trade costs, which are often not taken into account in standard trade models; for instance, trade costs are asymmetric and not proportional to distance. Waugh (2010) has argued that asymmetric trade costs are necessary to explain some empirical regularities regarding trade flows and prices across rich and poor countries. We also contribute to a literature that has considered the role and features of the (container) shipping industry; e.g. Hummels and Skiba (2004) explore the relationship of exporter product prices at different destinations and shipping costs; Hummels, Lugovskyy and Skiba (2008) explore market power in container shipping; Wong (2017) incorporates container shipping prices featuring a "round-trip" effect in a trade model.[4] Finally, recent work has explored the matching of importers and exporters under frictions (Eaton, Jinkins, Tybout and Xu (2016)).

Our paper is also related to both old and new work on the role of geography in international trade (Krugman (1991), Head and Meyer (2004), Allen and Arkolakis (2014)), as well as the impact of transportation infrastructure and networks (Donaldson (2012), Allen and Arkolakis (2016). We extend this literature by exploring how the location of each country and all its neighbors interact with the functioning of transportation agents and thus shape up trade costs and flows. We illustrate that both a country's location (i.e. distances from all other countries) and raw material inheritance are key features of the equilibrium. The latter is, to our understanding, novel and it ends up being particularly important as it leads to imbalanced trade; this natural asymmetry is crucial in determining a country's exports and trade costs. One feature of the trade literature that our paper is missing is that we do not determine commodity prices and input prices in equilibrium. The latter may be reasonable as we focus only on commodities and so wages and capital prices may be taken as exogenous. The former would require additional data on exporters and is an interesting avenue for future research.

Second, our paper relates to the search and matching literature (see Rogerson, Shimer and Wright (2005) for a survey of the literature). On one hand, our model is essentially a search model in the spirit of Mortensen and Pissarides (1994) where firms and workers (randomly) meet subject to search frictions and Nash bargain over a wage. An important addition in our case is the spatial nature of our setup: there are several interconnected markets at which agents (ships) can search. Lagos (2000, 2003) and Buchholz (2016) have also adopted similar spatial search models for taxi cabs; an important difference here is that prices are set in equilibrium, while in the taxi market prices are exogenously set by regulation. In our setup this is important, since by endogenizing the trade costs we can consider how they change trade flows in the different counterfactuals.

[4]Wong (2017) is also exploring the impact of endogenous trade costs, but her approach is different and complementary to ours. She is looking at container, rather than bulk shipping and uses market-level US rather than micro-level world data. Moreover, the paper mostly explores how allowing for the "round-trip effect" in bilateral trade costs changes the theoretical predictions of a standard trade model. The "round-trip effect" refers to the fact that containerships have specific itineraries and are thus forced to make return trips.

Third, we relate to the literature on industry dynamics (Hopenhayn (1992), Ericson and Pakes (1995), etc.). Consistent with this research agenda, we study the long-run industry equilibrium properties, in our case the spatial distribution of ships and exporters. Moreover, our empirical methodology borrows from the literature on the estimation of dynamic setups (e.g. Rust (1987), Aguirregabiria and Mira (2007), Bajari, Benkard and Levin (2007), Pakes, Ostrovsky and Berry (2007)) in matching conditional choice probabilities that involve value functions (applications include Ryan (2012) and Collard-Wexler (2013)). Buchholz (2016) and Frechette, Lizzeri and Salz (2016) also explore dynamic decisions in the context of taxi cabs' search choices and shift ending decisions respectively. Finally, Kalouptsidi (2014) has also looked at the shipping industry, albeit at the entry decisions of shipowners and the resulting investment cycles in new ships.

The rest of the paper is structured as follows. Section 2 provides a description of the industry and the datasets used. Section 3 presents a number of facts that document the importance of geography, establish price patterns and explore the presence of search frictions. Section 4 describes the model. Section 5 lays out our empirical strategy, while Section 6 presents the estimation results. In Section 7 we provide the counterfactuals and in Section 8 we conclude. The Appendix provides additional tables and figures, proofs to our propositions, as well as further data and estimation details.

2 Industry and Data Description

2.1 Trade in Dry Bulk Commodities

Dry bulk shipping involves vessels designed to carry a homogeneous unpacked dry cargo, for individual shippers on non-scheduled routes. The entire cargo belongs to one cargo owner (the exporter). Bulk carriers operate like taxi cabs: a specific cargo is transported individually in a specific ship, for a trip between a single origin and a single destination. Dry bulk shipping involves mostly commodities, such as iron ore, steel, coal, bauxite, phosphates, but also grain, sugar, chemicals, lumber and wood chips.[5]

There are four different categories of dry bulk carriers based on size: Handysize (10,000–40,000 DWT), Handymax (40,000–60,000 DWT), Panamax (60,000–100,000 DWT) and Capesize (larger than 100,000 DWT). The industry is unconcentrated, consisting of a large number of small shipowning firms (see Kalouptsidi (2014)): the maximum fleet share is around 4% (3% for Handysize, 7% for Handymax, 3%

[5]It is worth noting that bulk ships are very different from containerships, which operate like buses: containerships carry cargos (mostly manufactured goods) from many different cargo owners in container boxes, along prespecified itineraries. It is technologically impossible to substitute bulk with container shipping.

for Panamax, 5% for Capesize), while the firm size distribution features a large tail of small shipowners. Moreover, shipping services are largely perceived as homogeneous. In his lifetime, a shipowner will contract with hundreds of different exporters, carry a multitude of different products and visit numerous countries; we discuss this further in Section 3.

Trips are realized through individual contracts: shipowners have vessels for hire, exporters have cargo to transport and brokers put the deal together.[6] Ships carry at most one freight at a time: the exporter fills up the hired ship with its own cargo. In this paper, we focus on spot contracts and in particular the so called "trip-charters", in which the shipowner is paid in a per day rate.[7] The exporter who hires the ship is responsible for the trip costs (e.g. fueling), while the shipowner incurs the remaining ship costs (e.g. crew, maintenance, repairs, insurance).

2.2 Data

In this paper we combine a number of different datasets to create a unique database of ship contracts and movements.

First, we employ a dataset of dry bulk shipping contracts, from 2001 to 2016, collected by Clarksons Research, a major global shipbroking firm. An observation is a transaction between a shipowner and a charterer, for transportation of a specific cargo, on specific dates, from a specific origin to a specific destination. We observe the name and age of the vessel, the identity of the charterer who hires the ship, the contract signing date, the agreed loading and unloading dates, the agreed upon trip price in dollars per day, as well as some information on the origin and destination (see below and the Appendix).

Second, we use satellite AIS (Automatic Identification System) data from exactEarth Ltd (from now on, EE) from 2009 to 2016, that we match to the Clarksons dataset and thus track the movements of these vessels. AIS transceivers automatically broadcast information, such as the ships' positions (longitude and latitude), speed, and level of draft (i.e. the vertical distance between the waterline and the bottom of the ship's hull), at regular intervals of at most 6 minutes. The level of draft allows us to determine whether a ship is loaded or not at any point in time. We use the EE dataset to construct the history of trips (empty or loaded) for each ship.

The EE dataset is useful for two reasons. First, it provides more accurate information on origins and

[6] Some shipowners have in-house brokers while others collaborate with independent brokers. Oftentimes, multiple brokers are involved in a negotiation from both sides.

[7] Trip-charters are the most common type of contract. Long-term contracts ("time-charters"), however, do exist: on average, about 10% of contracts signed are long-term. Interestingly, though, ships in long-term contracts, are often "relet" in a series of spot contracts, suggesting that arbitrage is possible.

destinations for the Clarksons contracts, since we can find the exact location of the ship under contract on the loading and unloading dates specified by the contract (see the Appendix for details on how trips and stops are constructed). Second, it allows us to determine when ships decide to travel someplace empty (ballast) to find cargo and when ships stay put in their current location.

We also use the ERA-Interim archive, from CMWF (European Centre for Medium-Range Weather Forecasts), to collect global data on daily sea weather. This allows us to construct weekly data on the wind speed (in each direction) on a 0.75° grid across all oceans. Finally, we employ several time series from Clarksons on e.g. the total fleet and fuel prices, as well as country-level imports/exports, production and commodity prices from numerous sources (e.g. UNCTAD, FAO, IEA, Comtrade).

Summary statistics Our final dataset involves 5,410 ships between 2012 and 2016.[8] We end up with 7,652 shipping contracts, for which we know the price, as well as the exact origin and destination. As shown in Table 1, the average price is 10,000 dollars per day (or 140,000 dollars if we take the trip duration into account), with substantial variation (the standard deviation is 5,000 dollars per day; more on this in Section 3). We have 233,580 ship-week observations at which the ship decides to either ballast someplace empty or stay at its current location. Loaded trips last on average 2.3 weeks, with little variation within an origin-destination pair. Ballast trips last less, 1.72 weeks on average. Contracts are signed on average 6 days prior to the loading date.[9] Upon signing a contract, about 42% of ships are already in the loading port. Ships that do not find a cargo, remain in their current port with probability 76%. Clarksons reports the product carried in a small subsample of the contract dataset (about 20%). The main commodity categories are grain (23%), iron ore (20%), coal (20%), alumina ore (6%), chemicals/fertilizer (6%) and minor bulks like wood chips and sands. Finally, it is worth noting that our sample period is one of low shipping demand and severe ship oversupply due to high ship investment between 2005 and 2008 (see Kalouptsidi (2014, 2016)).

[8]We only use contracts during the period that we also have satellite data. Moreover, to estimate the matching function we drop the first two years (2009-2011), as during this period, the geographic regions covered by satellites is growing as new satellites are launched.

[9]As practitioners say, "a ship is not a train"; it is not possible for a ship to promise too far in advance arrival to load at a specific port, due to the uncertainties of prior travels (predominantly weather conditions, but also port/canal congestion, port strikes, etc.).

	Mean	St. Dev	Median	Min	Max
Contract price per day (10^5 US dollars)	0.105	0.054	0.095	0.01	0.7
Contract trip price (10^5 US dollars)	1.417	9	1.17	0.07	8.315
Contracts per ship	2.108	1.445	2	1	12
Loaded trip duration (weeks)	2.50	2.06	2.02	0.14	10.41
Empty trip (ballast) duration (weeks)	1.74	1.63	1.29	0.09	8.98
Days between contract signing and loading date	6.11	6.686	4	0	28
Ship size (1,000 DWT)	74	35	74	13	455
Prob of ship staying at port conditional on being unmatched	0.77	0.12	0.76	0.59	0.95

Table 1: Sample summary statistics.

3 Facts

In this section, we present a number of novel facts about the bulk shipping industry and world seaborne trade more generally. We provide some key features of vessel movements, we document geographic patterns of trade, we explore the nature of trade costs (i.e. shipping prices), and finally we discuss some descriptive evidence of search frictions in this market. Our findings here guide the model formulation in Section 4.

3.1 Trade Flows, Geography and Ballasting

Figure 1 plots the message count from each location of the globe during a 10 day period. It reveals that some of the most frequent voyages are between Australia and China, Brazil and China, as well as Northwest America and China. This graph does not distinguish between loaded and ballast voyages; Figure 15 in the Appendix presents a chart of the loaded trips between world regions. The most popular loaded trip is from Australia to China (around 5% of loaded trips in our dataset). The most popular ballast trip is the reverse, from China to Australia (5.7% of ballast trips). It is no accident that China dominates the observed flows: in recent years, Chinese growth has led to massive imports of raw materials for industrial expansion and infrastructure building. In turn, Australia, Brazil and Northwest America are big exporters of minerals, grain, coal, etc.

China's example suggests that global trade features a substantial imbalance, partly owing to the natural inheritance of different countries in raw materials. Indeed, we next illustrate that most countries in the world are either net importers or net exporters of commodities. Figure 2 depicts the difference between the number of ships departing loaded minus the number of ships arriving loaded, over the sum of the two.

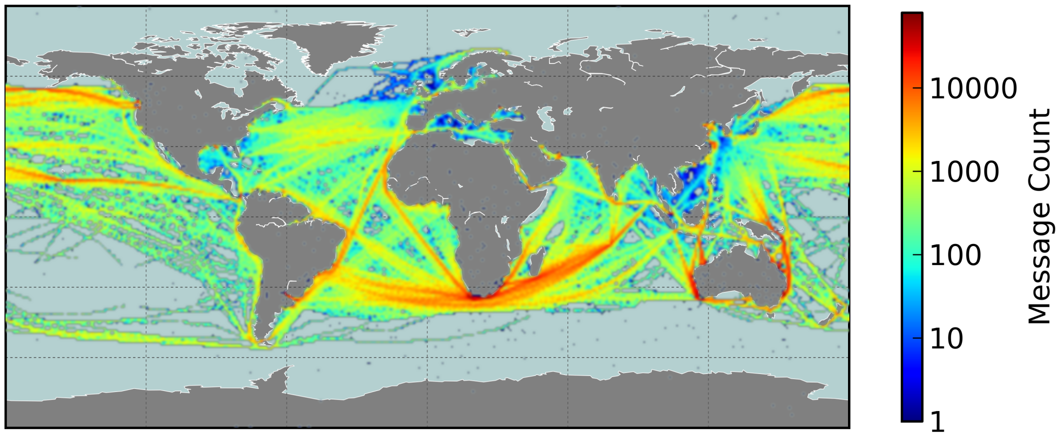


Figure 1: Number of AIS signals received in a ten day period.

A positive ratio indicates that a country is a net exporter of dry bulk commodities, while a negative ratio suggests that the country is a net importer; a ratio close to zero implies balanced trade. As shown, trade flows in most countries in the world are considerably imbalanced, rather than balanced. Australia, Brazil and Northwest America are big exporters, whereas China and Europe are big importers. This feature of trade is not unique to raw materials; container shipping exhibits similar asymmetries; the direction of the imbalance, however, may be different (e.g. China is a big exporter in containers rather than a big importer).

As a consequence of the imbalanced nature of international trade, ships spend much of their time traveling ballast, i.e. without cargo. The fraction of the miles a ship travels ballast over the total miles traveled while present in our dataset is about 40%.

The imbalanced nature of trade, although an important empirical feature, is often overlooked in the trade literature. In this paper, we do not assume balanced trade and in fact this asymmetry is a key driver of our model and empirical findings.

3.2 Trade Costs (Shipping Prices)

We next turn to the nature of trade costs that exporters face. A quick inspection of the data reveals that trade costs are asymmetric: for instance, a trip from China to Australia costs on average 7,500 dollars per day, while a trip from Australia to China costs substantially more, at 10,000 dollars per day on average, net of fuel costs.[10] In fact, most trips exhibit substantial asymmetry: the average ratio of the price from

[10]This price asymmetry has been documented also in container shipping; see for instance Wong (2016) and references therein.

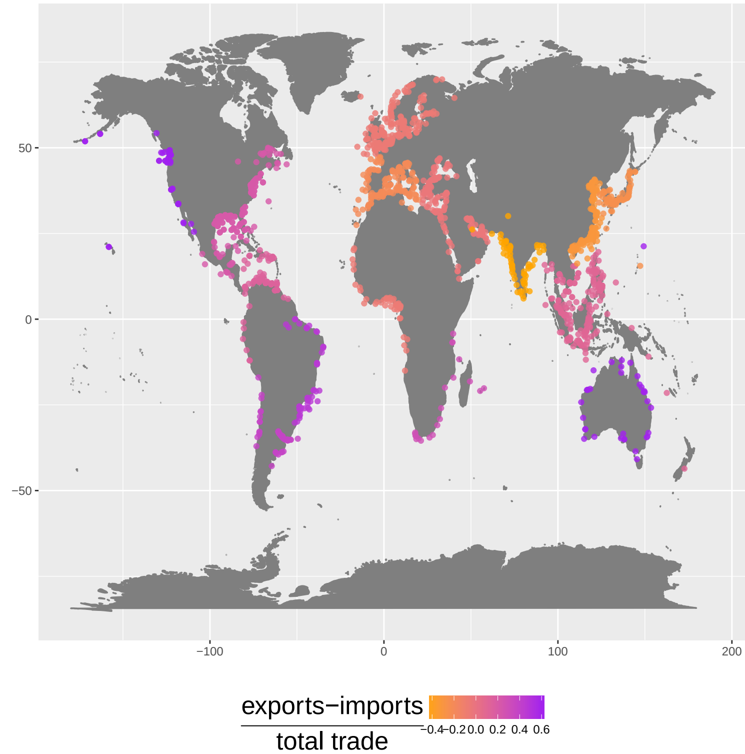


Figure 2: Trade Imbalance. Difference between exports (ships leaving loaded) and imports (ships leaving empty) over total trade (all ships).

origin i to destination j to the price from j to i (highest over lowest), is equal to 1.6 and can be as high as 4.1.[11] This empirical pattern suggests that bilateral distance is not the only determinant of trade costs and travel direction matters.

More generally, shipping prices are determined by the entire network of origins and destinations, rather than just bilateral distances. Although this will be properly documented via our structural model of Section 4 later on, we present here some regressions to identify some determinants of prices. The first column presents the results of a log-price regression on ship types and country of origin fixed effects which already account for 66% of the price variation, suggesting that geography is important in explaining trade costs. This is not surprising, as local demand and supply at the origin (products that the location produces, the number of ships present etc.) are clearly important determinants of shipping prices. The second column adds destination fixed effects and, interestingly, the fraction of price variation increases. This suggests that ships may demand a premium to travel towards a destination with low exports, to compensate for the difficulty of finding a new cargo originating from that destination (and vice versa for destinations with high exports). To control more directly for this effect, in the third column of Table 2 instead of a destination fixed effect, we include (i) the probability that a ship leaves ballast from the destination specified in the contract; and (ii) conditional on leaving ballast, the miles a ship travels ballast

[11]This is calculated using the 15 geographical regions employed in our empirical exercise below (see Section 6), to guarantee sufficient data per origin/destination.

from the destination on average. As expected, we find that both variables are positive and significant and lead to substantially higher prices. Indeed, a 1% increase in the average distance traveled ballast after the destination, is associated to an increase prices by 0.16%; while exporting to a destination where the probability of a ballast trip afterwards is ten percentage points higher, costs on average 2.3% more.[12]

Finally, this is a good point to emphasize that ship heterogeneity is not an important consideration. First, even in our small dataset, the majority of ships (80%) are seen carrying at least 2 of the 3 main products (coal, ore and grain), which suggests that ships do not specialize on certain products. Similarly, we observe that most ships travel to most regions, suggesting that they do not specialize geographically either. Ship fixed effects have no explanatory power in either price regressions or ballast probability regressions. Finally, this is consistent with much of the evidence provided in Kalouptsidi (2014); for instance hedonic regressions of ship resale prices suggest that unobserved heterogeneity is not an important consideration.

3.3 Search Frictions

In this section we present some descriptive evidence suggesting that search frictions inhibit the matching of all available ships and exporters. Overall, it is not straightforward to know a priori whether a market (here, the market for sea transport) suffers from search frictions. In labor markets, where search frictions are generally thought to be present, there are two main empirical regularities that are offered as evidence: (i) the coexistence of unemployed workers and vacant firms; and (ii) wage inequality among observationally identical workers. In this section we show that a similar set of empirical conditions hold in shipping, suggesting that search frictions are present here as well.

We first consider whether there is evidence of unrealized matches. While in labor markets the coexistence of unemployed workers and vacant firms in labor markets is strong evidence of the presence of search frictions, our data reports only ships and matches, not searching exporters, so we cannot make a similar argument; we can however consider a different moment that has a very similar flavor.

Figure 3 displays the weekly number of ships that arrive in a country empty and load, as well as the number of ships that leave a country empty, for the case of Norway and Chile. Both countries are net exporters. In Norway, several ships arrive empty and load, but almost no ship departs empty. In Chile, however, the picture is quite different: we see both ships that arrive empty to load, as well as ships that depart empty. In other words, an empty ship arrived, picked up cargo, while at the same time another ship departed empty. This fact suggests the presence of search frictions in Chile: why didn't the ship that

[12] To confirm that the result is not driven by the different composition of products exported toward different destinations, the last column of Table 2 also controls for the type of product carried for the subsample of contracts reporting the product.

	I	II	III	IV
		log(price)		
Handyamax	-0.148**	-0.136**	-0.123**	0.027
	(0.014)	(0.014)	(0.014)	(0.120)
Handysize	-0.397**	-0.330**	-0.343**	-0.209**
	(0.017)	(0.018)	(0.017)	(0.124)
Panamax	-0.223**	-0.214**	-0.212**	-0.117
	(0.013)	(0.013)	(0.013)	(0.119)
Coal				0.088**
				(0.045)
Fertilizer				0.245**
				(0.051)
Grain				0.131**
				(0.048)
Ore				0.124**
				(0.045)
Steel				0.135**
				(0.049)
Probability of ballast			0.234**	0.556**
			(0.030)	(0.081)
Average duration of ballast trip (log)			0.166**	0.065**
			(0.014)	(0.032)
Constant	10.304**	10.284**	9.127**	8.915**
	(0.068)	(0.103)	(0.099)	(0.408)
Destination FE	No	Yes	No	No
Origin FE	Yes	Yes	Yes	Yes
Product FE	No	No	No	Yes
Quarter FE	Yes	Yes	Yes	Yes
Obs	11,014	11,014	11,011	1,662
R^2	0.663	0.694	0.674	0.664

$^{**}p < 0.05$, $^{*}p < 0.1$

Table 2: Trade costs (shipping price) regressions.

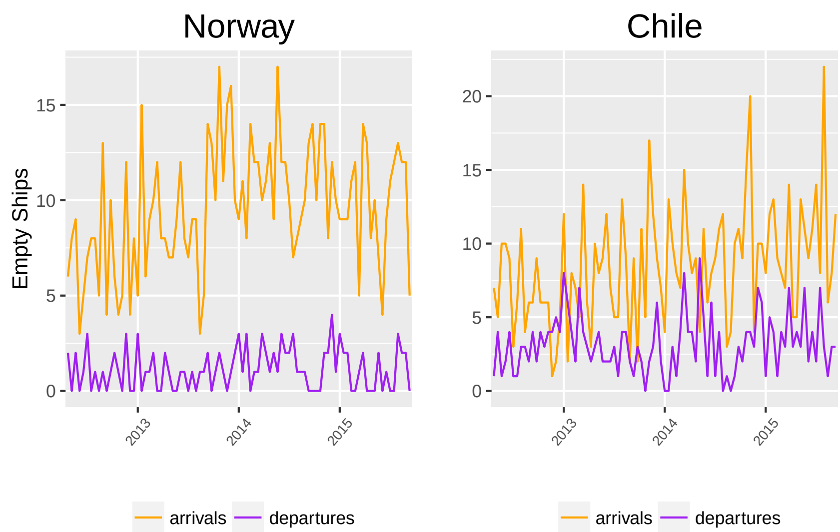


Figure 3: Flow of ships arriving empty and ships leaving empty within 2 weeks intervals

departed empty, pick up the cargo, instead of having another ship arrive from elsewhere to pick it up?[13]

We next perform this exercise for all exporting countries, by computing the bi-weekly ratio of the incoming empty and loading ships over the outgoing empty ships for a given country. If there are no search frictions this ratio should be close to zero for countries that are net exporters. However, as shown in Figure 4 which depicts the histogram of these ratios, most countries are more similar to Chile, than Norway. Indeed, the average ratio is well above zero and for some countries it is even above 0.5. In addition, this pattern has proved very robust in a number of dimensions.[14] While in labor markets, as some researchers have argued, observed or unobserved heterogeneity may explain part of the co-existence of unemployment and vacancies (a vacancy for a chemical engineer may not be of interest to a high school dropout), in this market the importance of heterogeneity is much more limited: as discussed above ships are widely considered to offer homogeneous services and do not specialize geographically or in terms of products.

[13]Recall that contracts are signed 6 days on average before the loading date and that with probability 42% the ship is already in the loading port.

[14]This figure is robust to alternative market definitions, time periods and ship types. Figure 16 in the Appendix presents this histogram by ship type: Capesize vessels exhibit somewhat larger mass towards zero, consistent with the higher concentration of charterers and the ships' ability to approach fewer ports. The figure is also the same if done by port rather than country. As mentioned above, ships tend to carry all products; thus we do not believe this pattern is explained by product switching. Labor contracts are usually about 5-8 months long and the crew flies between their home and the relevant port. To control for repairs we remove stops longer than 6 weeks. Finally, we only consider as "ships arriving empty" the ships arriving empty and sailing full toward another market, and we consider as "ships leaving empty" ships sailing empty toward a different country. This definition also implies that refueling cannot explain the fact either; although there are very small differences in fuel prices across space anyway (less than 10%).

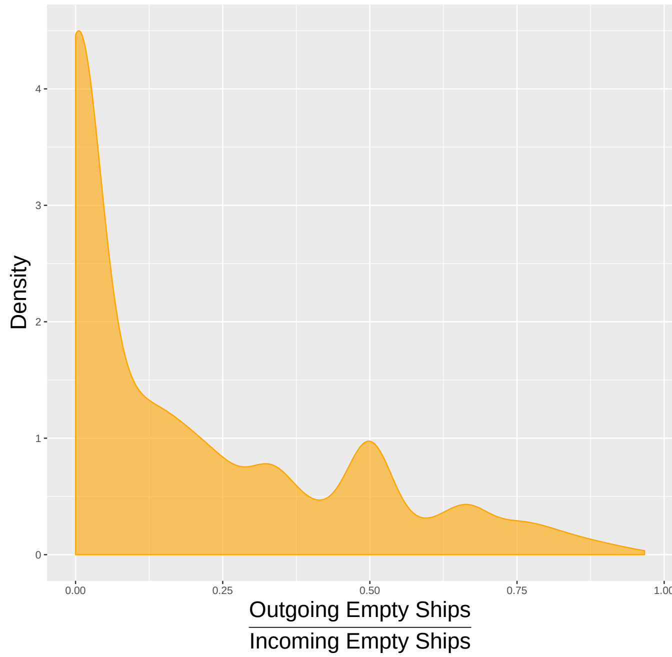


Figure 4: Histogram of the ratio of incoming and outgoing empty ships in net exporting countries.

Again inspired by the labor literature, we investigate a second aspect of this market that is suggestive of search frictions: dispersion in prices. In markets with no frictions, the law of one price holds so that there is a single price for the same service. This does not hold in labor markets, where there is substantial wage dispersion among workers who are observationally identical. This observation has generated a substantial and widely influential literature on frictional wage inequality, i.e. wage inequality that is driven by search frictions.[15]

In the shipping market a similar empirical regularity is present. As we already saw in Table 2 there is substantial price dispersion in shipping spot contracts. More specifically, at best we can account for about 70% of price variation, controlling for ship size, as well as quarter, origin and destination fixed effects. Moreover, the coefficient of variation of prices within a given quarter, origin and destination triplet is about 30% (23%) on average (median).[16] For instance, in the most popular trip from Australia to China the coefficient of variation is on average 37% and ranges from 21% to 55% across quarters.

In addition, it is worth noting that the type of product carried affects the price paid and overall more valuable goods lead to higher contracted prices.[17] In the absence of frictions, if there are more ships than

[15] See for instance Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), Mortensen (2005) and references therein.

[16] Again here we use the 15 geographical regions employed in our empirical exercise below (see Section 6), to guarantee sufficient data per origin/destination. Results are robust if we only keep quarter-origin-destination triplets with more than 10 observations.

[17] Recall that we have significantly fewer contract observations for which the product is reported, so the results in the last column of Table 2 are less precise.

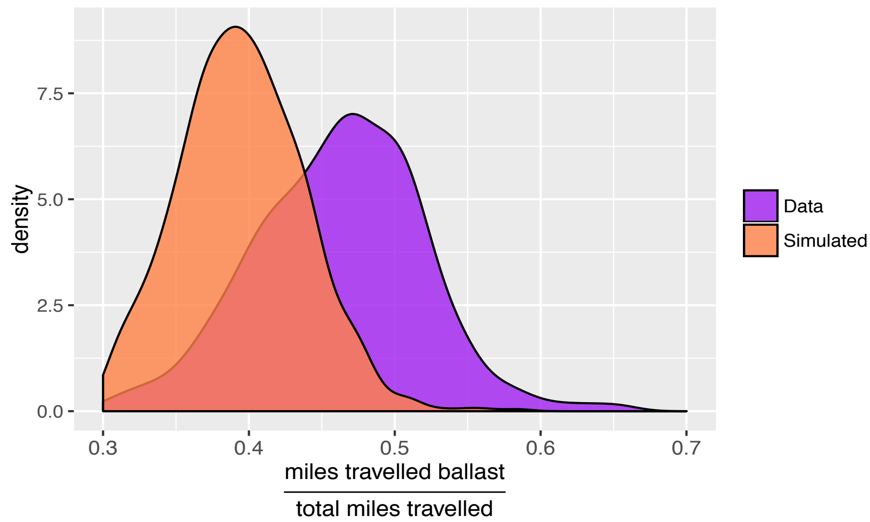


Figure 5: Histogram of fraction of miles traveled ballast. Actual data and simulation.

exporters (as is the case arguably during our sample period), we would expect prices to be bid down to the marginal ship's costs. In contrast, in markets with frictions and bilateral bargaining, this is no longer true and as shown formally in the model of Section 4, the buyer's outside option affects the price she pays. In the context of the shipping market, our model predicts that exporters with higher valuations pay more when there are matching frictions, consistent with evidence shown in Table 2.

Finally, we perform a simple "dispatcher" simulation to quantify the extent of search frictions. We simply assign every exporter to the ship that is geographically closest to it. If there are no frictions, the simulation should lead to an allocation not far from the observed. That is not the case however. Figure 5 plots the histogram of the fraction of distance a ship ballasts in the simulation and compares it to the actual one. We find that in the data ships ballast significantly more. Note that this simulation does not employ the model of Section 4, so ships do not optimally choose where to search; we return to this in the counterfactuals of Section 7.[18]

We revisit search frictions in Section 5.1, where we both provide further reduced form evidence, and we estimate a non-parametric matching function, thus flexibly measuring the extent of search frictions. We close this section by briefly discussing what the nature of search frictions may be. The mere existence of brokers suggests that search frictions have been an issue in this market. Information frictions may still prevail though; when a ship is searching in a certain geographical region (e.g. east Americas) her broker

[18]The initial condition in the simulation is the ships' first registered position. Then, we assign every observed loaded trip to the ship that was closest to the trip's originating port, rather than the ship that actually executed the trip. We force ships to wait for cargo at the port at which they arrived loaded and only ballast when assigned to a cargo, rather than choose to ballast towards a more desirable destination.

may not "meet" the broker of a specific exporter in one of the ports. It can also be the case that a broker receives many messages and he may not be able to sort through them efficiently in time, similarly to an unemployed worker who cannot sort through all available vacancies that are posted on the various job finding websites. In a market with a small number of large exporters, it might be easier to for them to be matched to the existing ships. Consistent with this, we find that price dispersion is negatively correlated with the Herfindahl Index of the observed ship charterers.

4 Model

We next introduce a dynamic spatial search model of the global shipping industry. Geography enters the model through different trip durations across different locations. There are two types of agents: exporters (or freights) and ships. Exporters choose whether and where to export, subject to different exporting costs by destination. Following the search and matching literature we model frictions between exporters and ships using a matching function, essentially a "black box" that captures the implications of frictions in the trading environment, in a parsimonious fashion (Pissarides (2000)). Prices are determined by Nash bargaining. Ships are homogeneous and forward looking: when negotiating a trip they also take into account matching opportunities at the destination. If unmatched, ships decide whether to wait at their location or ballast someplace else, taking into account their expected discounted stream of profits at each location. We describe first the environment of the model and then the behavior of the agents.

4.1 Environment

Exporters Time is discrete. There are I locations/markets and each location is denoted by $i \in \{1, 2, ..., I\}$. At each location i and period t, there are f_{it} freights that need to be delivered to another location. We use the words freight, exporter and cargo interchangeably. Exporters have heterogeneous valuations, v, received upon delivery to their destination. The valuation of a freight going from i to j is drawn from the distribution F^v_{ij} with mean μ_{ij}. Unmatched freights survive with probability $\delta > 0$. Each period, at each location i, $\mathcal{E}_i$ potential exporters decide whether and where to export, and pay production and export costs, κ_{ij}.

Ships There are S ships in the world.[19,20] Ships are homogeneous, risk neutral and have discount factor β. In every period, a ship is either traveling loaded or ballast, from some location i to some location j, or it is at port in some region i. A ship at port in location i incurs cost c_i^u, while a ship sailing from i to j incurs cost c_{ij}^s. The duration of a trip between region i and region j is stochastic: a ship arrives at region j within a period with probability ξ_{ij}, so that the average duration of the trip is $1/\xi_{ij}$.[21]

Matching Freights can only be delivered to their destination by ships and each ship can carry (at most) one freight. In each region, freights and ships need to search for each other and search is random. Every period a ship (freight) can match with at most one freight (ship). The number of new matches at time t and market i is given by the matching function,

$$m_{it} = m_i\left(f_{it}, s_{it}\right)$$

where f_{it} is the number of unmatched freights searching in i and s_{it} is the number of unmatched ships searching in i. The matching function $m_i\left(.\right)$ is increasing in both arguments. Let λ_{it} denote the probability with which an unmatched ship in location i meets a freight; $\lambda_{it} = m_{it}/s_{it}$. Similarly, let λ_{it}^f denote the probability with which an unmatched freight meets a ship; $\lambda_{it}^f = m_{it}/f_{it}$.

Search frictions generate rents to realized matches that are split between the freight and the ship via the price-setting mechanism. We assume that the price paid to the ship delivering a freight of valuation v from region i to destination j, π_{ijv}, is determined by generalized Nash bargaining, with $\gamma \in (0,1)$ denoting the exporter's bargaining power. The price is paid upfront and the ship commits to begin its voyage immediately to region j.[22]

Timing The timing of each period is as follows:

[19] We follow Kalouptsidi (2014) and assume constant returns to scale so that a shipowner is a ship. Similarly, a freight owner is a freight, so that he does not choose the export tonnage. We also ignore the different ship sizes in the model; as described in Section 6 estimating the model separately for each ship type yields similar results.

[20] In this paper, we do not model ship entry or exit. Exit is overall very small, while due to long construction lags in shipbuilding, the fleet is close to being fixed in the short run (see Kalouptsidi (2014, 2016)).

[21] Recall that there is little variation in the duration of a trip from i to j (see Section 2.2). The above assumption preserves model tractability without changing any of the results compared to a model with deterministic trip duration. In particular, ships are risk neutral so it does not affect their value functions. In addition, the law of large numbers implies that in the steady state the number of ships that arrive/depart from a market every period is equal to that in a model with deterministic trip duration.

[22] Note that since all unmatched ships at a port are homogeneous, an exporter (v, j) either forms a match with any ship it meets or cannot agree to a mutually acceptable price with any ship in a steady state. In what follows we restrict attention to freights that are acceptable and thus always sign a contract once a meeting takes place. More generally, since no failed negotiations are observed, it would not be possible to allow for this feature in the empirical analysis; during our sample period, which is characterized by severe ship oversupply, we do not expect freights to be rejected by ships.

1. In each market i, existing ships and exporters match.

2. In each market i, unmatched ships pay a cost to stay in port and draw additive iid preference shocks $\epsilon = [\epsilon_1, ..., \epsilon_I]$ from distribution F^ϵ for each market and decide whether to (i) stay in their current region and wait for freight; or (ii) ballast toward some destination j (i.e. sail empty and look for cargo in j)[23]

3. Unmatched ships from each market that decided to ballast away begin traveling to their chosen destination. All ships already traveling from i to j arrive at j with probability ξ_{ij}.[24] Existing exporters disappear with probability $1 - \delta$.

4. In each market i, $\mathcal{E}_i$ potential exporters decide whether and to which destination to export. The exporters that do enter the market draw their valuations from F^v_{ij}, and join the pool of searching exporters next period.

States and Transitions The state variable of a ship in region i includes its current location i, as well as the vector (s_t, f_t, s^w_t) where $s_t = [s_{1t}, ..., s_{It}]$, $f_t = [f_{1t}, ..., f_{It}]$ and the $I^2 - I$ dimensional vector s^w_t, with entries s^w_{tij}, denotes the number of ships traveling from i to j in period t. The state variable of an existing exporter in i includes his location i, valuation v and destination j (chosen upon entry), as well as the vector (s_t, f_t, s^w_t). Exporters in market i at time t transition as follows:

$$f_{it+1} = \delta \left(f_{it} - m_i \left(s_{it}, f_{it} \right) \right) + d_i \tag{1}$$

with d_i the (endogenous) flow of new freights (discussed below). Ships at location i transition as follows:

$$s_{it+1} = \left(s_{it} - m_i \left(s_{it}, f_{it} \right) \right) P_{ii} + \sum_{j \neq i} \xi_{ji} s^w_{jit} \tag{2}$$

where P_{ij} is the probability of an unmatched ship ballasting from i to j (determined endogenously from ship choices, see below). In words, out of s_{it} searching ships, m_{it} ships get matched and leave i, while out of the ships that did not find a match, P_{ii} percent chooses to remain at i rather than ballast away;

[23]Ships do indeed sail towards a destination without having already signed a contract (recall that the average trip duration is 2-3 weeks, while contracts are signed on average 6 days in advance). According to practitioners, this is mostly due to the uncertainties of traveling and arrival times (see also Footnote 9), though it is overall intuitive- why would the ship wait in a region where chances of reloading are slim (unlike airplanes, it is very costly to let the ship wait in a port)? Note also, that in our empirical exercise we will focus on broad geographical regions so that one can think of ships as searching there (although even at the port level, 42% of ships are already in the loading port when signing a contract).

[24]Note that a ship that begins traveling today cannot arrive in the same period.

moreover, out of the ships traveling towards i, ξ_{ij} percent arrive. Finally, ships that are traveling from i to j, s^w_{ijt} evolve as follows:

$$s^w_{ijt+1} = (1 - \xi_{ij})\, s^w_{ijt} + P_{ij}\,(s_{it} - m_i\,(s_{it}, f_{it})) + G_{ij} m_i\,(s_{it}, f_{it}) \tag{3}$$

where G_{ij} is the probability of going loaded from i to j (determined endogenously from exporter choices, see below). In words, ξ_{ij} percent of the traveling ships arrive, P_{ij} percent of ships that remained unmatched in location i chose to ballast to j and finally, G_{ij} of ships matched in i depart loaded to j.

4.2 Behavior

We derive the optimal behavior of exporters and ships, as well as the equilibrium prices. In this paper, we consider the steady state of this model.[25] This assumption makes sense for the data at hand, which covers a period (2012-2016) that is uniformly characterized by ship oversupply and relatively low demand for shipping services. More specifically, we assume that agents view the spatial distribution of ships and freights as fixed and make decisions based on their steady-state values: it does not feel unreasonable that they ignore aggregate long-run shocks when making weekly ballasting decisions.

Ships Let W_{ij} denote the value of a ship traveling from i to j (empty or loaded). Then:

$$W_{ij} = -c^s_{ij} + \xi_{ij}\beta U_j + (1 - \xi_{ij})\beta W_{ij} \tag{4}$$

In words, the ship that is traveling from i to j, pays per period cost of transit c^s_{ij}; with probability ξ_{ij} it arrives at its destination j, where it will begin unmatched with value U_j defined below; finally, with the remaining probability $(1 - \xi_{ij})$ the ship does not arrive at its destination and keeps traveling. The ship arrives at destination j after $1/\xi_{ij}$ periods on average.

Consider now a ship in market i. This ship obtains:

$$U_i = -c^u_i + \lambda_i E_{i,v} V_{ijv} + (1 - \lambda_i) J_i \tag{5}$$

In words, the ship is matched with probability λ_i, in which case it obtains the value of a matched ship V_{ijv} defined below. The ship takes expectation over the type of freight it meets, i.e. its value and destination.

[25] Given that a ship can travel to most ports in the world in under a month, any transition dynamics to a new steady state will be very short. This is very convenient in our counterfactual analysis of Section 7, where we are able to compare steady states without worrying about the transition dynamics.

With the remaining probability, $1-\lambda_i$, the ship does not find a freight and it obtains the value J_i also defined below. Note that the probability of finding a cargo is given by the matching function and depends on the number of ships and cargos available in that market. Finally, the ship pays the per period port cost c_i^u.

If matched with an exporter with destination j and value v, the ship receives the agreed upon price and begins traveling, so that:

$$V_{ijv} = \pi_{ijv} + W_{ij} \tag{6}$$

If the ship remains unmatched, it faces the choice of either staying at i and matching there the following period with probability λ_i, or ballasting away from i in search of better opportunities. In the latter case, the ship can choose among all possible destinations. In particular, if unmatched, the ship receives a vector of iid preference shocks ϵ, for all possible destinations j, as well as its current location i, drawn from a double exponential distribution F^ϵ, with mean zero and variance σ. The unmatched ship's value function is :

$$J_i(\epsilon) = \max\left\{\beta U_i + \sigma\epsilon_i, \max_{j\neq i} W_{ij} + \sigma\epsilon_j\right\} \tag{7}$$

and let:

$$J_i \equiv E_\epsilon J_i(\epsilon) = \sigma \log\left(\exp\frac{\beta U_i}{\sigma} + \sum_{j\neq i}\exp\frac{W_{ij}}{\sigma}\right) + \sigma\gamma^{euler}$$

where γ^{euler} is the Euler constant.[26] In words, if the ship stays in its current market i, it searches in i again the following period, thus obtaining value U_i; otherwise the ship chooses another market and begins its trip there. Let $P_{ii} = \Pr(i|i)$ denote the probability that a ship in location i chooses to remain there, while $P_{ij} = \Pr(j|i)$ denote the probability that a ship at location i chooses to ballast to j. Since ϵ is iid and and follows a double exponential, we can write:

$$P_{ii} = \frac{\exp\left(\beta U_i/\sigma\right)}{\exp\left(\beta U_i/\sigma\right) + \sum_{l\neq i}\exp\left(W_{il}/\sigma\right)} \tag{8}$$

and

$$P_{ij} = \frac{\exp\left(W_{ij}/\sigma\right)}{\exp\left(\beta U_i/\sigma\right) + \sum_{l\neq i}\exp\left(W_{il}/\sigma\right)}. \tag{9}$$

[26]The formula for the ex ante value function $J_i = E_\epsilon J_i(\epsilon)$ is the closed form expression for the expectation of the maximum over multiple choices, and is obtained by integrating $J_i(\epsilon)$ over the double exponential distribution of ϵ.

Exporters We start with existing exporters and then consider exporter entry. The value of an unmatched exporter in market i, with destination j and valuation v is,

$$J_{ijv}^{f} = \lambda_i^f V_{ijv}^{f} + \left(1 - \lambda_i^f\right) \beta\delta J_{ijv}^{f}. \tag{10}$$

In words, the exporter gets matched with probability λ_i^f and gets the value of being matched, V_{ijv}^{f} (defined below); with the remaining probability he remains unmatched; in that case, he survives with probability δ and waits for next period.

An exporter that is matched in market i receives value:

$$V_{ijv}^{f} = v - \pi_{ijv}, \tag{11}$$

in words, he obtains his delivery value, v and pays the agreed price, π_{ijv}, which is discussed below.

There are $\mathcal{E}_i$ ex ante homogeneous potential exporters in market i in each period. Each potential entrant n in market i, makes a discrete choice between not exporting, as well as which destination j to export to, subject to production and exporting costs κ_{ij}, as well as random preference shocks ϵ_{nj}^{f}, all j which are distributed according to a double exponential distribution.[27] Upon deciding to become an existing exporter in i with destination j, the entrant draws a valuation v from F_{ij}^{v}. Therefore, potential entrant n solves:

$$J_i^{ef} = \max\left\{\epsilon_{n0}^{f},\ \max_{j\neq i}\left\{\beta E_v J_{ijv}^{f} - \kappa_{ij} + \epsilon_{nj}^{f}\right\}\right\}$$

where we denote with 0 the (outside) option of not exporting and normalize the payoff in that case to zero.[28]

Potential exporter n's behavior is given by the choice probabilities:

$$\tilde{G}_{ij} \equiv \frac{\exp\left(J_{ij}^{f} - \kappa_{ij}\right)}{1 + \sum_{l\neq i} \exp\left(J_{il}^{f} - \kappa_{il}\right)} \tag{12}$$

and

$$\tilde{G}_{i0} \equiv \frac{1}{1 + \sum_{l\neq i} \exp\left(J_{il}^{f} - \kappa_{il}\right)} \tag{13}$$

[27] Since each exporter is one shipment, they do not export to multiple destinations. Adding independent exporter data to further study the intensive margins of exporting (e.g. multiple destinations, size of shipments) is an interesting avenue for future research.

[28] It is also possible to allow the potential exporters to know their valuations (across destinations) before making their exporting choice. It would simply make estimation computationally somewhat more demanding.

where $J_{ij}^{f} \equiv E_v J_{ijv}^{f}$.

The number of entrant exporters in i ends up being:

$$d_i = \mathcal{E}_i \left(1 - \tilde{G}_{i0}\right) \tag{14}$$

It is worth noting that the distribution of export destinations conditional on exporting used is given by

$$G_{ij} \equiv \frac{\tilde{G}_{ij}}{1 - \tilde{G}_{i0}} \tag{15}$$

This is the distribution that ships employ when forming expectations over the potential matches in different markets in equation (5).[29]

Prices As discussed above, the rents generated by a match between a freight and a ship, are split via Nash bargaining. This implies the usual surplus sharing condition:

$$\gamma \left(V_{ijv} - J_i\right) = (1 - \gamma) \left(V_{ijv}^{f} - J_{ijv}^{f}\right) \tag{16}$$

We use the above condition to solve out for the equilibrium price π_{ijv}, in the following lemma:

Lemma 1. *The price agreed upon between a ship and an exporter with valuation v and destination j in location i is given by:*

$$\pi_{ijv} = \frac{\gamma \left(1 - \beta\delta \left(1 - \lambda_i^{f}\right)\right)}{1 - \beta\delta \left(1 - \gamma\lambda_i^{f}\right)} \left(J_i - W_{ij}\right) + \frac{(1 - \gamma)(1 - \beta\delta)}{1 - \beta\delta \left(1 - \gamma\lambda_i^{f}\right)} v \tag{17}$$

Proof. Substitute V_{ijv}, V_{ijv}^{f}, J_{ijv}^{f} and U_i in (16). □

In other words, the price is a linear combination of the exporter's value, v, and the difference between the ship's value of delivering the freight, W_{ij}, and its outside option, J_i.

It is worth noting that if the value of a ship traveling from i to j, W_{ij}, is low, then, all else equal, the price is higher. It is worth remembering that $W_{ij} = -\frac{c_s}{1-(1-\xi_{ij})\beta} + \frac{\xi_{ij}\beta}{1-(1-\xi_{ij})\beta} U_j$, includes both the conditions at the destinations through U_j, but also the importance of distance captured by ξ_{ij}. In other words destinations that are unappealing to ships because there are few freights there and the probability of

[29] In this paper, we assume away from the determination of commodity prices; in other words, we take μ_{ij} to be exogenous. μ_{ij} is meant to capture the revenue of the exporter in i that sells her commodity in j. To determine this object in equilibrium would require additional data on exporters; this is an interesting avenue for future research.

ballasting afterwards is high, would command higher prices. This is consistent with the evidence presented in Table 2. The same holds for destinations that are further away (low ξ_{ij}). In addition, note that the price between i and j depends on all countries rather than just i and j through the outside option of the ship, J_i. Therefore, in a model of endogenous trade costs, the shipping price from region i to region j depends on all countries, rather than just i and j. This feature will be important in counterfactuals, as shown in Section 7.

In addition, exporters that have a higher value, v, pay higher prices, again consistent with evidence in Table 2. As discussed in Section 3.3, this is true because of search frictions and the fact that the law of one price no longer holds.

Steady State Equilibrium We next define equilibrium for this model and prove that a steady state equilibrium exists.

Definition. A steady state equilibrium, (s^*, f^*, s^{w*}), is a distribution of ships and exporters over markets, that satisfies the following conditions:

(i) Ships' optimal behavior, $P_{ij}(s^*, f^*)$ follows (8) and (9) and expectations employ (15).

(ii) Potential exporters' behavior, $\tilde{G}_{ij}(s^*, f^*)$, follows (12) and (13) and entrants are determined from (14).

(iii) Prices are determined by Nash bargaining, according to (17).

(iv) Ships and freights satisfy the following steady state equations:

$$s_i^* = \sum_j P_{ji}(s^*, f^*)\left(s_j^* - m_j\left(s_j^*, f_j^*\right)\right) + \sum_{j \neq i} G_{ji}(s^*, f^*)\, m_j\left(s_j^*, f_j^*\right) \tag{18}$$

$$f_i^* = \delta\left(f_i^* - m_i\left(s_i^*, f_i^*\right)\right) + E_i\left(1 - \tilde{G}_{i0}(s^*, f^*)\right) \tag{19}$$

$$s_{ij}^{w*} = \frac{1}{\xi_{ij}}\left(P_{ij}(s^*, f^*)\left(s_i^* - m_i^*\right) + G_{ij}(s^*, f^*)\, m_i^*\right)$$

Proposition 1. *Suppose that the matching function is continuous, the preference shocks ϵ and ϵ^f have full support, $\mathcal{E}_i$ and S are finite and $f_i \leq \mathcal{E}_i/(1-\delta)$. Then, an equilibrium exists.*

Proof. See the Appendix. □

Finally, we characterize the steady state trade flow between regions i and j, which is equal to:

$$\mathcal{E}_i\tilde{G}_{ij} = \mathcal{E}_i\frac{\exp\left(J_{ij}^f - \kappa_{ij}\right)}{1+\sum_{l\neq i}\exp\left(J_{il}^f - \kappa_{il}\right)} = \mathcal{E}_i\frac{\exp\left(\alpha_i\left(\mu_{ij}-\pi_{ij}\right)-\kappa_{ij}\right)}{1+\sum_{l\neq i}\exp\left(\alpha_i\left(\mu_{il}-\pi_{il}\right)-\kappa_{il}\right)}$$

where $\alpha_i = \lambda_i^f / \left(1-\beta\delta\left(1-\lambda_i^f\right)\right)$, since in the steady state:

$$J_{ij}^f \equiv E_v J_{ijv}^f = E_v\frac{\lambda_i^f\left(v-\pi_{ijv}\right)}{1-\beta\delta\left(1-\lambda_i^f\right)} = \frac{\lambda_i^f\left(\mu_{ij}-\pi_{ij}\right)}{1-\beta\delta\left(1-\lambda_i^f\right)}$$

where $\pi_{ij} = E_v\pi_{ijv}$ and μ_{ij} is the average valuation from i to j. This equation is reminiscent of a "gravity equation"; it delivers the trade flow (in quantity rather than value) from i to j as a function of two components. First, the primitives $\{\lambda_i^f, \mu_{ij}, \kappa_{ij}, \mathcal{E}_i\}$ not just for i, j but for all countries; this is reminiscent of the analysis in Anderson and Van Wincoop (2003) who show that the gravity equation in a trade model needs to include a country's overall trade disposition. Second, the endogenous trade costs, π_{ij}, for all j. The important addition here is that the trade flow depends on all countries through the endogenous trade cost π_{ij}; indeed, recall from the price equation (17) that the shipping price π_{ij} depends on all possible destinations from i through the outside option of the ship that can ballast anywhere. Therefore, the network of countries affects bilateral trade flows not just directly through the primitives, but also indirectly through the endogenous trade cost.

5 Empirical Strategy

In this section we lay out the empirical strategy followed to estimate the model of Section 4. The main model primitives we wish to recover are: the matching function and searching exporters, the ship costs of traveling and waiting at different ports, as well as the distribution of exporter valuations and their costs. We exploit data on the numbers of ships, the numbers of matches and weather conditions to recover the matching function and exporters. Then, using the ballast decisions of ships, the prices and the flows of loaded trips, we recover the remaining primitives. Here, we describe this empirical strategy, while Section 6 presents the results.

5.1 Matching Function Estimation

A sizable literature has estimated matching functions in several different contexts (e.g. labor markets, marriage markets, taxicabs).[30] For instance, in labor markets, one can use data on unemployed workers, job vacancies and matches to recover the underlying matching function. In the market for taxi rides, one observes taxis and their rides, but not hailing passengers; in recent work, Buchholz (2015), and Frechette, Lizzeri and Salz (2016) have used such data, coupled with a "parametric" assumption on the matching function to recover the hailing passengers.[31]

Similar to the taxi market, we observe ships and matches, but not searching exporters. Here, we adopt a novel approach to simultaneously recover both exporters, as well as a nonparametric matching function. Our approach makes two contributions to the literature. First, we do not take a stance on the presence and magnitude of search frictions in the industry. Why is this important? Consider the case of no search frictions, so that the matching function is

$$m_{it} = \min\left(s_{it}, f_{it}\right) \tag{20}$$

In words, all potential matches are realized. In contrast, if there are search frictions, we have:

$$m_{it} = m\left(s_{it}, f_{it}\right) \leq \min\left(s_{it}, f_{it}\right) \tag{21}$$

so that some potential matches are not realized. If one side of the market is unobserved (in this case freights) or mismeasured (arguably, in labor markets searching workers are imperfectly measured due to on-the-job search, unobserved search effort or noisy participation decisions; Lange and Papageorgiou (in progress)) it is not straightforward to differentiate (20) from (21). Indeed, when a ship/taxi is traveling empty is it because no exporter/passenger was searching or because an exporter/passenger was there but did not get to meet the ship/taxi due to frictions? Our approach, allows us to disentangle the two.

Our second contribution is to avoid imposing parametric restrictions on the matching function. The literature has imposed functional forms such as the Cobb-Douglas. The desire to be non-parametric is not just "stylistic" when it comes to matching functions: parametric restrictions are directly linked to welfare implications. For instance, it has been shown in a wide class of labor market models, that the condition

[30] See Petrongolo and Pissarides (2001) for a survey in the context of labor markets.

[31] Buchholz (2016) assumes an "urn-ball" matching function. Frechette, Lizzeri and Salz (2015) construct a numerical simulation of taxi drivers that randomly meet passengers over a grid that resembles Manhattan; this spatial simulation essentially corresponds to the matching function, and can be inverted to recover hailing passengers.

for constrained efficiency depends crucially on the elasticity of the matching function with respect to the search input (Hosios (1990)). In most of the matching function estimation literature this elasticity has been restricted to be constant.[32]

Our approach borrows from the literature on nonparametric identification (Matzkin (2003)) and delivers both a nonparametric matching function, as well as the number of searching exporters. Roughly, the method leverages (i) an invertibility assumption between matches and freights, (ii) the observed relationship between ships and matches, (iii) an instrument that shocks the ships exogenously, and (iv) a restriction on the matching function that allows us to disentangle monotonic transformations. To provide some intuition, we outline a simple version of the methodology. We then formalize the argument. We refer the interested reader to Matzkin (2003) for further details.

Suppose that (i) the matching function $m_i(s_i, f_i)$ is continuous and strictly increasing in f_i and, (ii) that s_i is independent of f_i. The first assumption is natural in our context: increasing the number of freights should lead to more matches, all else equal. The second assumption will prove useful in presenting the estimation methodology, but it is likely not valid in our case, as the spatial distribution of ships and freights is determined jointly in equilibrium; we relax this assumption below.

Let $F_{m|s}$ denote the distribution of matches conditional on ships, and F_f the distribution of freights, f. Then at a given point (s_{it}, f_{it}, m_{it}) we have:

$$\begin{aligned} F_{m|s}\left(m_{it}|s_{it}\right) &= \Pr\left(m\left(s,f\right) \leq m_{it}|s_{it}\right) \\ \text{monotonicity} &= \Pr\left(f \leq m^{-1}\left(s, m_{it}\right)|s_{it}\right) \\ \text{independence} &= \Pr\left(f \leq m^{-1}\left(s_{it}, m_{it}\right)\right) \\ &= F_f\left(f_{it}\right) \end{aligned} \tag{22}$$

In words, the conditional distribution of matches (outcome) on ships (observed covariate) at a point (m_{it}, s_{it}) is equal to the distribution of freights at the corresponding (unobserved) point f_{it}. Equation (22) is our main relationship for the identification and estimation of both freights and the matching function. However, (22) alone is not sufficient: it is not possible to distinguish monotonic transformations of f and $m(\cdot)$. To do so, a restriction on either the distribution F_f or the matching function is required. In this paper we assume that the matching function is homogeneous of degree one, so that: $m_i(\alpha s_i, \alpha f_i) = \alpha m_i(s_i, f_i)$,

[32] Petrongolo and Pissarides (2001)

all $\alpha > 0$.[33] The intuition behind the identification argument is as follows: the correlation between s_i and m_i informs us on $\partial m_i(s_i, f_i)/\partial s_i$, since the sensitivity of matches to changes in ships is observed; then, due to homogeneity, this derivative also delivers the derivative $\partial m_i(s_i, f_i)/\partial f_i$; and once these derivatives are known, so is the matching function, which can now be inverted to provide the freights as well.[34]

Finally, as mentioned above, independence of ships and freights is not a natural assumption in our setting. We do, however, have a plausibly valid instrument: sea weather shifts the arrival of ships in a port without affecting the number of freights. Hence, we instrument with weather conditions. We assume that ships s are a function of the instrument z and a shock η such that ships and freights are independent conditional on η. This allows us to modify (22) by conditioning on $\hat{\eta}$ as well as on s to obtain the distribution of freights.

Proposition 2 formalizes these arguments:

Proposition 2. *(i) Suppose that $m(s, f)$ is continuous, (positively) homogeneous of degree 1 and strictly increasing in f. Suppose further that s and f are independent. Finally, suppose that m is known for a specific pair (s^*, f^*) so that $m^* = m(s^*, f^*)$, with $m^* \neq 0$. Then, the function $m(\cdot)$ is identified.*
(ii) Suppose there exists an instrument z such that

$$s = h(z, \eta) \tag{23}$$

with z independent from (f, η). Assume that proper conditions hold so that η can be uniquely recovered from (23). Then, s and f are conditionally independent given η. The distribution F_f and the matching function can be recovered from:

$$F_{f|\eta}(\phi) = F_{m|s=\frac{\phi}{f^*}s^*,\eta}\left(\frac{\phi}{f^*}m^*\right)$$

$$F_f(\phi) = \int F_{f|\eta}(\phi) f_\eta(\eta) d\eta$$

$$m(s, \phi) = \int F^{-1}_{m|s,\eta}(F_{f|\eta}(\phi)) f_\eta(\eta) d\eta$$

In Section 6, we implement this methodology separately for each market and present the results.[35]

[33] In the labor search literature, where there have numerous studies estimating matching functions, most estimates find support for constant returns to scale (see Petrongolo and Pissarides, 2001). Given that the nature of search frictions is not that different (in both cases it is a shortcut for information frictions about which ships/freights may be available), we consider this a reasonable starting point.

[34] We could alternatively impose an assumption on the distribution F_f. For example, if we assume that F_f is uniform on $[0, 1]$, we can use (22) to recover f_{it} pointwise by the conditional distribution of m on s; once freights are recovered, we also instantly know the (inverse) matching function. Bajari and Benkard (2005) employ this methodology to nonparametrically estimate hedonic price equations and unobserved product quality in the case of personal computers.

[35] We interpret the observed time-series variation as driven by short-run deviations from the steady state values. Ships and

5.2 Ship Costs

We next turn to the ships' sailing cost, c^s_{ij}, and port costs, c^u_i, as well as the variance of the shocks, σ.[36] We obtain the parameters of interest, $\theta = \{c^s_{ij}, c^u_j, \sigma\}$, from the ships' optimal ballast choice probabilities given by (8) and (9). As shown in these two conditions, ships' ballast choice probabilities depend on the ships' value functions (U_i,W_{ij}), which in turn depend on the parameters of interest, θ. We estimate θ via Maximum Likelihood. We use a nested fixed point algorithm to solve for the ship value functions at every guess of the parameter values (Rust (1987)), compute the predicted choice probabilities and then calculate the likelihood.

Since our model features a number of inter-related value functions (W, U, V), it does not fall strictly into the standard Bellman formulation. Hence, we provide Lemma 2, which proves that our problem is characterized by a contraction map and thus the value functions are well defined.

Lemma 2. *For each value of the parameter vector θ, the map $T_\theta : \mathbb{R}^n \to \mathbb{R}^n$, $U \to T_\theta(U)$ with,*

$$T_\theta(U) = -c^u_i + \lambda_i \pi_i + \lambda_i \sum_{j \neq i} G_{ij} \left[-\frac{c^s_{ij}}{1 - \beta(1 - \xi_{ij})} + \beta \xi_{ij} \frac{U_j}{1 - \beta(1 - \xi_{ij})} \right] + (1 - \lambda_i) J_i(\theta, U)$$

is a contraction and $U(\theta)$ is the unique fixed point.

Proof. See the Appendix. $\square$

In brief, our estimation algorithm proceeds in the following steps:

1. Guess an initial set of parameters $\{c^s_{ij}, c^u_i, \sigma\}$.

2. Solve for the ship value functions via a fixed point:

   (a) Set an initial value U^0. Then at each iteration m:

   (b) Solve for W^m from:
   $$W^m_{ij} = \frac{-c^s_{ij} + \xi_{ij} \beta U^m_i}{1 - \beta(1 - \xi_{ij})}$$

   (c) Update J^m from:
   $$J^m_i = \sigma \log \left(\exp \frac{\beta U^m_i}{\sigma} + \sum_{j \neq i} \exp \frac{W^m_{ij}}{\sigma} \right) + \sigma \gamma^{euler}$$

freight make their decisions based on the long-run, steady state distributions (s^*, f^*, s^{w*}).

[36] Note that here σ is identified because the observed prices pin down the scale of utility (in dollars).

(d) Update U^{m+1} from:

$$U_i^{m+1} = -c_{ui} + \lambda_i E_{v,j}\pi_{ijv} + \lambda_i \sum_{j\neq i} G_{ij} W_{ij}^m + (1-\lambda_i)\, J_j^m$$

where we use the actual average prices from i to j, i.e., $E_{v,j}\pi_{ijv} = \sum_{j\neq i} G_{ij}\pi_{ij}$, where π_{ij} is the average observed price from i to j. Note that λ_i is known (it is simply the average ratio m_i/s_i). Similarly, the matrix G_{ij}, whose element (i,j) is the probability that an exporter ships from i to j (conditional on exporting), is obtained directly from the data, simply by computing frequencies of trade flows (loaded trips); we further discuss this in Section 5.3.

3. Form the likelihood using the choice probabilities:

$$\mathcal{L} = \sum_i \sum_j \sum_m y_{ijm} \log P_{ij}(\theta) \tag{24}$$

where y_{ijm} is 1 if ship m chose to go from i to j, and $P_{ij}(\theta)$ are given by (8) and (9).[37]

Identification As is always the case in dynamic discrete choice models, not all parameters of interest are identified and some restriction needs to be imposed. Here, we have I^2+1 parameters and I^2-I choice probabilities, so we require $I+1$ restrictions; we show this formally, borrowing from the analysis of Kalouptsidi, Souza-Rodrigues and Scott (2016) in the Appendix. The additional restrictions amount to using the observed fuel price to determine c_{ij}^s; details in Section 5.2. We estimate the waiting costs $c_1^u, ..., c_I^u$, which may be capturing heterogeneous costs that are difficult to measure (actual port wait costs, ability to wait outside of port, etc.), as well as σ.

We present our results in Section 6.

5.3 Exporter Valuations and Exporting Costs

We begin with exporter valuations v, and then turn to their costs κ. Due to the data on shipping prices, we are able to back out exporter valuations flexibly and in a straightforward manner. Indeed, consider

[37] We assume that our data comes from one steady state. During our sample period of 2012-2016 the industry did not experience any major shocks. Moreover, we estimate the model separately by season to allow for seasonal time variation and find our results to be robust (see Section 6).

the equilibrium price (17) solved with respect to the exporter's valuation:

$$v = \frac{1-\beta\delta\left(1-\gamma\lambda_i^f\right)}{(1-\gamma)(1-\beta\delta)}\pi_{ijv} - \frac{\gamma\left(1-\beta\delta\left(1-\lambda_i^f\right)\right)}{(1-\gamma)(1-\beta\delta)}(J_i - W_{ij}) \tag{25}$$

Notice that once $\theta = \{c_{ij}^s, c_i^u, \sigma\}$ is known, so is J_i and W_{ij}. Moreover, π_{ijv} is observed, while λ_i^f is known from the matching function estimation (λ_i^f is the average ratio $\frac{1}{T}\sum m_{it}/f_{it}$ where f_{it} is estimated). We calibrate the freight survival probability to $\delta = 0.99$ so that freights survive with high probability; and we calibrate the discount factor to $\beta = 0.995$. Now, equation (25) has two unknowns: the value v and the bargaining coefficient γ. Our approach amounts to obtaining γ using some prior knowledge on the average value of commodity trading and then using this parameter to recover v from (25) pointwise. In particular, we collect information on the average price of the five most common commodities and multiply it with the average tonnage carried by a bulk carrier; this resulting average value is 7 million dollars. We take the average of (25) over i, j and solve for γ, finding that $\gamma = 0.3$.[38] Now, given this estimate, we recover exporter valuations point-wise from (25) and obtain their distribution, F_{ij}^v, nonparametrically.[39] Note that valuations are drawn from an origin-destination specific distribution, which allows for arbitrary correlation between a cargo's valuation and destination.

The exporter costs κ_{ij} capture both the cost of production as well as any export costs beyond shipping prices and are estimated from exporters' chosen destination given by the choice probabilities $\tilde{G}_{ij}$ defined in (12). Indeed, given $\tilde{G}_{ij}$ we can recover κ_{ij} as follows (see Berry (1994)):

$$\kappa_{ij} = J_{ij}^f - \left(\ln \tilde{G}_{ij} - \ln \tilde{G}_{i0}\right) \tag{26}$$

where J_{ij}^f is now known: use (10) and the recovered distribution F_{ij}^v

$$J_{ij}^f \equiv E_v J_{ijv}^f = E_v \frac{\lambda_i^f (v - \pi_{ijv})}{1-\beta\delta\left(1-\lambda_i^f\right)} = \frac{\lambda_i^f (\mu_{ij} - \pi_{ij})}{1-\beta\delta\left(1-\lambda_i^f\right)}$$

[38] We collect the average price during our sample period for iron ore, coal, grain, steel and urea from Index Mundi, and obtain $\bar{\mu}$ as their weighted average based on each commodity's frequency in shipping contracts. The average cargo load is equal to the average bulker size times its utilization rate (see Footnote 53). Then, solving (25) with respect to γ and averaging over i, j, v yields:

$$\gamma = \frac{(1-\beta\delta)(\overline{\mu} - \overline{\pi})}{\beta\delta E_{ijv}\lambda_i^f \pi_{ijv} + (1-\beta\delta)\overline{\mu} - E_{ij}\left(1-\beta\delta\left(1-\lambda_i^f\right)\right)(J_i - W_{ij})}$$

where $\overline{\mu}$ is calculated from above to equal 7 million US dollars and $\overline{\pi}$ is the average observed price.

[39] We have implicitly assumed that the exporter obtains zero payoff when he does not find a match. It is not possible to separately identify valuations from such inventory costs or scrap values; the assumption that the exporter obtains zero payoff when he does not find a match means that valuations are interpreted with reference to inventory costs or scrap values.

where $\pi_{ij} = E_v \pi_{ijv}$ and μ_{ij} is the average valuation from i to j. $\tilde{G}_{ij}$ corresponds to the observed frequency of loaded trips going from i to j. We do however need to determine the probability of the outside option; as described in detail in Section 6, we use external data on the total commodity production by country.

6 Results

In this section we present the results from our empirical analysis. Throughout the estimation, we consider 15 geographical regions.[40] To determine these regions, we employ a clustering technique that minimizes the distance between individual ports.[41] Figure 17 in the Appendix depicts the constructed regions. Pointwise confidence intervals and standard errors are computed via bootstrap samples.

6.1 Matching Function

Search Frictions Test Before presenting the main results, we provide some further reduced-form evidence for search frictions, inspired by the model and the empirical methodology outlined in Section 5.1. Suppose that for some market i, it is known that there are more ships than exporters ($s_{it} > f_{it}$), i.e. $\min(f_{it}, s_{it}) = f_{it}$. If there are no search frictions, so that $m_{it} = \min(f_{it}, s_{it}) = f_{it}$, exogenously changing the number of ships does not affect the number of matches. In contrast, if there are search frictions, any exogenous change in the number of ships changes the number of matches. We can thus test for search frictions by using unpredictable changes in ocean weather conditions to explore whether changing the number of ships in markets with a lot more ships than exporters affects the realized number of matches. To proxy for weather conditions, we employ the unpredictable component of wind at sea in all directions.[42] Table 4 in the Appendix presents the results for the west coast of South America, which reveal that indeed, matches seem to be affected by weather conditions, and that thus search frictions are present in this market. We run similar regressions for all markets with significantly more ships than

[40] The trade-off here is that we need a large number of observations per region, while allowing for sufficient geographical detail. The markets are: west coast of North America, east coast of North America, Central America, west coast of South America, east coast of South America, West Africa, Mediterranean, North Europe, South Africa, Middle East, India, Southeast Asia, China, Australia, Japan-Korea. We ignore inter-regional trips.

[41] We have 1492 ports. Let D denote a matrix of distances, so that D_{ij} represents the sea distance among each region pair i, j. We use D as an input to cluster ports through a hierarchical clustering algorithm: in each iteration, the algorithm produces clusters that minimize within-region dispersion.

[42] In particular, we divide the sea surrounding each market into at most 4 different zones. For each zone we use information on the wind speed at different distances from the coast and in different directions. To obtain the unpredictable component of weather we run a VAR regression of these weather indicators on their lag component and season fixed effects. We experiment with the lag structure and the results are overall robust. Finally, the results are robust to running the VAR jointly for neighboring zones.

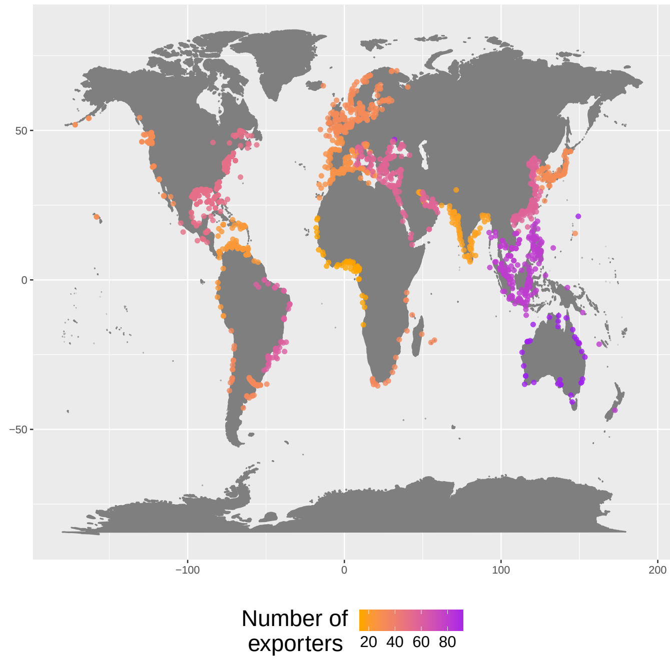


Figure 6: Average weekly number of exporters.

freights, and find similar results.[43]

Matching Function Estimates We now turn to the results from using our main methodology of Section 5.1.[44] Figure 6 presents the weekly average number of exporters in the world. Exporters are concentrated in Australia, the east coast of South America and Southeast Asia; while India, Africa and central America have the lowest freights available. This is consistent with the raw data patterns of Section 3 which documents the global trade imbalance. To evaluate our results, we collect external country-level data on total commodity exports and find that they in fact correlate well with our estimated exporters (see Figure 19 in the Appendix).

To visualize the matching function, Figure 7 plots matching rates for ships and exporters, for the west coast of South America as one example. The top panel plots the matching rate for exporters, λ^f, as a function of the number of exporters searching and for different levels of ships. Note that, as expected, λ^f declines as the market gets crowded with exporters looking for a ship. Similarly, the bottom panel plots the probability that a ship finds a match, λ, as a function of the number of ships and for different levels

[43] It is not straightforward to interpret the signs of the coefficients as each region may have ports facing different directions and so wind from a certain direction will affect them differently.

[44] Following Proposition 2, we need a known point (m^*, s^*, f^*), such that $m(\alpha s^*, \alpha f^*) = \alpha m^*$. We choose $1 = m(s^*, 1)$, so that one exporter is always matched when there are s^* ships. We set s^* such that in all markets $m_i \leq f_i$ and we iterate over s^*. Note that this approach delivers a conservative bound on search frictions, since in principle we could allow for higher levels of exporters.

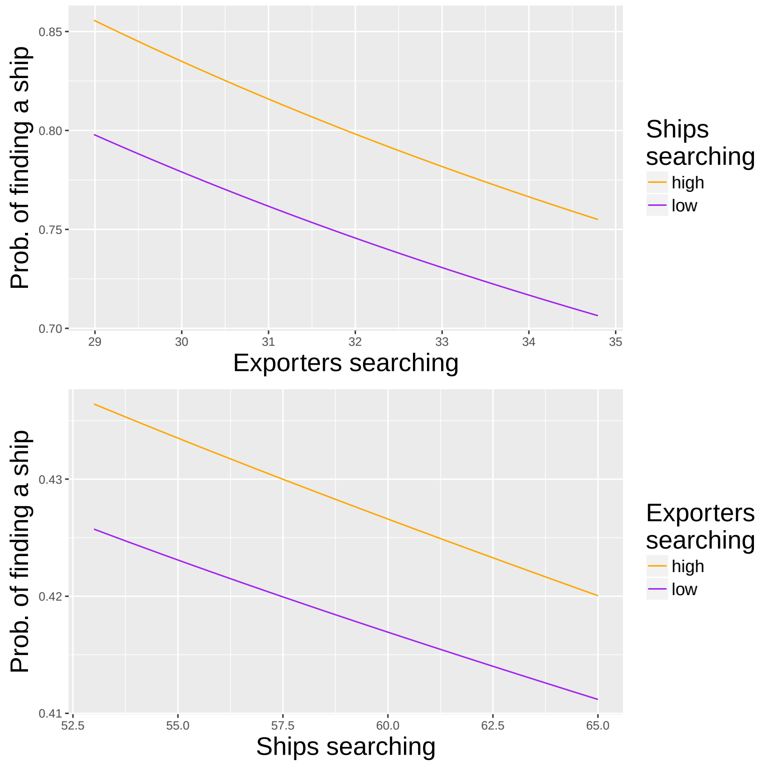


Figure 7: Estimated match finding probabilities for ships and exporters, λ and λ^f in the west coast of South America. "High" corresponds to 40^{th} percentile, while "low" corresponds to 60^{th} percentile.

of exporters. Again, this probability declines in the number of searching ships. It is also worth noting, that exporters have substantially higher chances of finding a match than ships, consistent with our sample period of high ship supply and low demand, as well as our conservative scale restriction on the exporters (see footnote 44). This is true for all markets.

To measure the extent of search frictions in different markets, we compute the average percentage of weekly "unrealized" matches; i.e. $(\min\{s_i, f_i\} - m_i) / \min\{s_i, f_i\}$. The results are plotted in Figure 8 and reveal that search frictions are heterogeneous over space and may be sizable, with up to 20% of potential matches "unrealized" weekly in regions like west Africa and parts of South America and Europe. On average, 17.2% of potential matches are "unrealized".[45,46]

We use the ratio of unrealized matches to evaluate our results. First, we find that the ratio of "unrealized" matches correlates well with the ratio of incoming and outgoing ships, which, as discussed in

[45] Results are overall robust if instead of imposing that the matching function is homogeneous of degree one, we fix the distribution of f; see Footnote 34. In our case, a $[0, 1]$ uniform distribution for freights does not sound plausible since we need to also satisfy $m_{it} \leq f_{it}$ for all i, t. Therefore, we instead experimented with a more flexible distribution (normal, log-normal) and calibrated its parameters so that this inequality is always just satisfied- this again yields the most conservatively estimated level of search frictions.

[46] It is worth noting that this does not imply that in the absence of search frictions we would have 17.2% more matches; this is simply a measure of the severity of search frictions in different markets. In Section 7 we address this question by shutting down search frictions and examining the corresponding change in trade.

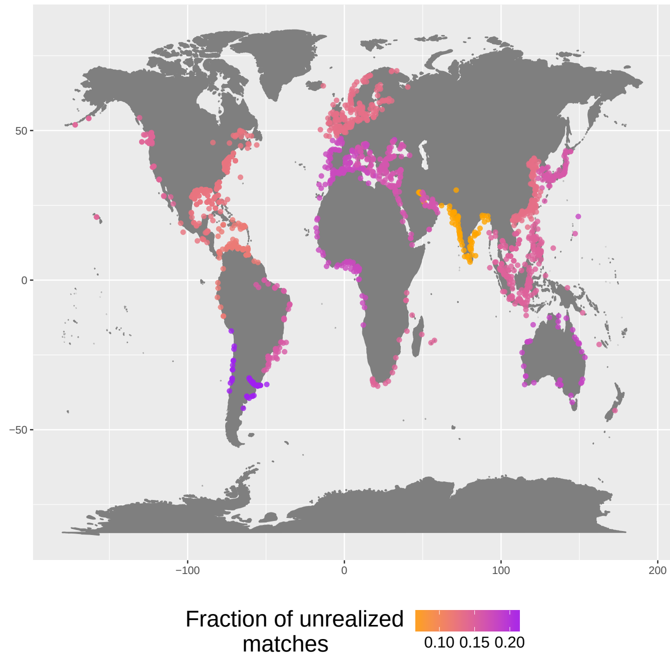


Figure 8: Average weekly share of unrealized matches due to search frictions.

Section 3.3, is also consistent with the presence of search frictions. As an example, comparing Figure 3 to Figure 8 we see that a substantial share of potential matches are unrealized in Chile, while fewer are unrealized in Norway. Second, the ratio of unrealized matches, is also positively correlated with the observed within-market price dispersion, another indicator of search frictions. These findings suggest that our estimates for the matching function and the searching exporters are reasonable. Finally, we estimate the matching function separately for Capesize (biggest size) and Handysize (smallest size) vessels and find that Capesize have somewhat lower ratios of unrealized matches; this is makes sense as the market for Capesize is thinner (results are overall robust across ship types).

6.2 Ship Costs

In our baseline specification, we construct seven groups for the sailing cost c_{ij}^s roughly based on the continent and coast of the origin, and we estimate all port wait costs c_i^u, for all i.[47] Note that c_{ij}^s is the *per week* sailing cost from i to j and its major component is the cost of fuel.[48] We set this cost for one of the groups (for trips originating from the east coast of North and South America) equal to the average

[47] The seven groups are: (i) Central America, west coast Americas; (ii) east coast Americas; (iii) west and south Africa; (iv) Mediterranean, Middle East and Baltic; (v) India; (vi) Australia and southeast Asia; (vii) China, Japan and Korea.

[48] The ship also incurs operating costs (crew, maintenance, etc.). However, these are fixed costs of operation; as such they do not affect the ships' decisions and can be ignored.

fuel price. [49] Note also that since the fuel cost is paid by the exporter when the ship is loaded, we add it to the observed prices.[50]

The first two columns of Table 3 report the results. Not surprisingly perhaps, sailing costs are fairly homogeneous. Port wait costs are more heterogeneous and large, ranging between 260,000 and 130,000 US dollars per week. Consistent with industry narratives, it is costly to let a ship waiting at port. Ports in the Americas are the most expensive, while ports in China, India, Southeast Asia, the Middle East are the cheapest. The estimate of the standard deviation of the preference shocks, σ, is about 11,000 US dollars, roughly 10% of price, which implies that the preference shocks do not account for a disproportionately large part of utility. As shown in Figure 18 of the Appendix, the fit is very good, as our predicted choice probabilities are very close to the observed ones. The results are robust when the estimation is performed separately for each ship type. Results are also robust when we estimate the costs separately by season.[51]

	Port Costs (c_u)	Cost of Travelling (c_s)	Exporters Valuations (μ_v)	σ
North America West Coast	2.604	0.692	15.962	
North America East Coast	2.269	0.691	17.196	
Central America	1.846	0.692	12.711	
South America West Coast	2.004	0.692	11.525	
South America East Coast	2.562	0.691	24.689	
West Africa	1.435	0.641	12.041	
Mediterranean	1.642	0.568	11.174	
North Europe	1.4	0.568	7.9	
South Africa	2.514	0.641	17.855	
Middle East	1.278	0.568	7.108	
India	1.482	0.624	15.109	
South East Asia	1.67	0.56	10.331	
China	1.445	0.558	11.729	
Australia	2.646	0.56	13.849	
Japan-Korea	1.554	0.558	10.626	
				0.1173

Note: all the estimates are in 100,000 USD

Table 3: Ship costs and exporter valuation estimates.

[49] We have experimented heavily with different types of identification restrictions and the results are robust. In particular, we have considered (i) one $c_{sij} = c_s$; (ii) coarser and finer groups; (iii) c_{sij} clustered by the distance between i and j and clustered by the weather between i and j (both capture nonlinear effects of distance on the sailing cost).

[50] The average weekly price of fuel is 69,100$.

[51] As our data comes from a period of historically low shipping prices and our estimated value functions are negative. This is partly due to the fact that we are not modeling ships' expectations so shipowners do not realize that under a mean-reverting demand for seaborne trade prices will go up eventually (see Kalouptsidi (2014)). If we compute the equilibrium under higher exporter valuations that lead to prices closer to the ones observed before 2010, the ship value function indeed becomes positive.

6.3 Exporter Valuations and Costs

In Figure 9 we plot the average exporter valuations across origins, while the third column of Table 3 reports the estimates. There is substantial heterogeneity in valuations across space. South and North America have the highest valuations, while India and southeast Asia have the lowest. This ranking is reasonable, as for instance, Brazil exports grain which is expensive, while Southeast Asia exports mostly coal, which is one of the cheapest commodities. We generalize this example by focusing on grain, the most expensive frequently shipped commodity. In particular, we use the small subsample of contracts that report the product shipped and explore whether countries that have a high share of grain exports tend to have higher estimated valuations. The results, shown in Figure 10 reveal that indeed there is a positive correlation between the two, suggesting that exporters with higher valuations may be producers of more expensive products.[52] Of course, there may be other factors determining the valuation of an exporter such as inventory control, just in time production, etc. On average, the average price π_{ij} is equal to 5% of the average valuation μ_{ij}, consistent with other estimates in the literature (e.g. Hummels, Lugovskyy and Skiba (2008)).

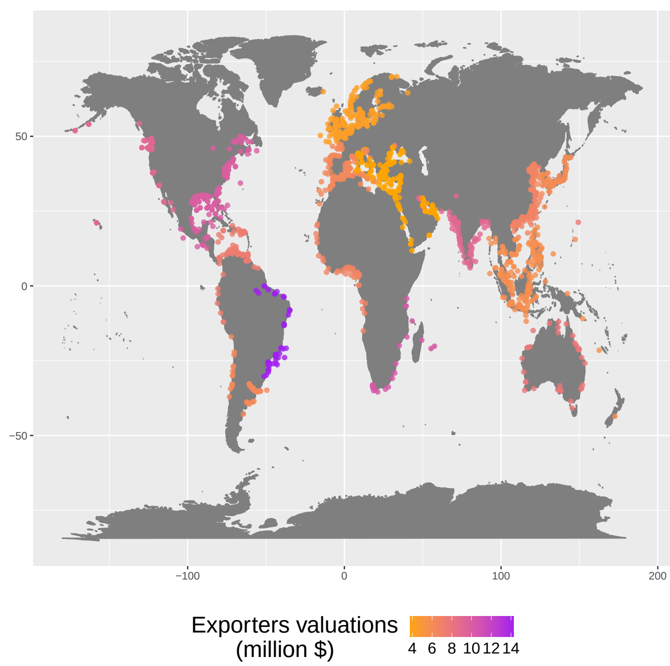


Figure 9: Average exporter valuations.

[52]As discussed in Section 3.3, the dependence of prices on the type of good is suggestive of search frictions in the market. In our estimation we are able to back out the valuations given our estimates for search frictions. The external validation discussed in this paragraph also supports our conclusions that search frictions are present in these markets.

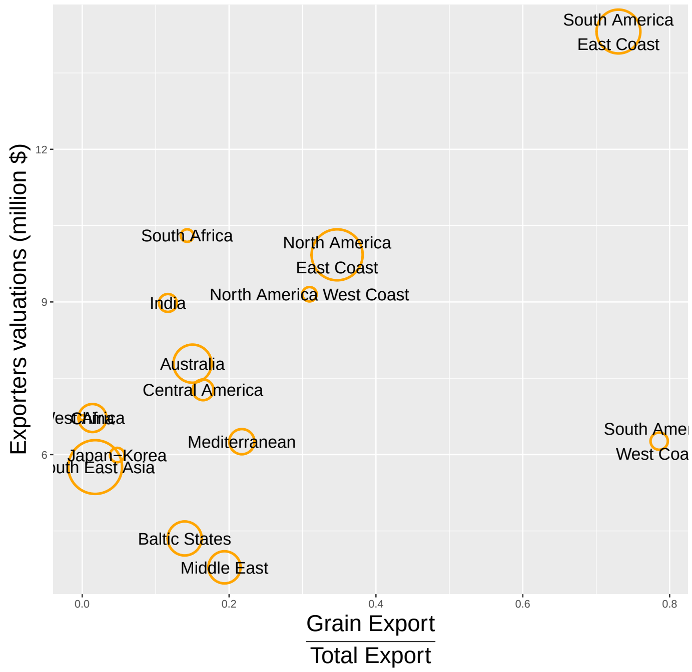


Figure 10: Average exporter valuations and share of exports in grain. The size of the circle proxies the number of observations.

Finally, we turn to exporter costs. We back out the cost of entering into exporting from origin i to destination j, κ_{ij}, from (26). As discussed above, the satellite data provides direct information on the frequencies G_{ij}; indeed, G_{ij} is equal to the proportion of loaded trips going from i to j (see Figure 15 of the Appendix). To obtain κ_{ij} though we also need the share of the outside option, or equivalently the number entrants d_i and potential entrants $\mathcal{E}_i$. We obtain the number of entrants by solving for d_i from the freight transition (1) and taking the average. The number of potential entrants $\mathcal{E}_i$ is set equal to the total production of the relevant commodities for each region i.[53] The share of the outside option is on average 62%. China and India feature the highest outside share, consistent with their large imports, while South America, Southeast Asia and Australia have the lowest outside share.

The estimated exporter costs exhibit substantial heterogeneity across destinations from a given origin, as well as across origins. On average κ_{ij} is the same order of magnitude as the average valuation μ_{ij}. Moreover, we find that exporter costs are lower between an origin i and a destination j if the same language is spoken at i and j, which is reasonable since κ_{ij} includes both production costs, as well as other exporting costs.

[53] We collect annual country-level production quantity data for grain (FAO), coal (EIA), iron ore (US Geological Survey), fertilizer (FAO), steel (World Steel Association). To transform the production tons into shipments, we divide by the total fleet times the average "active" ship size; a ship operates on average 340 days per year (due to maintenance, repair, etc.) and has a deadweight utilization of about 65%. A region's total production serves as an upper bound to the region's exports.

7 Counterfactuals (prelim)

In this section, we use our estimated model to explore a number of questions of interest through the lens of endogenous trade costs. We calculate the trade elasticity with respect to transportation costs. We shed light on how shocks propagate through the entire network of countries; in particular, we ask how a Chinese slow-down affects the world. We compute the impact of the opening of the Northwest Passage on trade costs and flows. Finally, we quantify the social loss and the trade reduction due to search frictions.

To perform the counterfactuals, we compute the steady state spatial equilibrium distribution of ships and exporters defined in Proposition 1. In the Appendix, we provide the computational algorithm employed.[54]

7.1 Trade Elasticity with respect to Transport Cost

We first consider how a decrease in the cost of shipping, c^s, affects shipping prices, π and trade flows. In particular, we reduce the sailing cost, c_s, by 10%. This change in transportation costs has two effects:

First, there is a direct increase in the surplus of all matches, since now a match between a ship and a freight is more valuable.[55] All else equal, this reduces export prices, π, which in turn increases the value of an unmatched exporter, J^f, and thus induces more potential entrants to enter the export market.

That's not all, however. Reducing c^s, implies that ballasting is now cheaper and ships can reallocate across space more freely. Therefore, their "outside option", J is now higher. Since Nash bargaining requires that both parties receive their outside option plus a share of the surplus, an increase in the ship's outside option leads, all else equal, to an increase in prices, π (see the price equation (17)). Put differently, reduced transportation costs implies that ships are less "tied" to the particular market and freights' monopsony power is reduced. This effect tends to mitigate the increase in freight entry driven by the direct effect on the surplus.

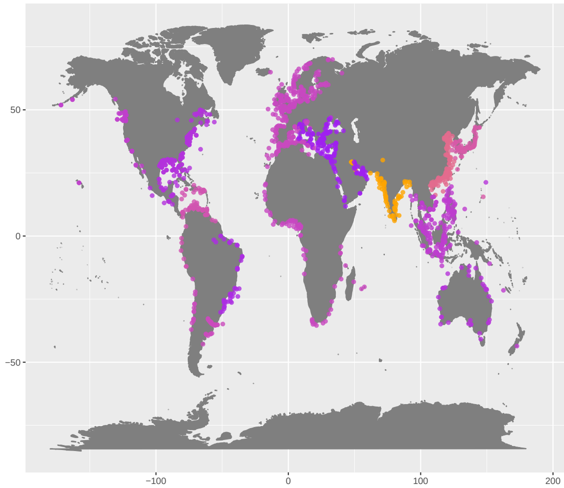
Figure 11 presents the results and showcases that there is substantial heterogeneity in different regions' reaction to a change in transportation costs. The Americas (especially the east coast), Australia and Europe see a 5-15% increase in exporting and a 2-12% decrease in export prices, consistent with the first effect. In contrast, China, India and Japan/S.Korea experience a 5-20% decline in exporting and a 7%

[54] Our algorithm always converged to the same solution, even from quite different starting values.

[55] Formally, using the ship and freight value functions, the match surplus is given by

$$S_{ijv} = v - \frac{c^s_{ij}}{1-(1-\xi_{i,j})\beta} + \frac{\xi_{ij}\beta}{1-(1-\xi_{ij})\beta}U_j - J_i - J^f_{ijv}.$$

A decline in c^s_{ij} directly increases $S_{ll'v}$.



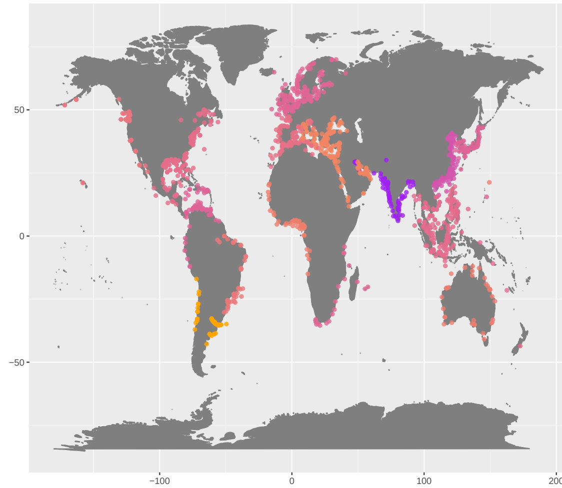


Figure 11: Change in exporting and trade costs under a 10% decline in transportation costs

increase in export prices, suggesting that here the second effect dominates. It is no accident that net exporters see their exporting rise, while net importers see their exporting fall. Ships ending a trip at net importing countries now face lower ballast costs and will thus tend to wait less to get a cargo; their outside option is higher and they can command higher prices. In contrast, net exporters benefit from the increased willingness of ships to ballast, as they see an increase in the number of ships ballasting there. In addition, we see that more distant exporters (South America) benefit more than non-remote exporters (Australia). Similarly, distant importers (India) are hurt more. Overall, consistent with the second effect, we observe that ships are more willing to ballast towards higher value destinations (e.g. east cost of North America, as well as Brazil): as transport costs, and thus distance, now matter less, freight valuations, v, become a relatively more important determinant of ships' decisions.

7.2 Chinese Slow-down

We next explore how shocks propagate in a world where trade costs are endogenous by considering a Chinese slow-down. In particular, we reduce $\mu_{i,china}$ by 10%. Figure 12 presents the resulting trade costs and trade flows across the world.

We begin discussing the results by looking at China itself: shipping prices from China increase by 2%, while exporting declines by 7%. This nicely illustrates the complementarity between imports and exports: the large Chinese imports, led to a large number of ships ending their trip in China and looking for freight there, which reduced the exporting costs for Chinese exporters. Therefore, when imports decline, fewer ships end up in China and Chinese exporters are hurt.

Second, as Figure 12 indicates, China's neighbors including Australia, South-East Asia and India also see their exports fall substantially (between 20 and 39%). On one hand, these countries export to China; therefore, they now lose an important trading partner in both volume and value. This will tend to reduce exporting. It also pushes down prices, both because China is a relatively expensive destination (high $\mu_{i,China}$), but also because ships' outside option is now lower. Again, however, that is not all. This region of the world consisted of some close-by large exporters (Australia, South-East Asia) and importers (China, India) and thus benefited from a "cheap" supply of ships that are stuck in the area ballasting and trading between these countries. China's slow-down tends to reduce this ship glut; this second effect may dampen or amplify the increase (decrease) in trade costs (trade flows). In particular, Australia and South-East Asia, as big net exporters can still attract these ships to export to other destinations (mainly they substitute to India). In contrast, India is a net importer who exported mostly to China; therefore, ships are even less willing now to go to India, which substitutes its exporting activities to the Middle East, an unattractive destination for ships; India thus sees its exporting prices increase. Indeed, if ships were not able to reallocate, India's decline in exporting would have been 19% rather than 39%, while Australia's would have been 41% rather than 35%.

Finally, distant countries are also affected by China's slow-down. The Americas, Europe and Africa experience a 2-10% decline in exporting. Similarly to above, the Americas lose a large and valuable trading partner and this non-surprisingly pushes down their exporting. At the same time, however, they are able to attract the ships that left China and its neighbors. Therefore, although overall their exporting declines, if shipping prices were held fixed, this decline would have been much larger (21% rather than 10% for Brazil, 14% rather than 6% for North America).

7.3 Opening the Northwest Passage

The Northwest Passage is a sea route connecting the northern Atlantic and Pacific Oceans through the Arctic Ocean, along the northern coast of North America. This route is not easily navigable due to Arctic sea ice; with global warming and ice thinning, there is public discussion about opening the passage to be exploited for shipping. The Northwest Passage would reduce the distance between America and the Far East.

To simulate the opening of the Northwest Passage, we reduce the distance between the East Coast of North America and China/Japan/S.Korea, as well as the distance between Europe and China/Japan/S.Korea

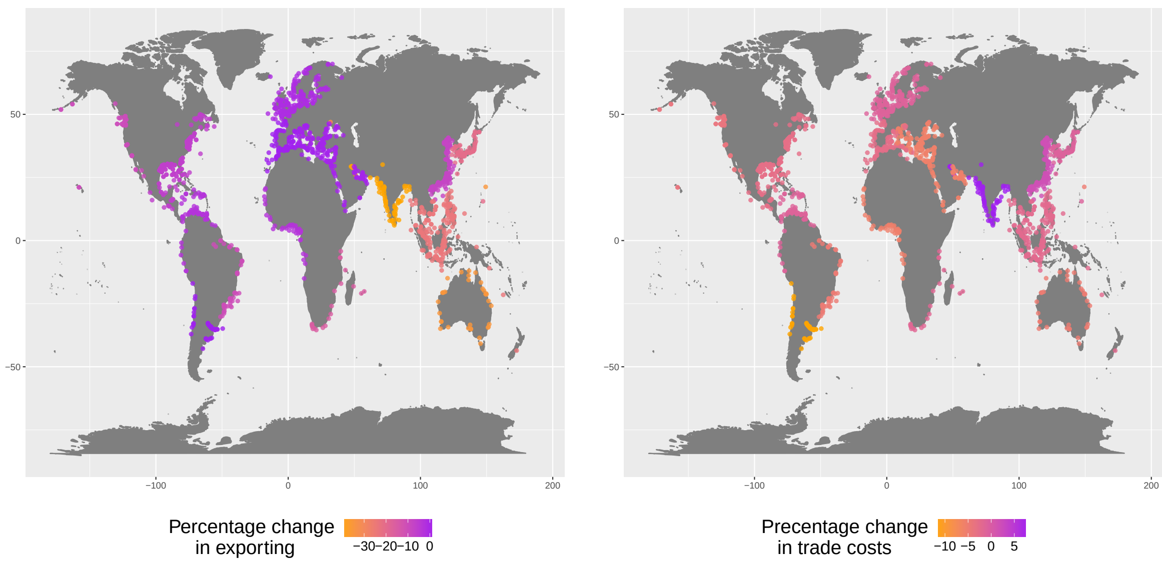


Figure 12: Change in exporting and trade costs under a Chinese slow-down

by 20%.[56] Figure 13 presents the resulting change in trade costs and exporting by region. Naturally, the East Coast of North America sees its exporting rise. Interestingly, China and Japan/S.Korea see their export prices rise and their exports fall. Why? On one hand, the import and export complementarity suggests that exports should increase when imports increase. On the other hand, ballasting is now less costly for ships and when in China ships can now ballast to America more cheaply; this increases shipping prices and decreases exporting. In this case, the second effect dominates.

Finally, Figure 13 reveals that other countries, not directly affected by the opening of the Northwest Passage experience changes in their trade. Strikingly, shipping prices in India increase and exports fall. Ships arriving loaded in India are now less likely to remain in port waiting for a cargo, as they can ballast to nearby exporters (Southeast Asia, Australia) and from there load to go to China, which is now a more attractive destination (U_{china} is higher since ships can then more easily ballast to North America). In contrast, other countries benefit from the Northwest passage; for instance, in Brazil and Australia shipping prices fall and exporting rises: as these countries export more to China and China is a more attractive destination now, ships require lower payment to go there. These effects are present because of the endogenous trade cost and demonstrate the spillover of such policy changes through the network of countries.

[56]Sailing distances between New York and Shanghai via the Northwest Passage are obtained from Østreng et al 2013, while to approximate the best alternative route we use an online voyage calculator (http://sea-distances.com/).

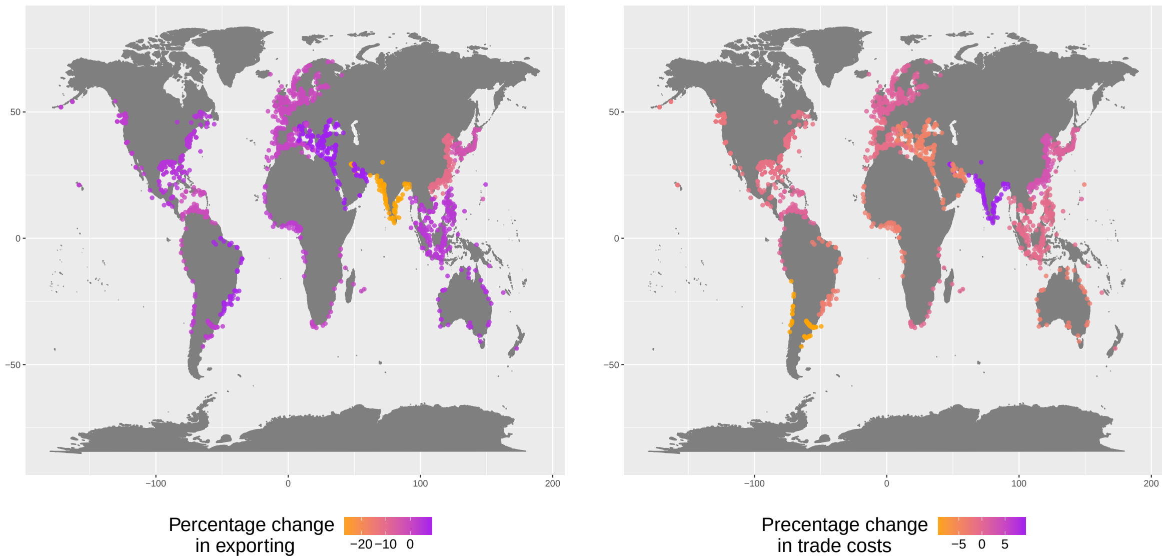


Figure 13: Change in exporting and trade costs if the Northwest Passage was a viable commercial route.

7.4 The Cost of Search Frictions

Finally, we quantify the trade lost because of search frictions. To do so, we shut down search frictions by setting $m(s, f) = \min\{s, f\}$. We find that exporting would be 10-60% higher across different regions in the world. While countries that experience more severe frictions, as captured in Figure 14, see large increase in exports, this is not always the case. In particular, one has to take into account the mechanisms illustrated in the previous counterfactuals and how eliminating search frictions affects the entire network of countries. For instance, high value exporters like Brazil and the east coast of North America experience disproportionally large increases in exports, as differences in frictions across markets are no longer relevant and value, μ, becomes a much more important determinant of trade. Indeed the correlation between the change in exports and μ_{ij} is 0.71.[57]

[57] The p-value for the correlation coefficient is 0.

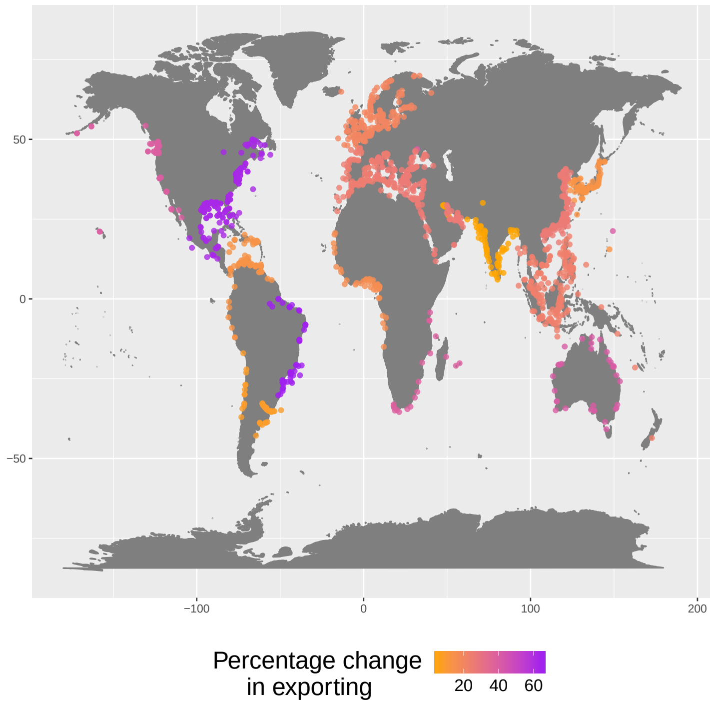


Figure 14: Change in exporting in a world without search frictions.

8 Conclusion

In this paper, we build a dynamic spatial search model for world ships and exporters. Using unique data on shipping contracts and ship movements we recover the main primitives of interest: the matching function between ships and freights, the distribution of searching exporters, ship costs, exporter valuations. Our methodology allows us to obtain the matching process flexibly, without relying on assumptions regarding the extent of search frictions or the parametric form of the matching function. We demonstrate that accounting for the endogeneity of trade costs is important in both descriptive analysis (e.g. elasticities, shock propagation), as well as policy analysis (e.g. transportation infrastructure planning). Finally, we find that search frictions substantially reduce world trade.

References

[1] Anderson James and Eric Van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle". The American Economic Review, Vol. 93, No. 1, pp. 170-192

[2] Arkolakis and Allen. 2014. "Trade and the Topography of the Spatial Economy". Quarterly Journal of Economics, 129(3).

[3] Arkolakis and Allen. 2016. “The Welfare Effects of Transportation Infrastructure Improvements”. Mimeo, Yale University.

[4] Aguirregabiria, Victor, and Pedro Mira. 2007. “Sequential Estimation of Dynamic Discrete Games”. Econometrica 75 (1): 1-53.

[5] Bajari, Patrick and Lanier C. Benkard. “Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach”. Journal of Political Economy 113.6 (2005): 1239-1276.

[6] Berry T. Steven. “Estimation Discrete-Choice Models of Product Differentiation”. The RAND Journal of Economics 25(2)(1994): 242-262.

[7] Bertsekas Dimitri P., and John N. Tsitsiklis. “Parallel and Distributed Computation:Numerical Methods”. MIT Libraries. 2003.

[8] Buchholz, Nicholas. “Spatial equilibrium, search frictions and efficient regulation in the taxi industry”. Working paper, 2016.

[9] Collard-Wexler, Allan. 2013. “Demand Fluctuations in the Ready-Mix Concrete Industry”. Econometrica 81 (3): 1003–1037.

[10] Donaldson, David. 2012. “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure”. American Economic Review, Forthcoming.

[11] Eaton Jonathan and Samuel Kortum. 2002. “Technology, Geography, and Trade”. Econometrica, Vol. 70, No. 5

[12] Ericson, Richard, and Ariel Pakes. 1995. “Markov-Perfect Industry Dynamics: A Framework for Empirical Work”. The Review of Economic Studies 62(1): 53-82.

[13] Frechette Guillaume R., Alessandro Lizzeri, and Tobias Salz. “Frictions in a Competitive, Regulated Market Evidence from Taxis”. Working Paper, 2016

[14] Hopenhayn, Hugo. 1992. “Entry, Exit and Firm Dynamics in Long Run Equilibrium”. Econometrica 60 (5): 1127-1150.

[15] Hummels, David and Sasha Skiba. “Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture”. Journal of Political Economy 112.6 (2004): 1384-1402.

[16] Hummels, David, Volodymyr Lugovskyy and Sasha Skiba. 2008. "The Trade Reducing Effects of Market Power in International Shipping". Journal of Development Economics 89 (2009) 84-97

[17] Hotz, V. Joseph, and Robert A. Miller. "Conditional Choice Probabilities and the Estimation of Dynamic Models". Review of Economic Studies 60.3 (1993): 497-529.

[18] Kalouptsidi, Myrto. "Time to build and fluctuations in bulk shipping." The American Economic Review 104.2 (2014): 564-608.

[19] Kalouptsidi, Myrto, Paul T. Scott, and Eduardo Souza-Rodrigues. "Identification of Counterfactuals and Payoffs in Dynamic Discrete Choice with an Application to Land Use". No. w21527. National Bureau of Economic Research, 2015.

[20] Lange, Fabian and Theodore Papageorgiou. 2017. "Estimating Matching Functions when Search Intensity is Unobserved"

[21] Lagos, Ricardo. 2000. "An Alternative Approach to Search Frictions". Journal of Political Economy. 108(5)

[22] Lagos, Ricardo. 2003. "An Analysis of the Market for Taxicab Rides in New York City". International Economic Review. 44(2)

[23] Matzkin, Rosa L. "Nonparametric estimation of nonadditive random functions". Econometrica 71.5 (2003): 1339-1375.

[24] Østreng, Willy, Eger, Karl Magnus, Fløistad, Brit, Jørgensen-Dahl, Arnfinn, Lothe, Lars, Mejlænder-Larsen, Morten and Wergeland, Tor. 2013. "Shipping in Arctic Waters: A Comparison of the Northeast, Northwest and Trans Polar Passages". Springer.

[25] Pakes, Ariel, Ostrovsky Michael and Steven T. Berry. 2007. "Simple Estimators for the Parameters of Discrete Dynamics Games". The RAND Journal of Economics 38 (2): 373-399.

[26] Petrongolo, Barbara, and Christopher A. Pissarides. "Looking into the black box: A survey of the matching function." Journal of Economic Literature 39.2 (2001): 390-431.

[27] Pissarides, C. A. (2000): *Equilibrium Unemployment Theory*, 2nd edition, Cambridge MA: MIT Press.

[28] Ryan, P. Stephen. 2012. "The Costs of Environmental Regulation in a Concentrated Industry". Econometrica 80(3): 1019–1062.

[29] Rust, John. "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." Econometrica: Journal of the Econometric Society (1987): 999-1033.

[30] Stopford, Martin. 1999. “Maritime Economics”. New York: Routledge.

[31] Waugh, Michael. 2010. “International Trade and Income Differences”. The American Economic Review, 100(5): 2093-2124.

9 Appendix

9.1 Construction of Ship Travel Histories

Here, we describe how the travel history of ships is constructed. The first task is to identify stops that ships make, using EE data, which provides the exact location of ships every 6 minutes. A stop is defined as an interval of at least 24 hours during which (i) the average speed of a ship is below 5 mph (the sailing speed is between 15 and 20 mph) and, (ii) the ship is located within 250 miles from the coast. A trip is the travel between two stops.

The second task is to identify whether a trip is loaded or ballast. To do so, we use the ship's draft: high draft indicates that a larger portion of the hull is submerged and therefore the ship is loaded. The distribution of draft for a given vessel is roughly bimodal, since as described in Section 2, a hired ship is usually fully loaded. Therefore, we can assign a “high” and a “low” draft level for each ship and consider a trip loaded if the draft is high (in practice, low draft is 70% of the highest draft). The data on draft is not as complete as that on location, as not all satellite signals contain the draft information. We therefore consider a trip ballast (loaded) if we observe a signal of low (high) draft during the period that the ship is sailing. If we have no draft information during the sailing time we consider the draft at adjacent stops. Finally, we exclude stops longer than 6 weeks, as such stops may be related to maintenance or repairs.

The third and final task is to refine the origin and destination information provided in the Clarksons contracts. Although the majority of Clarksons contracts provide some information on the origin and destination of the trip, this information is often vague (e.g. “Far East”, “Japan-SKorea-Singapore”), especially in terms of destination. We use the EE data to refine the contracted trips' origins and destinations by matching each Clarksons contract to the identified stop in EE, that is closest in time and, when possible, location.[58] First we use the loading date annotated on each contract to find a stop in the ship's movement

[58] Some contracts require a ship to ballast from a “delivery port” to a “loading port” (ballast leg). In this case, we consider as origin the loading port.

history that corresponds to the beginning of the contract. This is done by finding the stop that is closest in time to the loading date. While we have good data about the beginning of the contract, the data on on the date or port in which the cargo is to be discharged is somewhat more noisy. Therefore, we search the ships history for a stop that we can classify as the end of the contract. In particular, we consider all stops within a three months window since the beginning of the contract. Among these stops we eliminate all those that (i) are in the same country in which the ship loaded the cargo and (ii) are in Panama, South Africa, Gibraltar or at the Suez canal and in which the draft of arrival is the same as the draft of departure (to exclude cases in which the ship is waiting to pass through a strait or a canal) [59]. To select the end of the contract among the remaining options we consider the following possibilities:[60]

1. If the contract reports a country of arrival and if there are stops in this country, select the first of these stop as the end of the trip;

2. If the country of arrival is "Japan-SKorea-Singapore", and if there are stops in either Japan, China, Korea, Taiwan or Singapore, we select the first among these as the end of the trip;

3. If the contract doesn't report a country of arrival and there are stops in which the ship arrives full and leaves empty, we select the first of these as the end of the trip;

9.2 Construction of Searching Ships

Here we describe how the vector of searching ships $s_t = [s_{t1}, ..., s_{tI}]$ and the vector of matches $m_t = [m_{t1}, \cdots, m_{tI}]$ are constructed, where s_{it} denotes the number of ships in market i and week t that are available to transport a cargo and m_{it} the realized matches in market i and week t. To construct s_{it} we consider all ships that ended a trip (loaded or ballast) in market i and week $t-1$. We exclude the first week post arrival in the market to account for loading/unloading times. To construct m_{it}, we consider the number of ships that began a loaded trip from market i in week t.

9.3 Additional Figures and Tables

[59] For the contracts that call for a ballast leg from a "delivery port" to a "loading port", we look for a stop following the start of the contract in which the ship arrives empty and leaves full, and eliminate all the preceding stops.

[60] We check the performance of the algorithm by comparing the duration of common trips obtained in this way with distances provided by https://sea-distances.org/, and find that we match trip durations well.

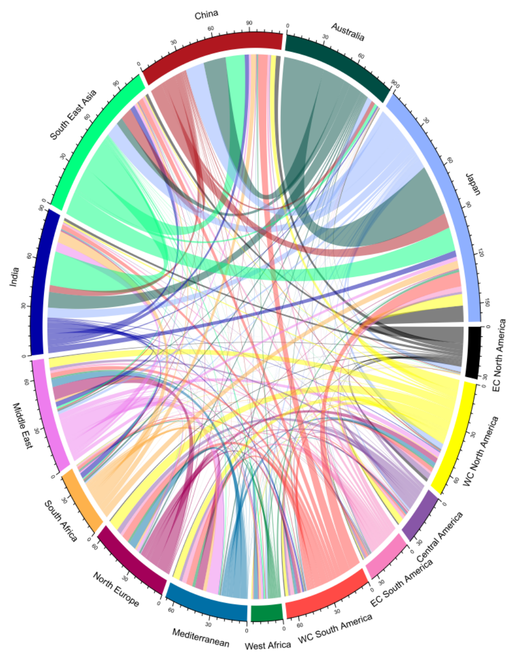


Figure 15: Flows of loaded trips. The colored bar along the perimeter of the circle is proportional to the total number of incoming and outgoing loaded ships for each market. The number of outgoing loaded ships from a market are represented as rays of the same color as the color bar, and are directed toward the market of destination. The width of each outgoing ray is proportional to the number of loaded ships headed from the market of origin to the market of destination.

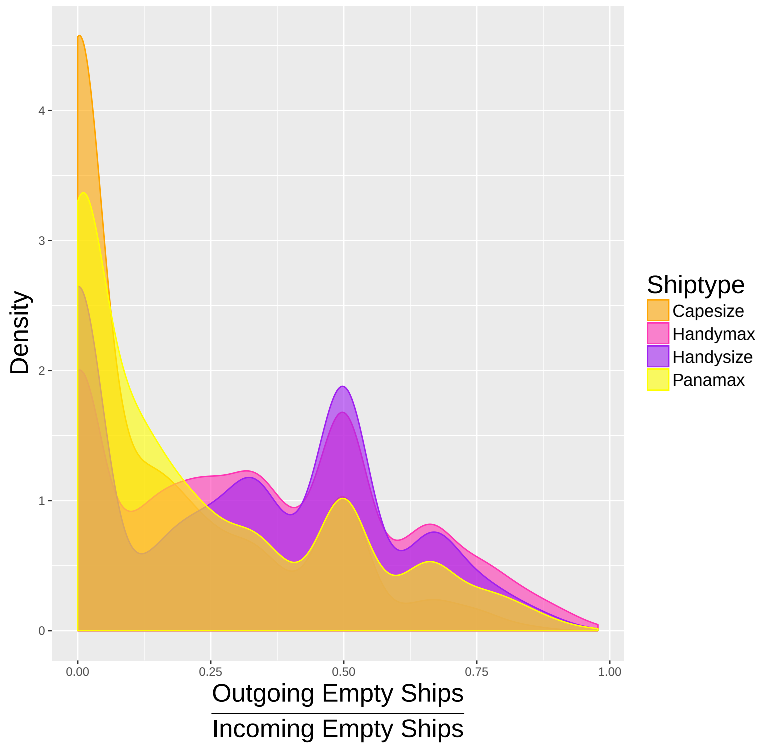


Figure 16: Histogram of the ratio of incoming and outgoing empty ships in net exporting countries, by ship types.

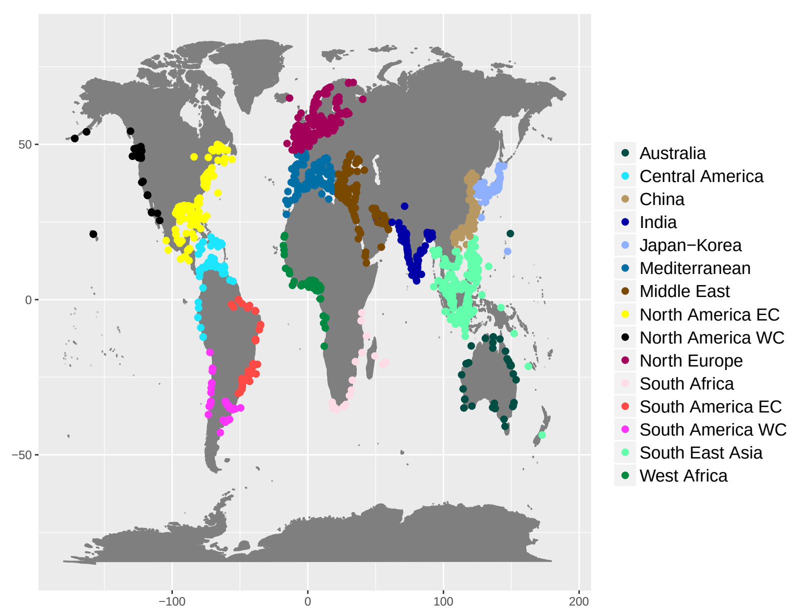


Figure 17: Definition of markets. Each color depicts on of the 15 geographical regions.

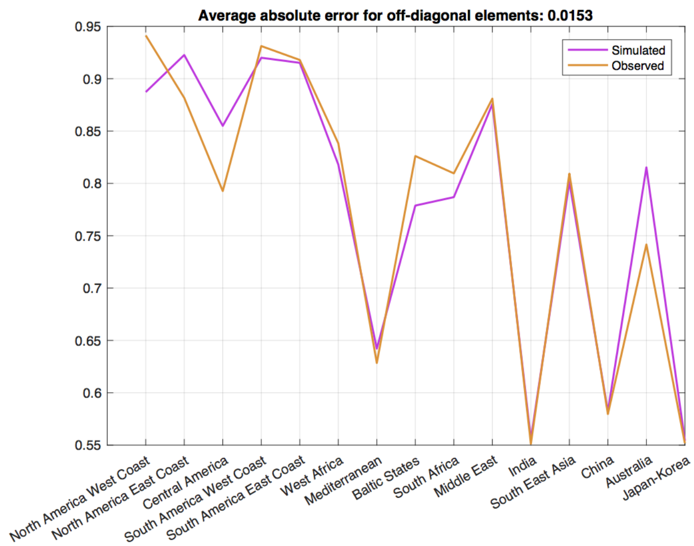


Figure 18: Estimated and simulated probability of waiting in port

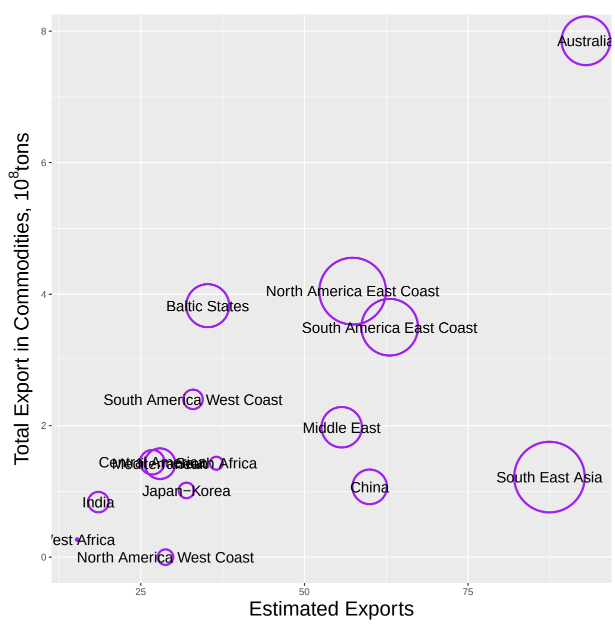


Figure 19: Correlation between the estimated number of exporters and external data on country level commodity exports. We download commodity export data (comtrade and EIA). We then divide the value of exports by the average price for the commodity, to proxy for the total volume (in metric tons). As this measure involves non-seaborne trade as well, it is potentially a good proxy for the number of exporters.

	matches m_{it} South America West Coast	
N/S component in t, from NW, at sea	26.448**	25.135**
	(0.534)	(0.644)
N/S component in t-1, from NE, at coast	-0.512	-1.248**
	(0.366)	(0.465)
N/S component in t-1, from SE, at sea	0.832**	0.601
	(0.403)	(0.49)
E/W component in t, from NW, at coast	0.498**	0.53*
	(0.223)	(0.275)
E/W component in t, from SW, at sea	1.173**	0.828
	(0.533)	(0.659)
E/W component in t-1, from NE, at coast	-0.762*	-0.708
	(0.413)	(0.473)
E/W component in t-1, from SE, at sea	-0.476**	-0.464*
	(0.228)	(0.268)
E/W component in t-1, from SW, at sea	1.258**	1.667**
	(0.52)	(0.679)
Obs	200	145
R^2	0.336	0.413
Joint Significance	0	0
F stat	2.645	2.465

$^{**}p < 0.05, ^{*}p < 0.1$

Table 4: Regression of the number of matches on current and lagged wind speed. Column I: weeks such that ships are at least 1.5 times higher than matches. Column II: weeks such that ships are at least two times higher than matches. As regressors we use the unpredictable component of weather in different zones of the sea sourrounding the west coast of South America. In particular, we divide the sea surrounding the coast into 8 different zones (North East, South East, South West and North West both close to the coast and in open sea), and we use the speed of the horizontal (E/W) and vertical (N/S) component of wind in each of these zones to proxy for the weather. Finally we run a VAR regression of these weather proxies on their lag component and season fixed effects and use the residuals in the above regression.

	ships s_{it}			
	North America East Coast	South America East Coast	South East Asia	China
E/W component in t, from NE, at coast	1.466	0.673	-	-
	(2.083)	(0.71)	-	-
E/W component in t, from NE, at sea	-3.231**	5.539**	-	-
	(1.118)	(1.863)	-	-
E/W component in t, from NW, at coast	-1.091	-	-	-5.555**
	(0.772)	-	-	(1.998)
E/W component in t, from NW, at sea	-	-1.194	-	-12.241*
	-	(0.737)	-	(6.238)
E/W component in t, from SW, at coast	-	1.144	-	-
	-	(1.024)	-	-
E/W component in t-1, from NE, at coast	-2.828	-	-	-0.927
	(1.794)	-	-	(0.677)
E/W component in t-1, from NE, at sea	-1.056	-	-	-
	(0.774)	-	-	-
E/W component in t-1, from NW, at coast	-	-	-	-0.888
	-	-	-	(0.704)
E/W component in t-1, from NW, at sea	-	-1.107	-	-
	-	(1.02)	-	-
E/W component in t-1, from SW, at coast	-	-3.396*	-	-
	-	(1.881)	-	-
E/W component in t-1, from SW, at sea	-	5.708**	-	-
	-	(1.857)	-	-
N/S component in t, from NE, at coast	-	1.628	-	1.065
	-	(1.142)	-	(0.731)
N/S component in t, from NW, at sea	-	-	207.974**	-
	-	-	(2.008)	-
N/S component in t, from SE, at sea	112.49**	157.273**	-	321.98**
	(1.513)	(2.442)	-	(2.321)
N/S component in t-1, from NE, at coast	-	-	-	1.052
	-	-	-	(0.728)
N/S component in t-1, from NW, at coast	-	-	-5.461**	-6.207**
	-	-	(1.6)	(1.997)
N/S component in t-1, from NW, at sea	-	-	-3.186**	-12.027*
	-	-	(1.232)	(6.121)
N/S component in t-1, from SE, at sea	-3.258**	3.295**	-	1.004
	(1.102)	(1.658)	-	(0.864)
N/S component in t-1, from SW, at sea	-	1.611	-	-
	-	(1.676)	-	-
Obs	200	200	200	200
R^2	0.106	0.178	0.081	0.176
Joint Significance	0.003	0	0.001	0
F stat	3.255	3.710	5.768	4.045

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5: First Stage, Matching Function Estimation

9.4 Proof of Proposition 1

We first derive (18) and (19). Suppose $f_{it}, s_{it}, s^w_{ijt}$ approach f_i, s_i, s^w_{ij}, d_i as $t \to \infty$. Then (2) becomes:

$$s_i = \left(s_i - m_i\left(s_i, f_i\right)\right) P_{ii} + \sum_{j \neq i} \xi_{ji} s^w_{ji} \tag{27}$$

while for a ship traveling from j to i, (3) becomes:

$$s^w_{ji} = \left(1 - \xi_{ji}\right) s^w_{ji} + P_{ji} s_j + \left(G_{ji} - P_{ji}\right) m_j\left(s_i, f_i\right) \tag{28}$$

or

$$\xi_{ji} s^w_{ji} = P_{ji} s_j + \left(G_{ji} - P_{ji}\right) m_j = P_{ji}\left(s_j - m_j\right) + G_{ji} m_j$$

where $m_i = m_i\left(s_i, f_i\right)$. Summing this with respect to $j \neq i$ we obtain:

$$\sum_{j \neq i} \xi_{ji} s^w_{ji} = \sum_{j \neq i} P_{ji}\left(s_j - m_j\right) + \sum_{j \neq i} G_{ji} m_j$$

and replacing in (27) we get (18).

Equation (19) is a direct consequence of (1) and (14).

The steady state equations (18) and (19) have a fixed point over a properly defined subset of $\mathbb{R}^{2I}$, by the Leray-Schauder-Tychonoff theorem (Bertsekas and Tsitsiklis (2003)) which states that if X is a non-empty, convex and compact subset of $\mathbb{R}^{2I}$ and $h : X \to X$ is continuous, then h has a fixed point. Indeed, let $h : \mathbb{R}^{2I} \to \mathbb{R}^{2I}$, $h = (h^s, h^f)$ with:

$$h^s_i\left(s, f\right) = \sum_{j=1}^{I} P_{ji}\left(s, f\right)\left(s_j - m_j\left(s_j, f_j\right)\right) + \sum_{j \neq i} G_{ji} m_j\left(s, f\right)$$

$$h^f_i\left(s, f\right) = \delta\left(f_i - m_i\left(s_i, f_i\right)\right) + E_i \sum_{j \neq 0, i} \tilde{G}_{ij}\left(s, f\right)$$

for $i = 1, ..., I$. Let $X = \prod_{i=1}^{I} \left[0, E_i/(1-\delta)\right] \times \Delta s$, where $\Delta s = \left\{ s_i \geq 0 \ : \ \sum_{i=1}^{I} s_i \leq S \right\}$. X is nonempty, convex and compact, while h is continuous on X We assume that the matching function is such that λ, λ^f are zero at the origin and continuous. It remains to show that $F(X) \subseteq X$. Let $(s, f) \in X$. Then,

$f_i \leq E_i/(1-\delta)$ and $\sum_{i=1}^{I} s_i \leq S$. Now,

$$h_i^s(s,f) = \sum_{j=1}^{I} P_{ji}(s,f)\left(s_j - \lambda_j(s_j,f_j)\, s_j\right) + \sum_{j\neq i} G_{ji}\lambda_j(s,f)\, s_j$$

or

Proof.

$$h_i^s(s,f) = \sum_{j=1}^{I} s_j \left[P_{ji}(s,f)\left(1-\lambda_j(s_j,f_j)\right) + G_{ji}\lambda_j(s,f)\right]$$

where for convenience let $G_{ii} = 0$ (we do not allow for inter-market trips). Summing over i gives:

$$\sum_{i=1}^{I} h_i^s(s,f) = \sum_{j=1}^{I} s_j \left[\sum_{i=1}^{I} P_{ji}(s,f)\left(1-\lambda_j(s_j,f_j)\right) + \sum_{i=1}^{I} G_{ji}\lambda_j(s,f)\right]$$

or

$$\sum_{i=1}^{I} h_i^s(s,f) = \sum_{j=1}^{I} s_j \left[1-\lambda_j(s_j,f_j) + \lambda_j(s,f)\right] \leq S$$

Hence $h_i^s(s,f) \in \mathit{\Delta} s$.

Finally, consider h^f; since $m_i \geq 0$, we have

$$h_i^f \leq \delta f_i + E_i \sum_{j\neq 0,i} \tilde{G}_{ij}(s,f) \leq \delta f_i + E_i \leq \delta \frac{E_i}{1-\delta} + E_i = \frac{E_i}{1-\delta}$$

Hence $h_i^f(s,f) \in \left[0, E_i/(1-\delta)\right]$. □

9.5 Proof of Proposition 2

(i) Following Matzkin (2003), two matching functions $m(\cdot)$ and $\tilde{m}(\cdot)$ are observationally equivalent if there a exists a strictly increasing and differentiable function $g(\cdot)$ such that:

$$\tilde{m}(s,f) = m(s,g(f))$$

Let $\lambda > 0$ and fix $\bar{s}, \bar{f}$. Then

$$\tilde{m}(\lambda\bar{s},\lambda\bar{f}) = \lambda\tilde{m}(\bar{s},\bar{f}) = \lambda\bar{m}$$

Furthermore,

$$\tilde{m}(\lambda\bar{s},\lambda\bar{f}) = m(\lambda\bar{s},g(\lambda\bar{f})) = \lambda m(\bar{s},\frac{1}{\lambda}g(\lambda\bar{f}))$$

Therefore,

$$\bar{m} = \tilde{m}(\bar{s}, \bar{f}) = m(\bar{s}, \frac{1}{\lambda} g(\lambda \bar{f}))$$

Invertibility implies that $\bar{f} = \tilde{m}^{-1}(\bar{s}, \bar{m})$ and $\frac{1}{\lambda} g(\lambda \bar{f}) = m^{-1}(\bar{s}, \bar{m})$, or

$$g(\lambda \bar{f}) = \lambda m^{-1}(\bar{s}, \bar{m})$$

Differentiate with respect to λ to obtain $\bar{f} g'(\lambda \bar{f}) = m^{-1}(\bar{s}, \bar{m})$, which for $\lambda = 1$ becomes $g'(\bar{f})\bar{f} = m^{-1}(\bar{s}, \bar{m}) = g(\bar{f})$. Therefore, the Euler condition is satisfied and $g(\cdot)$ is homogeneous of degree 1. Since $g(\cdot)$ is a function of a real variable, the only possibility is $g(f) = cf$ with $c > 0$, a constant. Finally, we use the a priori knowledge of the point $m^* = m(s^*, f^*)$ to establish that $c = 1$. Indeed, by definition, $m(s^*, f^*) = \tilde{m}(s^*, f^*) = m^*$. But also, $m(s^*, cf^*) = \tilde{m}(s^*, f^*)$. Therefore, $cf^* = f^*$ and since $f^* \neq 0$, $c = 1$.

(ii) Conditional on η, s is a function of z which in turn is by assumption independent from f. It follows that s and f are conditionally independent given η. At a point ϕ we have that:

$$F_{f|\eta}(\phi) = \Pr\left(f \leq \phi | \eta\right) = \Pr\left(f \leq \phi | \eta, s\right) = \Pr\left(m(s, f) \leq m(s, \phi) | \eta, s\right) = F_{m|s,\eta}\left(m(s, \phi)\right)$$

Hence,

$$m(s, \phi) = F^{-1}_{m|s,\eta}(F_{f|\eta}(\phi))$$

which we integrate over η to obtain the result. Let $\phi = \frac{\phi}{f^*} f^*$. Then,

$$F_{f|\eta}(\phi) = F_{m|s=\frac{\phi}{f^*}s^*,\eta}\left(m\left(\frac{\phi}{f^*}s^*, \frac{\phi}{f^*}f^*\right)\right) = F_{m|s=\frac{\phi}{f^*}s^*,\eta}\left(\frac{\phi}{f^*}m^*\right)$$

9.6 Proof of Lemma 2

Fix θ. Let $\phi_{ij} = \frac{1}{1-\beta(1-\xi_{ij})}$. The map $T_\theta(U)$ is differentiable with respect to U with Jacobian:

$$\frac{\partial T_\theta(U)}{\partial U} = \beta \left(DG + (I - D)\, P\right) \odot Z \tag{29}$$

where D is a diagonal matrix with λ_i it's i diagonal entry; P is the matrix of choice probabilities, G is the matrix of matched trips, Z is an $L \times L$ matrix whose (i, j) element is $\phi_{ij}\xi_{ij}$ and $\odot$ denotes the pointwise product. Indeed, the (i,j) entry of $\frac{\partial T}{\partial U}$ is

$$\left(\frac{\partial T}{\partial U}\right)_{ij} = 1\{i=j\} - \beta\lambda_i G_{ij}\xi_{ij}\phi_{ij} - (1-\lambda_i)\frac{\partial J_i}{\partial U_j}$$

Now,

$$\frac{\partial J_i}{\partial U_j} = \sigma \frac{1}{\sum_k e^{\frac{W_{ik}}{\sigma}}} e^{\frac{W_{ij}}{\sigma}} \frac{\partial W_{ij}}{\partial U_j} = \beta P_{ij}\xi_{ij}\phi_{ij}$$

and thus

$$\left(\frac{\partial T}{\partial U}\right)_{ij} = 1\{i=j\} - \beta\left(\lambda_i G_{ij} + (1-\lambda_i) P_{ij}\right)\xi_{ij}\phi_{ij}$$

which in matrix form becomes (29) (as a convention set $\xi_{ii} = 1$). Let $H = (DG + (I-D)P) \odot Z$. Take $||H|| = \max_i \sum_j |H_{ij}|$. Note that G, P are stochastic matrices and the diagonal matrix D is positive with entries smaller than 1. Thus $DG + (I-D)P$ is stochastic. It is also true that $0 < \xi_{ij}\phi_{ij} \leq 1$. Thus,

$$\sum_j |H_{ij}| = \sum_j \left(\lambda_i G_{ij} + (1-\lambda_i) P_{ij}\right)\xi_{ij}\phi_{ij} \leq \sum_j \left(\lambda_i G_{ij} + (1-\lambda_i) P_{ij}\right) \leq 1$$

and therefore $||H|| \leq 1$. We deduce that $||\frac{\partial T_\theta(U)}{\partial U}|| \leqq \beta < 1$. The mean value theorem then implies

$$||T_\theta(U) - T_\theta(U')|| \leq \beta ||U - U'||$$

9.7 Identification of Ship Port and Sailing Costs

Proposition 3. *Given the choice probabilities $P_{ij}(\theta)$, the parameters $\{\frac{c^s_{ij}}{\sigma}, \frac{c^u_i}{\sigma}, \frac{1}{\sigma}\}$ satisfy a $(I^2 - I) \times (I^2 + 1)$ linear system of equations of full rank $I^2 - I$. Hence, $(I+1)$ additional restrictions are required for identification.*

Proof. Let $\phi_{ij} = \frac{1}{1-\beta(1-\xi_{ij})}$. The Hotz and Miller (1993) inversion states:

$$\sigma \log \frac{P_{ij}}{P_{ii}} = W_{ij}(\theta) - W_{ii}(\theta)$$

Substituting from (4)-(5) we obtain:

$$\sigma \log \frac{P_{ij}}{P_{ii}} = -\phi_{ij} c^s_{ij} + \beta\xi_{ij}\phi_{ij} U_j(\theta) + c^u_i - \beta U_i(\theta) \tag{30}$$

It also holds that (see Kalouptsidi, Scott and Souza-Rodrigues (2016)):

$$\log P_{ij} = \frac{W_{ij}}{\sigma} - \frac{J_i}{\sigma} + \gamma^{euler}$$

or:

$$\sigma \log P_{ij} = -\phi_{ij} c^s_{ij} + \beta \xi_{ij} \phi_{ij} U_j(\theta) - J_i + \sigma \gamma^{euler} \tag{31}$$

and

$$\sigma \log P_{ii} = \beta U_i(\theta) - J_i + \sigma \gamma^{euler} \tag{32}$$

Now, replace W_{ij} from (31) into the definition of U, (5) to get:

$$U_i(\theta) = -c^u_i + \lambda_i \pi_i + \sigma \lambda_i \sum_{j \neq i} G_{ij} \log P_{ij} - \sigma \lambda_i \gamma^{euler} + J_i$$

and substitute J_i from (32):

$$U_i(\theta) = -\frac{1}{1-\beta} c^u_i + \frac{\sigma}{1-\beta} \left((1-\lambda_i)\gamma^{euler} + \lambda_i \sum_{j \neq i} G_{ij} \log P_{ij} - \log P_{ii} \right) + \frac{1}{1-\beta} \lambda_i \pi_i$$

so that given the CCP's, U_i is an affine function of c_u and σ. Next, we replace this into the Hotz-Miller inversion (30) to obtain:

$$c^s_{ij} = \frac{\beta}{\phi_{ij}(1-\beta)} c^u_i - \frac{\beta}{1-\beta} \xi_{ij} c^u_j +$$

$$+\sigma \left(\frac{\beta}{1-\beta} \left(\xi_{ij} \left[(1-\lambda_j)\gamma^{euler} + \lambda_j \sum_{l \neq j} G_{jl} \log P_{jl} - \log P_{jj} \right] - \frac{1}{\phi_{ij}} \left[(1-\lambda_i)\gamma^{euler} + \lambda_i \sum_{l \neq i} G_{il} \log P_{il} - \log P_{ii} \right] \right) \right)$$

$$-\frac{\sigma}{\phi_{ij}} \log \frac{P_{ij}}{P_{ii}} + \frac{\beta}{1-\beta} \xi_{ij} \lambda_j \pi_j - \frac{\beta}{(1-\beta)\phi_{ij}} \lambda_i \pi_i$$

Note that

$$\frac{1}{(1-\beta)\phi_{ij}} = \frac{1-\beta(1-\xi_{ij})}{1-\beta} = 1 + \frac{\beta \xi_{ij}}{1-\beta}$$

and set $\rho_{ij} = \frac{\beta \xi_{ij}}{1-\beta}$, then $\frac{1}{(1-\beta)\phi_{ij}} = 1 + \rho_{ij}$.

We divide by σ:

$$\frac{c^s_{ij}}{\sigma} = \beta\left(1+\rho_{ij}\right) \frac{c^u_i}{\sigma} - \rho_{ij} \frac{c^u_j}{\sigma} - \left[\beta\left(1+\rho_{ij}\right)\lambda_i \pi_i - \rho_{ij} \lambda_j \pi_j\right] \frac{1}{\sigma} +$$

$$+\rho_{ij}\left[(1-\lambda_j)\,\gamma^{euler}+\lambda_j\sum_{l\neq j}G_{jl}\log P_{jl}-\log P_{jj}\right]-\beta(1+\rho_{ij})\left[(1-\lambda_i)\,\gamma^{euler}+\lambda_i\sum_{i\neq j}G_{il}\log P_{il}-\log P_{ii}\right]$$

$$-\frac{1}{\phi_{ij}}\log\frac{P_{ij}}{P_{ii}}$$

This is a linear system of full rank in the parameters $\{\frac{c^s_{ij}}{\sigma},\frac{c^u_i}{\sigma},\frac{1}{\sigma}\}$, since $\frac{c^s_{ij}}{\sigma}$ can be expressed with respect to $\{\frac{c^u_i}{\sigma},\frac{1}{\sigma}\}$. □

9.8 Algorithm for computing the steady state equilibrium

In this appendix we describe the algorithm employed to compute the steady state of our model to obtain the counterfactuals of Section 7.

1. Make an initial guess for $\{s^0, f^0, U^0\}$.
2. At each iteration m, inherit $\{s^m, f^m, U^m\}$
   (a) Update the ship's and exporter's optimal policies using the following steps for K times.[61]
      i. Solve for W^{m+1} from:
      $$W^{m+1}_{ij}=\frac{-c^s_{ij}+\xi_{ij}\beta U^m_i}{1-\beta\left(1-\xi_{ij}\right)}$$
      ii. Update J^{m+1} from:
      $$J^{m+1}_i=\sigma\log\left(\exp\frac{\beta U^m_i}{\sigma}+\sum_{j\neq i}\exp\frac{W^m_{ij}}{\sigma}\right)+\sigma\gamma^{euler}$$
      iii. Compute the equilibrium prices using
      $$\pi^m_{ij}=(1-\gamma)\left(\frac{1-\beta}{1-\beta+\gamma\beta\lambda^{f,m}_i}\right)\mu_{ij}+\gamma\frac{1-\beta\left(1-\lambda^{f,m}_i\right)}{1-\beta+\gamma\beta\lambda^{f,m}_i}\left(J^{m+1}_i-W^{m+1}_{ij}\right),$$
      iv. Update $\tilde{G}$:
      $$\tilde{G}^{m+1}_{ij}\equiv\frac{\exp\left(\frac{\lambda^f_i\left(\mu_{ij}-\pi^m_{ij}\right)}{1-\beta\delta\left(1-\lambda^f_i\right)}-\kappa_{ij}\right)}{1+\sum_{l\neq i}\exp\left(\frac{\lambda^f_i\left(\mu_{ij}-\pi^m_{ij}\right)}{1-\beta\delta\left(1-\lambda^f_i\right)}-\kappa_{il}\right)}$$

[61] K is chosen to accelerate the convergence, in the spirit of standard modified policy iteration methods.

v. Update U^m:

$$U_i^{m+1} = -c_i^u + \lambda_i E_{v,j} \pi_{ijv} + \lambda_i \sum_{j \neq i} \left(\frac{\tilde{G}_{ij}^{m+1}}{1 - \tilde{G}_{i0}^{m+1}} \right) W_{ij}^{m+1} + (1 - \lambda_i) J_j^{m+1}$$

vi. Obtain the ships ballast choices $\left(P_{ij}^{m+1} \right)_{i=1:I, j=1:I}$

3. Update to $\{\tilde{s}^{m+1}, \tilde{f}^{m+1}\}$ from:

$$\tilde{f}_i^{m+1} = \delta_i \left(f_i^m - m_i^m \right) + \mathcal{E}_i \left(1 - \tilde{G}_{i0}^{m+1} \right)$$

and

$$\tilde{s}_i^{m+1} = \sum_j P_{ji}^{m+1} \left(s_j^m - m_j^m \right) + \sum_j \frac{\tilde{G}_{ij}^{m+1}}{1 - \tilde{G}_{i0}^{m+1}} G_{ji}^{m+1} m_j^m$$

4. If $\left\| \tilde{s}^{m+1} - s^m \right\| < \epsilon$, $\left\| \tilde{f}^{m+1} - f^m \right\| < \epsilon$ and $\left\| U^{m+1} - U^m \right\| < \epsilon$, stop, otherwise update freights and ships as follows:

$$\begin{aligned} s^{m+1} &= \alpha s^m + (1 - \alpha) \tilde{s}^{m+1} \\ f^{m+1} &= \alpha f^m + (1 - \alpha) \tilde{f}^{m+1}, \end{aligned}$$

where α is a smoothing parameter.