# Shipping inside the Box: Containerization and Trade[*]

A. Kerem Coşar

University of Virginia
Department of Economics

Banu Demir

Bilkent University
Department of Economics

July 2018

## Abstract

We quantify the effect of container technology on transport costs and trade by estimating the modal choice between containerization and breakbulk shipping using micro-level trade data. The model is motivated by novel facts that relate container usage to shipment, destination and firm characteristics. We find container transport to have a higher first-mile cost and a lower distance elasticity, making it cost effective in longer distances. At the median distance across all country pairs, the container decreases variable shipping costs between 16 and 22 percent. Containerization explains a significant amount of the global trade increase since its inception: a quantitative exercise suggests that, in the absence of containers, Turkish and U.S. maritime exports in a typical sector to the average destination market would have been about two-thirds, and aggregate maritime exports 14 to 21 percent lower than what they are today, respectively.

# 1 Introduction

The introduction of containers in the second half of 1950s marked a major innovation in transportation: the standard container (referred to within the industry as "the box") improved efficiency by allowing automation in cargo handling, connecting sea transport with intermodal inland transport, and reducing spoilage/pilferage on and off the ship. All these benefits generated economies of scale and slashed transit times (Levinson, 2008; Hummels, 2007). Despite its ubiquity, the mechanisms through which containerization affected world trade are still unexplored. Understanding the drivers of container usage at the decision-making level is key to the measurement of transportation costs affecting the volume and pattern of international trade. We provide the first such analysis using micro-level data on Turkish exports at the firm, product and destination level for the year 2013.

We start by documenting novel facts from Turkish micro data and U.S. aggregate data: despite the perception that international maritime trade is now *highly* containerized, there is still an important margin of modal choice for exporters between container and breakbulk.[1] As of 2013, of the total maritime Turkish and U.S. exports of general cargo (excluding oil, fertilizers, ore, and grain) by weight, only 41 and 46 percent were containerized, respectively. By export value, these figures were 54 and 58 percent.[2] To the best of our knowledge, comparable data on container share in world maritime trade *by value* is not available. Available statistics on worldwide usage are usually *by weight* or *by volume*. For instance, Rua (2014) documents that the global share of containers in general cargo (i.e., excluding oil, fertilizers, ore, and grain) *by volume* reaches 70 percent by mid-2000s; see figure I in her paper.

The data shows large variation in container usage across firms, products and destinations. We find four patterns in this variation: first, it is by and large explained by firms, rather than by products and destinations. Second, container usage increases with distance to the destination. Third, container usage also increases with shipment size but decreases with unit prices. Finally, container

---

[1]Breakbulk is defined as shipping goods in bags, bales, packed in cartons or pallets in ships' hold, instead of in standardized containers. This should not be confused with bulk cargo such as grains, coal and ores.

[2]By dropping bulk commodities, the definition of general cargo eliminates some but not all the products for which container shipping is not feasible due to their physical properties. Such goods are typically transported by specialized vessels, such as roll-on/roll-off (ro-ro) ships for cars and trucks. To account for this, we also check the container shares by restricting the sample to containerizable goods according to two available classifications. Using the 1968 German Engineers' Society classification from Bernhofen, El-Sahli, and Kneller (2016), 54 and 53 percent of Turkish and U.S. 2013 maritime exports (by value) were containerized. Using a more recent and restrictive OECD definition by Korinek (2011), the respective figures become 75 and 67 percent.

usage increases with firm size and labor productivity. These findings imply that, conditional on physical feasibility due to product characteristics and the necessary infrastructure being available in both the origin and the destination, exporting firms still face a choice on the mode of maritime transportation and only some of them find it profitable to ship in containers.

Informed by these facts, we propose and estimate a model of self-selection into containerized shipping by heterogenous firms. In the first stage, we estimate the variable cost of container shipping relative to breakbulk without making any assumption on firms' productivity distribution, using observed firm-product-destination level export revenues by mode. In the second stage, we use these estimates along with additional structure on firm productivities and parameter values from the literature to recover relative fixed trade costs by mode.

Our contribution is threefold. First, to the best of our knowledge, this is the first paper to use micro-level data at the firm-product-destination level to estimate the structural parameters of shipping technologies in a model with heterogeneous firms making modal and product quality decisions. The firm margin in shipping has so far not been considered in the empirical international trade literature. The detailed micro-level data enables us to estimate the parameters of interest in a very precise way using demanding specifications and controlling for various factors that affect maritime transport modal choice such as product characteristics, destination country characteristics, and more importantly, self-selection of firms into destination countries.

Second, the structural estimation on micro-level data allows us to do quantitative counterfactual analysis to evaluate the impact of reduced trade costs on trade: In the absence of containers, Turkish and U.S. maritime exports in a typical sector to the average destination would have been about two-thirds, and aggregate maritime exports 14 to 21 percent lower than their level in 2013, respectively. In the case of full adoption of container technology, Turkish and U.S. maritime exports to the average destination would increase by 13 and 19 percent in a typical sector, respectively. We also show that these quantitative results are invariant to whether transport costs are additive or multiplicative as long as the former specification takes into account the endogenous quality choice for exports.

Third, we provide supporting evidence for the conjecture that container shipping is subject to a higher scale but has a lower distance elasticity, facilitating increased trade with more distant destinations. At the median distance across all country pairs (10,400 km), the container decreases

variable shipping costs between 16 to 22 percent. We corroborate this evidence using direct measures on insurance and freight charges on U.S. imports.

Our paper relates to an empirical literature that investigates the effect of technological advancements in transportation on trade. Using data from 19th century India, Donaldson (2018) estimates that railroads reduced the cost of trading, narrowed inter-regional price gaps, and increased trade volumes. Pascali (2017) estimates the impact of steamships on the first wave of trade globalization. Focusing on airplanes, Harrigan (2010) investigates how geography and the choice of shipping mode interact in shaping comparative advantages and trade patterns. Hummels and Schaur (2013) use variation in transport modes across US imports at the origin-product level to identify the ad-valorem equivalent time costs of shipping. Micco and Serebrisky (2006) estimate that the liberalization of air cargo markets reduces air transport costs by about 9 percent by enabling the efficiency gains. Clark, Dollar, and Micco (2004) estimate the cost-reducing effects of port efficiency and containerization in US maritime imports. Brancaccio, Kalouptsidi, and Papageorgiou (2017) analyze the effect of matching frictions on trade costs and volumes in bulk shipping.

As to the impact of containerization, Hummels (2007) estimates that doubling the share of containerized trade decreases shipping costs between 3 and 13 percent. Moreover, he finds no evidence of decline in maritime liner shipping price index over the past five decades (see his figure 4), conjecturing that unmeasured quality changes in transportation—faster, more precise delivery services—could explain this finding. In a study of port efficiency, Blonigen and Wilson (2008) find that a ten percentage-point increase in the share of containerized trade between US and foreign ports reduces import charges by around 0.6 percent. These seemingly low estimates pose a puzzle in the face of container shipping being praised as a technological revolution. In contrast, Bernhofen, El-Sahli, and Kneller (2016) find a large effect. Using a panel data of industry-level bilateral trade for 157 countries, they identify the effect of containerization through countries' differential dates of adoption of container facilities. Their results suggest that containerization contributed more to the increase of world trade during the 1962-1990 period than trade policy changes such as GATT tariff cuts and regional trade agreements.

Finally, our model shares with Rua (2014) the widely-used element of depicting a technology—in this case containerization—which lowers variable costs in return for a high fixed cost in a Melitz (2003) type heterogeneous firm model of trade (e.g., Bustos 2011). As a result, only more productive

firms prefer containerized shipping to break-bulk. Rua (2014) embeds this firm-level problem into a model of country-level technology adoption in order to derive testable implications on the international diffusion of container technology. Using country-level data, she finds that fixed costs and network effects are the main determinants of the adoption of containerization. In contrast, we focus on the firms' choice of shipping mode and estimate the cost structure of shipping in order to quantify the role of containerization in trade volumes. The structural estimation in turn allows us to do quantitative counterfactual analysis.

The next section introduces the data and documents the facts motivating our model and estimation.

## 2  Data and Four Empirical Facts

### 2.1  Basic Features of the Data

The confidential micro-level data accessed from the premises of the Turkish Statistical Institute is based on customs forms and contains all Turkish export transactions that took place in 2013.[3] Each transaction records the identity of the exporting firm, 8-digit Harmonized System (HS) product code, value, weight (in kilograms) and quantity (in specified units, e.g. pair, number, liter, etc.) of the shipment, destination country, and the mode of transportation (truck, rail, vessel, air, pipeline). A separate binary variable informs us whether goods were shipped in a container or not. For reasons related to disclosure restrictions, our data excludes HS heading 27—mineral fuels, oils, waxes, and bituminous substances—and HS heading 93—arms and ammunition, parts and accessories. In the baseline analysis, we keep all other products and capture their containerizability by fixed effects. Following common practice, we also drop small transactions (firm-product-destination exports with an annual value of less than USD 5,000) from the dataset as they are likely to introduce noise into our estimates.

Unsurprisingly, containerization is associated with maritime shipments: 97.8 percent of all containerized exports by value are by sea. Only 0.3 percent and 1.2 percent of air and land exports are containerized, respectively. Therefore, we restrict the sample to vessel exports to coastal coun-

---

[3]The entire dataset spans 10 years from 2003 to 2013. We work with the latest available year when the container usage in Turkey peaks and the effects of the Great Recession subsides.

tries. Excluded landlocked destinations constitute a small share of exports (8 percent), and an even smaller share of containerized exports (1.9 percent).

Table 1 presents further relevant summary statistics from our data. Our dataset covers 27,241 exporters, 5,557 8-digit HS products, and 139 destination countries. The top panel of the table shows the fraction of observations with no containerization or full containerization. The respective fractions are small at the destination or product level: share of containerized exports lies strictly between zero and one to almost all destinations in about 75 percent of 8-digit HS product codes. Nevertheless, the extensive margin contributes significantly to the variation in container use at the firm-level: about one-third of Turkish exporters never shipped in containers, and another one-third shipped only in containers in the year 2013. The share of containerized exports is either zero or one for about 90 percent of observations (firm-product-destination level). The lower panel of table 1 presents the relevant summary statistics from a smaller sample that includes intrinsically container-izable products only. We now proceed to a more nuanced analysis that controls for compositional effects in order to tease out salient patterns on container usage from our data.

## 2.2 Four Facts about Container Usage

We now present four facts about the use of containerization. In each case, we first summarize the stylized evidence and then explain the underlying analysis. These facts subsequently guide our modeling choices in estimating the parameters of shipping technologies.

To check for similar patterns, we use publicly available aggregate U.S. maritime trade data at the level of 10-digit HS products, trade partner, port, and container use (Schott, 2008). While it does not inform us about firms, the import side of the U.S. data reports freight and insurance charges, which we subsequently use in the empirical analysis.

**Fact 1: A large share of the variation in containerization is explained by firms, rather than by products and destinations.**

As reported above, around half of all annual vessel exports are containerized, with varying fractions of full or no containerization across products, destinations and firms. We now explore the components of the overall variation in container usage. A priori, one may expect product characteristics to be the primary determinant of whether a shipment will be containerized or not. After all, bulk

5

commodities such as ores or grains are hardly fit for the standard container, whereas anecdotes of global trade convey the image that some goods, such as apparel and consumer electronics, are stackable and thus highly containerized. Similarly, one may expect that the characteristics of the destination country, such as the existence of the appropriate infrastructure or level of development, to be key determinants since the technology is presumably expensive and dependent on specialized ports and intermodal logistics.

For visual inspection, we plot the intensive-margin distribution of container share—i.e., excluding extensive margin mass points at zero and one—in vessel exports aggregated over months in figure 1. Evidently, there is large variation across shipments (Panel A), with high heterogeneity across all dimensions (Panels B-D). For statistical analysis, we run a series of fixed-effect regressions and analyze their fit in table 2. We denote firms by $a$, products by $j$, and destination countries by $d$.[4] The direct effects in the first column are the adjusted $R_k^2$'s from regressing container shares, including the mass points at zero and one, on single fixed-effects $k \in \{a, j, d\}$ in the top panel and on pair fixed effects $k \in \{aj, ad, jd\}$ in the bottom panel. In order to purge out compositional effects, we report in the second column the coefficients of partial determination isolating the unique contribution of each component.[5]

We first conduct the analysis using the Turkish micro export data, and report the results in panel A of table 2. In terms of individual effects, product categories and destinations have relatively little explanatory power. Firm-specific factors, both in terms of direct and isolated effects, account for a substantial fraction of the variation.[6] Looking at joint effects, the partial coefficient of firm-destination pairs equals 0.9, suggesting that containerization in international trade is predominantly determined by firms' modal choices that vary across countries.

Using the aggregate U.S. export data, which does not contain information about exporting firms, pairwise combinations of individual effects explain no more than 15 percent of the variation

---

[4]To account for potential seasonal effects in container ship schedules, we pair destinations with months but suppress the time subscript, i.e., $d$ refers to a destination-month pair.

[5]That is, for each $k$, we first regress container shares on fixed effects $(\mu_k, \mu_{-k})$ and find the fit. For instance, for firms ($k = a$), that would be the $R_{a,jd}^2$ of the regression $ContShr_{ajd} = \mu_a + \mu_{jd} + \epsilon_{ajd}$, where $\mu_{jd}$ represents product-destination pair fixed effects. We then drop the factor of interest $k = a$ to find $R_{jd}^2$ from the regression $ContShr_{ajd} = \mu_{jd} + \epsilon_{ajd}$. The coefficient of partial determination is the ratio of the difference $R_{a,jd}^2 - R_{jd}^2$ over the variation unexplained $1 - R_{jd}^2$ without contribution of the factor of interest $a$. This method is robust to sequence sensitivity when comparing the fit across specifications that add fixed effects in a progressive manner (Gelbach, 2016).

[6]These factors do not include firm location and access to ports as all international ports in Turkey also have container terminals. In other words, firm location does not matter for the relative access to container shipping.

in container usage (panel B). In the Turkish data, pairwise combinations involving firms explain as much as 91 percent of the variation. The large portion of the variation left unexplained in the U.S. data is consistent with the importance of firms.

**Fact 2: Container usage is increasing in distance.**

Given the substantial contribution of the destination margin to the firm-destination variation in container usage presented in panel A of table 2, we aggregate Turkish and U.S. export data to a common sample of destination countries. Regressing container shares in maritime trade to the logarithm of the sea distance to the trade partner, we find that containerization increases with distance (table 3). The positive and significant distance gradients are of remarkably comparable magnitudes for the two exporter countries and remain robust to controlling for other destination characteristics such as income per capita, adjacency and being in a free-trade agreement with Turkey or the U.S.

**Fact 3: Container usage is increasing in shipment size and decreasing in unit value of the shipment.**

Container shipping displays economies of scale due to high infrastructure costs and the large vessel sizes required to utilize these investments (see Stopford, 2009, chapter 13). Also called the "first-mile cost" in logistics, the decline in unit costs with scale and distance is a key characteristic of how transportation and shipping technologies affect trade. This cost structure is plausibly passed on from shipping companies to trading firms—as corroborated by minimum shipment requirements and differential pricing practices for full-container load and less-than-container load shipments. We can thus expect container usage and its geographic determinants to correlate with parcel size.

Table 4 confirms this conjecture: container usage is increasing in transaction size. It also shows that container usage is correlated with unit value of shipments, defined as shipment value per quantity measured in physical units.[7] In particular, controlling for shipment weight, lower unit values within a firm-product pair are associated with higher container usage. Second and third columns split the sample to differentiated and non-differentiated goods according to the Rauch (1999) classification. Results show that the negative association of container usage with unit values holds for differentiated goods only, suggesting a relationship between firms' quality and transport

---

[7]Unit of measurement does not vary within a given 8-digit HS product code.

mode choices. In particular, if transport costs are additive and unit shipping costs are higher in breakbulk, for a given firm-product pair, Alchian-Allen effect implies a negative relationship between container usage and quality. To account for this mechanism, we will incorporate endogenous quality differentiation to the model presented in the next section.

**Fact 4: Container usage is increasing in total sales, employment and productivity of the exporter, with no economies of scope.**

Per Fact 1, the most significant factor in explaining modal choice is the identity of exporting firms. Theoretically, heterogeneity in productivity or quality, together with fixed costs of container shipping, could induce firms to sort into using the technology (Rua, 2014). To investigate this, table 5 reports the results from regressing container usage on various firm-level characteristics. Across all specifications, it is important to control for shipment size to ensure that the effect is not going through the shipment-specific scale economies documented in Fact 3, and for compositional effects through product-destination fixed effects. Total firm exports, employment, sales, and sales per worker are all positively and significantly correlated with container usage (columns 1-4). When total exports and sales per worker are jointly controlled for in column 5, only the former is significant, suggesting that total export volume is a better proxy for firm selection in our data. In column 6, we include the number of 8-digit HS products exported by the firm to a given destination and find no evidence of economies of scope for container usage.

In concluding this section, we reiterate that the micro-level trade data show substantial variation in containerization within narrow product categories and destination countries, with the overall variation largely accounted for by firms. Moreover, container usage is systematically increasing in firm productivity and distance to the destination. Next section presents a transport mode choice model for heterogeneous firms that is consistent with these stylized facts and is amenable to estimation.

# 3    Model

To explain the choice between containerization and breakbulk at the firm-level, we now present a simple partial-equilibrium trade model with heterogeneous firms. Taking as given the demand in

export destinations, monopolistically competitive firms make optimal pricing, quality and mode of transport choices. By depicting containerization as a high fixed-low marginal cost shipping technology, we share some widely-used elements with the model proposed by Rua (2014). In line with mounting evidence (Hummels and Skiba, 2004; Irarrazabal, Moxnes, and Opromolla, 2015), we assume per-unit transport costs as our baseline specification. To take into account the potential Alchian-Allen effect—increased relative demand for high-quality products in the presence of additive transport costs—we incorporate quality differentiation to the model. This framework helps us characterize the conditions under which there is positive selection into container usage, and yields estimable equations that pin down the structural parameters of the two transport technologies. We later present an alternative version of the model with iceberg trade costs and without quality differentiation. As will be demonstrated and explained below, the choice between the two versions of the model does not affect our empirical strategy, and the predictions obtained from a counterfactual exercise about the effect of containerization on trade remain invariant to the specification of transport costs.

*Demand*    There is one source country exporting to multiple destinations indexed by $d$. It is populated by a large number of firms, which are heterogenous in productivity $a$ and produce a continuum of horizontally and vertically differentiated varieties. As now standard in the literature, we use the productivity index to represent varieties produced by these monopolistically competitive firms.

Consumer preferences in destination $d$ are represented by a quality-augmented CES aggregate as in Baldwin and Harrigan (2011) and Kugler and Verhoogen (2012):

$$Q_d = \left[ \int \left[ z_d(a) q_d(a) \right]^{\frac{\sigma-1}{\sigma}} dG(a) \right]^{\frac{\sigma}{\sigma-1}},$$

where $z_d(a)$ denotes the quality, $\sigma$ the elasticity of substitution, $q_d(a)$ quantity consumed and $G(a)$ the distribution of firm productivity. Utility maximization yields the following demand for each differentiated variety:

$$q_d(a) = X_d P_d^{\sigma-1} z_d(a)^{\sigma-1} \tilde{p}_d(a)^{-\sigma}, \tag{1}$$

where $X_d$ is the spending allocated to imports from the source country in destination $d$, $\tilde{p}_d(a)$ is the consumer price, and $P_d$ is a quality-augmented CES price index defined as:

$$P_d^{1-\sigma} = \int \left(\frac{\tilde{p}_d(a)}{z_d(a)}\right)^{1-\sigma} dG(a).$$

*Supply*    Consumer prices (c.i.f.) differ from the producer prices (f.o.b.) because of trade costs, which have a specific component $t_d^m$ that varies by destination and endogenously chosen transport mode $m = \{b, c\}$ (for break-bulk or container), and an exogenous ad-valorem component $\tau_d > 1$ that depends only on the destination:[8]

$$\tilde{p}_d^m(a) = \tau_d \big[p_d^m(a) + t_d^m\big]. \tag{2}$$

In modeling quality production, we follow Feenstra and Romalis (2014) and assume that a firm with productivity $a$ uses $l$ units of labor (expressed in efficiency units of labor) to produce one unit of product with quality $z(a)$:

$$z(a) = (a \cdot l)^\theta,$$

where $\theta \in (0, 1]$ represents diminishing returns in the production of quality. The quality production function implies marginal cost of production given by

$$C(a, z) = \frac{z(a)^{1/\theta}}{a} w, \tag{3}$$

where $w$ is the unit cost of labor input and the numéraire. We now describe firms' optimal pricing, quality and transport mode decisions.

*Optimal Price and Quality*    Given mode $m$, firms maximize operating profits by solving

$$\pi_d^m(a) = \max_{p_d, z} \left\{ q_d^m(a) \cdot \big[p_d^m(a) - C(a, z)\big] \right\},$$

where $q_d^m(\cdot)$ captures the dependence of demand (equation 1) on $m$ through the consumer price

---

[8]A vast majority of countries apply tariffs on transport inclusive prices—see footnote 10 in Feenstra and Romalis (2014).

(2). First-order condition with respect to price yields

$$p_d^m(a) = \frac{\sigma C(a, z_d) + t_d^m}{\sigma - 1}.$$  (4)

Similarly, first-order condition with respect to quality yields

$$z_d^m(a)^{1/\theta} = \frac{\theta}{1 - \theta} \cdot a \cdot t_d^m,$$  (5)

which varies by destination due to transport costs. Substituting (5) into (3) and then into (4) gives the following optimal price as a function of unit transport costs:

$$p_d^m(a) = \chi \cdot t_d^m,$$  (6)

where $\chi = \frac{1}{\sigma - 1}\left(\frac{\sigma\theta}{1-\theta} + 1\right)$ is common to all firms.[9] Equation (6) describes a simple linear relationship between f.o.b. prices and specific transport costs.

Given the profit-maximizing price and quality, the revenue of a firm with productivity $a$ exporting to destination $d$ using transport technology $m$ is given by:

$$r_d^m(a) = \Theta_d \cdot a^{\theta(\sigma-1)} \cdot (t_d^m)^{-(\sigma-1)(1-\theta)},$$  (7)

where $\Theta_d = \chi(\chi+1)^{-\sigma}\left(\frac{\theta}{1-\theta}\right)^{\theta(\sigma-1)} X_d P_d^{\sigma-1} \tau_d^{-\sigma}$. Subtracting variable costs at the optimal quality choice and rearranging terms, operating profits are given by

$$\pi_d^m(a) = \frac{\chi + 1}{\chi\sigma} \cdot r_d^m(a).$$  (8)

*Choice of Transport Mode*    A firm exporting to destination country $d$ pays a mode-specific

---

[9]All else being equal, optimal quality (equation 5) is increasing in the specific cost $t$. To see the intuition, note that the (fob) origin price elasticity of quality-adjusted demand $zq$ is $\epsilon = \sigma \cdot p/(p+t)$. At a given price $p$, the firm perceives its demand to be less elastic for higher values of $t$. To meet the quality-adjusted demand $zq$, with $t$ being paid per unit of $q$, the firm would vary $z$ as well. Of course, price and quality are both set optimally, which leads to an equalization of perceived demand elasticities across firms at $\epsilon = \sigma \cdot \chi/(\chi+1)$. We thank the referee for pointing this out.

fixed cost $f_d^m > 0$. Net export profit for using each transport mode is simply:

$$\Pi_d^m(a) = \pi_d^m(a) - f_d^m. \tag{9}$$

A firm exports to a destination in a container if $\Pi_d^c(a) \geq \Pi_d^b(a)$. The following condition is necessary and sufficient to induce productivity-based selection into containerization and thus make the model consistent with Fact 4:

$$\frac{f_d^c}{f_d^b} > \left(\frac{t_d^c}{t_d^b}\right)^{-(\sigma-1)(1-\theta)}. \tag{10}$$

Similar to Rua (2014), this restriction on relative fixed and variable trade costs is a modified version of the condition for selection into exporting in Melitz (2003): only sufficiently productive exporters choose container to breakbulk shipping by trading off higher fixed costs with lower variable costs. Note that for the relevant levels of variable costs satisfying $t_c < t_b$, a higher $t_c$ pushes toward selection by relaxing the constraint (10) on relative fixed costs: as the variable cost advantage of containerization diminishes, firms have to be more productive to self-select into container shipping. Moreover, this force gets stronger for lower returns to quality production $\theta$: the costlier it is to produce quality, the more productive an exporter has to be in order to pass the selection threshold.

Finally, the marginal exporter uses break-bulk and can be characterized by $\tilde{a}_d^b$ satisfying $\Pi_d^b(\tilde{a}_d^b) = 0$. The marginal *containerized* exporter $\tilde{a}_d^c$ is defined by $\Pi_d^c(\tilde{a}_d^c) = \Pi_d^b(\tilde{a}_d^c)$, and satisfies $\tilde{a}_d^c > \tilde{a}_d^b$.

## 4   Estimation

The model of firm selection into exporting and containerization is based on two novel sets of parameters that we wish to estimate: mode-dependent variable and fixed export costs $(t_d^m, f_d^m)$. Progressing in two stages, we first parameterize and estimate relative variable transport costs using observed firm-product-destination level export revenues by mode of shipping, controlling for selection through appropriate fixed effects. This flexible approach allows us to estimate variable costs without making a distributional assumption for firm productivity. In the second stage, we use these estimates along with additional structure on firm productivities to recover relative fixed trade costs by mode.

## 4.1 Estimation Strategy

To derive estimating equations from model-based firm revenue and mode-choice rules, we specify variable transport costs by:

$$t_d^m = \bar{t}_m \cdot dist_d^{\delta_m}, \tag{11}$$

where $dist_d$ is the distance to destination country $d$. The parameters $\bar{t}_m$ and $\delta_m$ capture the mode-specific first-mile costs and distance elasticities, respectively. Based on Fact 2 documented above, we anticipate containerization to have a higher first-mile cost ($\bar{t}_c > \bar{t}_b$) and a lower distance elasticity ($\delta_c < \delta_b$).

Under this parameterization, log revenues can be written as (see Appendix A1 for details)

$$
\begin{aligned}
\ln r_d(a) = \ln\left(\frac{r_d^b(\tilde{a}_d^b)}{(\tilde{a}_d^b)^{\theta(\sigma-1)}}\right) &+ (\sigma-1)\theta \ln a \\
&+ (1-\sigma)(1-\theta)\big[\ln(\bar{t}_c) - \ln(\bar{t}_b)\big] \cdot CONT \\
&+ (1-\sigma)(1-\theta)(\delta_c - \delta_b) \cdot CONT \cdot \ln dist_d, \tag{12}
\end{aligned}
$$

where the indicator function $CONT$ denotes container usage, i.e., $CONT = 1$ for $a \geq \tilde{a}_d^c$ and zero otherwise.

Equation (12) forms the basis of our estimation. To implement an empirical specification, we have to consider two issues. First, each firm produces a single variety $j$ in the model, whereas in the data, many firms operate in multiple sectors and export multiple products belonging to a given sector $s(j)$. As discussed above, Fact 4 motivates our abstraction from economies of scope in container usage: multi-product firms make independent shipping mode decisions for each product. There may be, however, economies of scope in other activities leading to the emerge of multi-product firms. Accordingly, we group 8-digit HS products under 4-digit HS sectors, and use appropriate firm and sector fixed effects to distinguish multi-product exporters' sales in different sectors.[10] This approach also allows us to take into account demand variation across sectors in a given destination, i.e., $X_{sd}$ denotes consumer spending on sector $s$. Second, to attenuate the noise in the monthly data, we aggregate firm-product-destination-mode level export sales to the annual level.

The first term in (12) is common to all firms in sector $s$ exporting to destination $d$, and thus can

---

[10]For instance, HS heading 8703 refers to "Motor vehicles for the transport of persons," which we consider as a sector. Finer 8-digit levels distinguish varieties according to body type, ignition type and engine capacity.

be captured by sector-destination fixed effect ($\alpha_{sd}$). The second term contains firm productivity, and thus can be captured by a firm fixed effect ($\alpha_a$). Our estimating equation can be written as:

$$
\begin{aligned}
\ln r_{ajdm} &= \underbrace{(1 - \sigma)(1 - \theta)\big[\ln(\bar{t}_c) - \ln(\bar{t}_b)\big]}_{\eta_1} \cdot CONT_{ajdm} \\
&+ \underbrace{(1 - \sigma)(1 - \theta)(\delta_c - \delta_b)}_{\eta_2} \cdot CONT_{ajdm} \cdot \ln dist_d \\
&+ \alpha_{sd} + \alpha_a + \epsilon_{ajdm},
\end{aligned}
\tag{13}
$$

where the indicator function $CONT$ is a dummy taking the value one if there is a containerized shipment in the observed firm-product-destination level flow. Equation (13) augments the theory-implied revenue equation (12) by an error term $\epsilon$, which captures i.i.d. revenue shocks realized after pricing and shipping decisions have been made, e.g., exchange rate movements, loss/damage of cargo in transit, or risk of non-payment. For another example where revenue shocks are orthogonal to the shipping decision, consider a buyer-seller pair contracting upon a minimum level of sales with a particular delivery time, such as an initial order to be delivered in a certain month, with the option of ordering an additional amount to the same shipment. The initial level of demand irreversibly determines the optimal shipping mode. Then the shock is realized and the exact level of demand— i.e., whether the additional order comes in or not—is determined. This affects revenue but not the pre-determined shipping mode.

## 4.2   Estimation Results

Table 6 presents the results from estimating equation (13). Our dependent variable is measured in terms of deviations from the respective 8-digit HS product means, $\ln(r_{ajdm}/\bar{r}_j)$, where $\bar{r}_j$ denotes the mean value of exports at the product-level. We do this to control for differences in price levels across products.[11]

In the first column, we start with the direct effect of containerization without the interaction term to gauge whether containerization is associated with larger trade flows. Controlling for demand-related factors with sector-destination fixed effects and supply-related factors (e.g., firm productivity) with firm fixed effects, containerized exports are indeed 35 percent ($e^{0.296} - 1$) larger

---

[11]Estimating with product fixed effects yields very similar estimates of both coefficients of interest.

than break-bulk exports.

The second column of table 6 presents results from the estimation of equation (13). Contrasted with the first column, adding the interaction between the container dummy and distance to destination reverses the sign of the coefficient $\eta_1$ on $CONT_{ajdm}$ to negative. This is consistent with our hypothesis that containerization has a higher first-mile cost than breakbulk shipping: since $\sigma > 1$ and $\theta < 1$, a negative $\eta_1$ estimate implies $\bar{t}_c > \bar{t}_b$. The coefficient $\eta_2$ on the interaction between $CONT_{ajdm}$ and $\ln dist_d$ is estimated to be positive and statistically significant, which implies a smaller elasticity of container shipping with respect to distance, $\delta_c < \delta_b$. In column 3, we replace firm fixed effects with firm-sector level fixed effects to account for the possibility that multi-product firms may have different productivities in different sectors (Eckel and Neary, 2010; Bernard, Redding, and Schott, 2011), as well as potential heterogeneity in productivity distributions or in the elasticity of substitution across sectors. The estimates of both $\eta_1$ and $\eta_2$ remain stable.

The specifications presented so far control for demand and supply factors related to sector-destination and firm-productivity pairs. While the latter fixed effects control for productivity-induced firm selection into containerization, they do not adequately control for selection at the relevant level of modal decision-making: a positive revenue shock at the firm-sector-destination level would increase the probability of containerization, creating an upward bias in the estimate of $\eta_1$ and driving it toward zero. Note that a firm operating in sector $s$ would prefer containerized exports if $\Pi^c_{sd}(a) \geq \Pi^b_{sd}(a)$, where net profits depend on revenues as derived in equation (9). Using the expressions for revenues in Appendix A1, profit gains from shipping in a container can be derived as follows:

$$\Pi^c_{sd}(a) - \Pi^b_{sd}(a) \;=\; \frac{\chi + 1}{\chi \sigma} \cdot \left[ \left( \frac{t^c_d}{t^b_d} \right)^{-(\sigma-1)(1-\theta)} - 1 \right] \cdot (f^c_d - f^b_d) \cdot r^b_{sd}(a).$$

Since the expression above varies at the firm-sector-destination level, selection into containerization can be accounted for by replacing sector-destination and firm-sector fixed effects in the estimating equation (13) with firm-sector-destination fixed effects $\alpha_{asd}$.

In this specification, the parameters of interest, $\eta_1$ and $\eta_2$, are identified from variation in container usage within a firm-sector-destination triplet across products. We can consistently estimate the parameters as long as firms face revenue shocks that do not systematically vary with the mode

15

of transport.

The fourth column of table 6 presents the results. Compared to the estimates in column 3, estimates of both $\eta_1$ and $\eta_2$ are larger in absolute value. In particular, the estimate of $\eta_1$ more than doubles (in absolute value) when selection into containerization is accounted for. This result is consistent with our prior that ignoring selection into containerization would drive the estimate of $\eta_1$ toward zero.

*Robustness Checks*    Table 7 presents a number of robustness checks. Columns 1 and 2 show the results from re-estimating our preferred specification (last column of table 6) on samples that exclude from the baseline intrinsically non-containerizable products based respectively on Korinek (2011) and Bernhofen, El-Sahli, and Kneller (2016). The coefficient estimates remain very close to their baseline values. We infer from this robustness check that the fixed effects included in baseline already take into account the nature of the products in terms of their containerizability. In another robustness check presented in column 3 of table 7, we investigate whether within-firm quality differentiation affects our baseline estimates. A potential concern is that firms might ship higher quality products to richer or more distant destinations, and such selection could interact with their choice of shipment technology and thus affect our estimates. To address this concern, we restrict the estimation sample to non-differentiated products, based on the classification suggested by Rauch (1999), since we expect that only firms in differentiated-good sectors to exhibit quality-based competence (Eckel, Iacovone, Javorcik, and Neary, 2015). The robustness of results rules out the possibility that our baseline estimates are driven by within-firm quality differentiation across destinations.

Column 4 of table 7 adds finer fixed effects (product-destination and firm-product) to the firm-sector-destination level fixed effects in the baseline. This specification relies on variation in the use of containers at a more disaggregated level to identify the coefficients of interest. The similarity of the estimates to the baseline increases our confidence that the latter identifies the coefficients of interest from variation in container usage within a firm-sector-destination triplet across products, relying mainly on product-specific random revenue shocks.

In the last column of table 7, we allow the relative distance elasticity to be piece-wise linear. While the difference in distance elasticity between breakbulk and container shipping increases with

the bilateral distance, the fit does not seem to improve compared to the parsimonious functional form we assume in the baseline.

In the next section, we will use our preferred estimates from the last column of table 6, parameter values from the literature, and further moments from the data to recover the unobserved relative variable and fixed costs of containerization. Recovering these costs will allow us to undertake model-consistent counterfactuals, yielding predictions for the contribution of containerization to the volume of trade.

## 5 Recovering Trade Costs

To recover transport technology parameters $(\bar{t}_m, \delta_m)$ from the estimates of $(\eta_1, \eta_2)$ in equation (13), we need to quantify $\sigma$ and $\theta$. As typical in the literature (Anderson and Van Wincoop, 2004; Coşar and Demir, 2016), the elasticity of substitution $\sigma$ cannot be separately identified from the distance elasticity of trade costs. In our case, an additional parameter $\theta$, capturing the supply of quality, enters the picture.

To proceed, we exploit further moments of the data related to the intensity of container usage in the aggregate and in the extensive margin. Under the assumption that the unconditional firm productivity $a$ is drawn from a Pareto distribution with domain $a \in [1, \infty]$ and shape parameter $k$ satisfying $k > \theta(\sigma - 1)$ similar to Chaney (2008), we can derive analytic expressions for destination-level share of containerized exports, $\mu_d$, and the fraction of firms using container shipping, $\mu_d^{ext}$ (see Appendix A2 for the details):

$$\mu_d = \frac{\int_{\tilde{a}_d^c}^{\infty} r_d(a) dG(a)}{\int_{\tilde{a}_d^b}^{\infty} r_d(a) dG(a)} = \frac{(\Delta t_d)^{-(\sigma-1)(1-\theta)} \left(\frac{\Delta f_d - 1}{(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{\theta(\sigma-1)-k}{\theta(\sigma-1)}}}{1 + \left(\frac{\Delta f_d - 1}{(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{\theta(\sigma-1)-k}{\theta(\sigma-1)}} \cdot \left[(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1\right]}, \quad (14)$$

and

$$\mu_d^{ext} = \frac{\int_{\tilde{a}_d^c}^{\infty} dG(a)}{\int_{\tilde{a}_d^b}^{\infty} dG(a)} = \left(\frac{\Delta f_d - 1}{(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{-k}{\theta(\sigma-1)}}, \quad (15)$$

where $\Delta f_d = f_d^c / f_d^b$ denotes relative fixed costs, and $\Delta t_d = t_d^c / t_d^b$ denotes relative variable trade costs. Using the functional form (11) for transport costs,

$$\Delta t_d = \frac{t_d^c}{t_d^b} = \left(\frac{\bar{t}_c}{\bar{t}_b}\right) \cdot dist_d^{(\delta_c - \delta_b)}. \quad (16)$$

17

The set of parameters to be calibrated is $(\sigma, \theta, k)$ as well as relative fixed costs $\Delta f_d$. The elasticity of substitution $\sigma$ and Pareto parameter $k$ are widely estimated in the literature. We take $\sigma = 4$ and $k = 4.25$ from Melitz and Redding (2015).

The following procedure pins down the quality production parameter $\theta$ along with relative fixed costs: from the data, we take the empirical moments $(\mu_d, \mu_d^{ext})$ as the median across sectors for each destination. For all values of $\theta \in (0, 1]$, we back out $(\bar{t}_c/\bar{t}_b)$ and $(\delta_c - \delta_b)$ from $(\eta_1, \eta_2)$ estimates of equation (13). Given distances, we use equation (16) to construct destination-specific relative variable transport costs, $\Delta t_d$. To back out $\Delta f_d$, we plug $\Delta t_d$ and the parameter values $(\sigma, k, \theta)$ into equation (14) and use the empirical moment $\mu_d$ on the left hand side. We then use $(\Delta t_d, \Delta f_d)$ along with parameter values $(\sigma, k, \theta)$ in equation (15) to calculate the model-implied extensive margin moment. The value of $\theta$ is picked such that the median of this moment across destinations matches its empirical value of $\text{median}_d \{\mu_d^{ext}\} = 0.74$, yielding $\theta = 0.267$. The resulting correlation between the observed and model-implied extensive margin across destinations is about 0.78.

*Variable Trade Costs*     Given $\sigma = 4$, the calibrated value $\theta = 0.267$ and estimates of $(\eta_1, \eta_2)$ from the last column of table 6 imply $(\bar{t}_c/\bar{t}_b) = 1.27$, i.e., the first-mile cost of container shipping is about 27 percent larger than that of breakbulk shipping. While it has a higher first-mile cost, container shipping has a smaller distance elasticity: $\delta_c - \delta_b = -0.05$.

Top panel of figure 2 plots $\widehat{\Delta t_d}$ against distance. The solid line is drawn using baseline parameter values, while the dashed line sets $\sigma = 6$ (and re-calibrates $\theta$) to check robustness to an empirically relevant higher value of $\sigma$. The negative gradient of relative variable container costs with respect to distance is consistent with the observed pattern in the data that container usage is increasing in distance to the destination (Fact 2). Using the benchmark estimates, variable cost savings from containerization reach 24 percent when the distance variable reaches 20,000 km. Cost savings are large for major trading pairs, amounting to 19.5 percent for Germany-USA and 22 percent for China-USA. For the latter pair, the lower bound for cost savings implied at the higher level of $\sigma = 6$ is around 13 percent.

Note that the combination of a higher first-mile cost and a lower distance elasticity implies that container shipping becomes cheaper beyond a breakeven distance. The horizontal line in the top panel of figure 2 marks the breakeven distance implied by our estimates, which is 103 km. This rather short breakeven distance is consistent with the raw data in that all destination countries, however close they are in proximity to Turkey, receive some containerized maritime exports (see the first row of table 1). Concurrently, it is consistent with the importance of firms, rather than destinations, in explaining container usage (Fact 1): if all destinations are situated beyond the breakeven distance, which is the case in our data, the large variation in container usage should come from firm-destination level heterogeneity. In the model, firm selection into container

shipping is driven by relative destination-dependent fixed cost of exporting $\Delta f_d = f_d^c / f_d^b$, which we present next.

*Fixed Trade Costs*    Bottom panel of figure 2 plots the histogram of calibrated relative fixed costs of containerized exports. Fixed cost of containerization is 40 to 100 percent higher than that of breakbulk, with a median of 70 percent across destinations.

Several channels could justify higher fixed costs associated with containerization. For instance, container links between ports are less frequent than break-bulk links. This implies, for any given shipment, the exporter has to spend additional effort to better manage the production and inventory scheduling. Another source is the importance of transaction-specific scale in container shipping: if the shipment size is large enough to fill a container (full-container-load), firms can schedule door-to-door shipping services. Otherwise, if the shipment is less-than-container load, exporters have to purchase additional services from freight-forwarders who consolidate and store shipments at ports. Due to the cargo handling involved, such services typically involve additional costs.[12]

*Iceberg Specification of Trade Costs*    The baseline model assumes that trade costs are composed of an ad-valorem part $\tau_d$ and an additive part $t_d^m$. We now investigate the implications of assuming that trade costs take the multiplicative "iceberg" form. Since multiplicative trade costs do not affect firm's choice of product quality, we set $z(a) = 1$ for all firms. We assume that trade costs can be written as the product of two terms: destination tariffs that do not vary between shipping modes $\tau_d$, and distance-mode dependent costs $t_d^m \geq 1$, such that $T_d^m = \tau_d \cdot t_d^m$ units of a good must be shipped to destination $d$ in order for one unit to arrive.

With this specification, consumer prices in destination $d$ are given by:

$$\tilde{p}_d^m(a) = T_d^m \cdot p(a).$$

Since we abstract from quality differentiation, firms maximize profits with respect to price only, yielding

$$p_d^m(a) = \frac{\sigma}{\sigma - 1} \cdot \frac{1}{a} \cdot T_d^m.$$

Given demand in destination $d$ for the variety exported by firm $a$, we can write firm revenues as

$$r_d^m(a) \;=\; \Theta_d \cdot a^{\sigma-1} \cdot (t_d^m)^{1-\sigma}, \tag{17}$$

---

[12]We abstract from potential heterogeneity in fixed costs across firms. Note that this does not bias the estimation of relative variable costs, which is identified from the variation in modal choices within firm-sector-destinations.

where $\Theta_d = \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} X_d P_d^{\sigma-1} \tau_d^{1-\sigma}$. Condition (10) for productivity-based selection into containerization becomes

$$\frac{f_d^c}{f_d^b} > \left(\frac{t_d^c}{t_d^b}\right)^{-(\sigma-1)}.$$

As in Appendix A1, we can write revenues of firm $a$ from its export sales to a destination country $d$ in terms of the revenues of the marginal firm exporting to the same destination:

$$r_d(a) = \begin{cases} r_d^b(\tilde{a}_d^b)\left(\frac{a}{\tilde{a}_d^b}\right)^{\sigma-1}\left(\frac{t_d^c}{t_d^b}\right)^{1-\sigma} & \text{if } a \geq \tilde{a}_d^c, \\ r_d^b(\tilde{a}_d^b)\left(\frac{a}{\tilde{a}_d^b}\right)^{\sigma-1} & \text{if } a < \tilde{a}_d^c. \end{cases} \tag{18}$$

Finally, container share for the model with iceberg-type trade costs becomes

$$\mu_d = \frac{(\Delta t_d)^{1-\sigma}\left(\frac{\Delta f_d-1}{(\Delta t_d)^{1-\sigma}-1}\right)^{\frac{\sigma-k-1}{\sigma-1}}}{1+\left(\frac{\Delta f_d-1}{(\Delta t_d)^{1-\sigma}-1}\right)^{\frac{\sigma-k-1}{\sigma-1}}\cdot\left[(\Delta t_d)^{1-\sigma}-1\right]}. \tag{19}$$

Note that letting transport costs to be multiplicative does not modify our estimating equation. Using the same specification for $t_d^m = \bar{t}_m \cdot dist_d^{\delta_m}$, revenue expressions (17)-(18) still imply the estimating equation (13) without the term containing $\theta$. In other words, the model with additive trade costs and quality differentiation delivers a reduced form equation that is isomorphic to the one derived from a model with multiplicative trade costs and no quality differentiation. This interesting result can be traced to the optimal pricing expression (6) derived under additive trade costs being multiplicative itself due to the presence of the quality margin.

Using the baseline estimates of $(\eta_1, \eta_2)$ from the fourth column of table 6, along with $\sigma = 4$, we back out $(\bar{t}_c/\bar{t}_b) = \exp(\eta_1)/(1-\sigma) = 1.19$ and $(\delta_c - \delta_b) = \eta_2/(1-\sigma) = -0.038$. Using the implied $\Delta t_d$, we then back out relative fixed costs $\Delta f_d$ from the empirical aggregate container shares $\mu_d$ using the expression above. Top and bottom panels of figure 3 plot relative variable trade costs $\Delta t_d$ and histogram of relative fixed trade costs $\Delta f_d$. Note that the breakeven distance remains unchanged from figure 2 since it only depends on our estimates of $\eta_1$ and $\eta_2$.[13] While relative variable costs look qualitatively and quantitatively similar to its counterpart at the top panel of figure 2, iceberg specification implies smaller variable cost savings from containerization. This is evident from the flatter cost gradient with respect to distance at top panels. The implied variable cost saving for a shipping between China and USA is 16 percent when transport costs are deemed to be multiplicative, compared to 22 percent when they are assumed to be additive. When the elasticity of substitution is set higher to $\sigma = 6$, the implied cost savings are much more similar between multiplicative and additive specifications (dashed lines at the top panels of both figures).

---

[13]This follows from $t_d^b(\widetilde{dist}) = t_d^c(\widetilde{dist})$, which yields $\widetilde{dist} = \exp\left(\frac{\ln(\bar{t}_c)-\ln(\bar{t}_b)}{\delta_b-\delta_c}\right) = \exp\left(-\hat{\eta}_1/\hat{\eta}_2\right).$

# 6 Additional Evidence from the U.S. Data

In this section, we use publicly available U.S. data to provide additional quantitative and qualitative support for our relative variable transport cost ($\Delta t_d$) estimates. The data, originally collected by the U.S. Census Bureau and made publicly available by Schott (2008), provides information about the value and weight of U.S. imports, as well as non-tariff charges (freight and insurance) on them, broken down by mode of shipment, source country, 10-digit product, and the customs district of entry. The mode of shipment is further broken down into containerized and non-containerized vessel imports. We use data for 2013 and aggregate product categories to 8-digit to facilitate comparison with our baseline estimates from the Turkish data. For the same reason, we exclude HS headings 27 and 93.

We start by checking the distribution of containerized import charges relative to breakbulk. $CIF_{js}^m$ denotes charges for insurance and freight per import weight for each product $j$, from each source country $s$ and by shipping mode $m \in \{container, breakbulk\}$. Figure 4 plots the histogram of $CIF_{js}^c / CIF_{js}^b$ across product-source country ($js$) pairs. With a median value of 0.804, import charges of containerized flows are lower for majority of import flows. This is direct evidence that for most destinations, breakbulk is the more expensive shipping mode in terms of variable costs. If, in addition, shipping costs have an additive component, the result is also consistent with the negative correlation between container usage and unit values (Fact 3) and justifies our modeling of endogenous quality differentiation to account for that.

Using the U.S. import data, we now directly estimate the transport cost function (11):

$$\ln(CIF_{js}^m) = \alpha_j + \alpha_s + \beta_1 \cdot CONT_{jsm} + \beta_2 \cdot CONT_{jsm} \cdot \ln(dist_s) + \epsilon_{jsm},$$

where $\alpha$'s represent the set of product and source country-port of entry fixed effects. Qualitatively, $\beta_1 > 0$ and $\beta_2 < 0$ would be consistent with our finding that containerization has a higher first-mile cost and a lower distance elasticity. Quantitatively, the magnitude of $\beta_2$ provides a direct estimate for $\delta_c - \delta_b$.

Table 8 presents the results. In column 1, we report the baseline estimate and in column 2, we include unit values in order to capture the effect of import charges varying with prices such as insurance. Not only do the coefficients have the anticipated signs, but the magnitude of $\hat{\beta}_2 = \delta_c - \delta_b$ in column 2 is consistent with the value of 0.05 estimated from the Turkish data.

# 7 Counterfactuals: Effect of Containerization on Trade

Having quantified the relative cost of maritime transport modes, we now explore the extent to which the availability of container shipping increases trade. To address this question, we calculate two statistics: we first compare the current level of exports to a destination to the counterfactual level that would obtain if

container shipping was not available:

$$\Delta EXP_1^d = \frac{\underbrace{\int_{\tilde{a}_d^b}^{\tilde{a}_d^c} r_d(a)dG(a) + \int_{\tilde{a}_d^c}^{\infty} r_d(a)dG(a)}_{\text{Current exports}}}{\underbrace{\int_{\tilde{a}_d^b}^{\infty} r_d(a)dG(a)}_{\text{Breakbulk only counterfactual exports}}}.$$

The second method assumes that the selection equation in (10) holds with equality, i.e., $\Delta f_d = \Delta t_d^{-(\sigma-1)(1-\theta)}$ for all $d$, due to a counterfactual decrease in the fixed cost of container shipping. In this case, all exporters to all destinations prefer containerization to breakbulk shipping. This statistic is given by

$$\Delta EXP_2^d = \frac{\underbrace{\int_{\hat{a}_d^c}^{\infty} r_d(a)dG(a)}_{\text{Container only counterfactual exports}}}{\underbrace{\int_{\tilde{a}_d^b}^{\tilde{a}_d^c} r_d(a)dG(a) + \int_{\tilde{a}_d^c}^{\infty} r_d(a)dG(a)}_{\text{Current exports}}},$$

where $\hat{a}_d^c$ denotes the ability of the marginal exporter to destination $d$ when $\Delta f_d = \Delta t_d^{-(\sigma-1)(1-\theta)}$. While $\Delta EXP_1^d$ is informative about how much international trade has increased from the pre-container era to the present due to the availability of the technology, $\Delta EXP_2^d$ is informative about potential future increases due to further improvements in the technology to the point of full adaption. In Appendix A3, we derive these statistics analytically:

$$\Delta EXP_1^d = 1 + \left(\Delta t_d^{-(\sigma-1)(1-\theta)} - 1\right) \left(\frac{\Delta f_d - 1}{\Delta t_d^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{\theta(\sigma-1)-k}{\theta(\sigma-1)}},$$

$$\Delta EXP_2^d = \frac{\Delta t_d^{-(\sigma-1)(1-\theta)}}{1 + \left(\Delta t_d^{-(\sigma-1)(1-\theta)} - 1\right) \left(\frac{\Delta f_d - 1}{\Delta t_d^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{\theta(\sigma-1)-k)}{\theta(\sigma-1)}}}.$$

We conduct these counterfactual exercises for both Turkey and the U.S. For the latter, we calculate $\Delta t_d$ using our estimates and the sea distances of export destinations to the U.S., and then back out $\Delta f_d$ from equation (14) using the median U.S. container share across sectors for each destination, $\mu_d$.

Figure 5 plots destination specific $\Delta EXP_1^d$ against sea distances. Since we recover variable trade costs from a specification with sector fixed effects and use the sectoral median container share $\mu_d$ to calibrate fixed costs, the interpretation of the counterfactual export change is also with respect to the median sector in terms of container usage. For both countries, containerization implies a significant increase in the exports of

the median sector: the average across 96 destinations in Figure 5 is 46 percent for Turkey, and 40 percent for the U.S. In other words, current trade levels would decrease by about a third in an average sector if container technology did not exist ($\approx 0.46/1.46$). The gains reach 78 percent for the most remote trade partners.[14]

Potential increases in trade if container shipping was fully adopted—to the point where it is preferred by all exporters—are relatively smaller: the mean $\Delta EXP_2^d$ across destinations is 13 percent and 19 percent for Turkey and the U.S., respectively. The implication is that for the typical sector, most of the trade increasing effect of containerization has already been realized. This is not surprising given that while only 54 percent of Turkish exporters use containerization in a typical sector, remaining firms export smaller amounts. Therefore, full adoption would increase Turkey's exports in that sector to the average destination by just 13 percent due to the diminishing returns caused by firm selection.

To gauge the total effect on maritime exports, we aggregate $\Delta EXP_1^d$ using as weights the export share of each destination country $d$. Since trade shares are decreasing in distance, the aggregate effect is smaller than the simple average reported above. We find that total Turkish and American *maritime* exports would have been 14 and 21 percent lower in the absence of containerization, respectively.

These results are invariant to transport costs being additive or multiplicative since $\Delta EXP_1^d$ and $\Delta EXP_2^d$ are functions of the observables $(\mu_d, dist_d)$ and the estimates of $(\eta_1, \eta_2)$—see Appendix A3 for details. To see the intuition, first note that the term $\Delta t_d^{-(\sigma-1)(1-\theta)}$ in $\Delta EXP_1^d$ and $\Delta EXP_2^d$ depends only on $dist_d$ and the estimates of $(\eta_1, \eta_2)$. This is a direct result of both specifications yielding the same estimating equation. Second, we back out $\Delta f_d$ from the same empirical container share moment $\mu_d$—depending on the cost specification, this is equation (14) or (19). As a result, the model with additive costs and endogenous quality choice yields the same trade increase due to containerization as the model with multiplicative trade costs and no quality margin, when both models are quantified using the same salient moments of the data. For the very same reason, the values chosen for $(\sigma, k)$ parameters do not matter for the magnitudes of $\Delta EXP_1^d$ and $\Delta EXP_2^d$.

The two specifications, however, differ in the implied elasticity of trade to shipping costs. In the baseline (specific transport costs), containerization decreases variable trade costs to a destination located at the median distance from Turkey or U.S. by about 20 percent. As mentioned above, exports to the average destination would be around 46 percent lower in a typical sector for both countries. This implies trade elasticity around 2.3 ($\approx 0.46/|-0.2|$). Since the estimated variable cost decrease is smaller under iceberg specification—see top panels of figures 2-3 and the discussion in the previous section—implied trade

---

[14]By construction, the exercise implies an increase of exports to all destinations. Such drastic changes in distance-related shipping costs, however, could potentially divert trade from nearby trade partners in a general equilibrium framework. We also abstract from endogenous price setting by carriers, which is empirically relevant in the transportation industry (Hummels, Lugovskyy, and Skiba, 2009). See Asturias (2018) and Wong (2018) for welfare analysis in general equilibrium when transportation costs are endogenous.

elasticities are larger, around 3 ($\approx 0.46/|-0.15|$).

# 8 Conclusion

This paper takes an important step toward better understanding and quantifying transport costs. Using micro-level Turkish export data, we estimate the effect of containerization on transport costs and trade by providing the first systematic estimation of firms' modal choice between containerized and breakbulk shipping. The results confirm the role of container technology in the global economy: it implies variable cost savings around 20 percent at distances that are relevant for major trading economies, with a fixed cost around 50 to 100 percent higher than its breakbulk alternative. A counterfactual exercise suggests that in the absence of containerization, Turkish and U.S. maritime exports in a typical sector to an average destination would decrease by around a third. Aggregating over destinations, we predict that total maritime exports would have been by about 14 and 21 percent lower for Turkey and the U.S., respectively. Container shipping is indeed a major driver of increased international trade in the past several decades.

We conclude by noting that our findings relate to the "distance puzzle" in international trade, i.e., the negative effect of distance on trade has not decreased over time (Disdier and Head, 2008). In describing the potential sources of the supposed death of distance, the literature often refers to technological advancements in transportation. Our result on container shipping costs being less distance-elastic provides a mechanism and support for this idea. We do not, however, offer a resolution to the puzzle itself. Neither do we interpret our result as predicting a decline in the distance elasticity of overall trade due to an increasing intensity of container usage over time. The distance elasticity of trade depends not only on the distance elasticity of trade costs but also on the trade elasticity with respect to trade costs. It is possible that changes in the sectoral or bilateral composition of world trade may have offset the decline in distance-related shipping costs. For instance, containerization may have helped countries that are located further apart start trading with each other. If trade between new partners is subject to higher non-transport trade frictions than those between existing trade partners, and if that discrepancy is correlated with distance (e.g., due to lingual or cultural differences), the compositional shift may offset the gains from lower shipping costs and keep the average distance elasticity constant (Head and Mayer, 2013). Investigating the channels contributing to the distance puzzle despite declining shipping costs thus remains an important question for researchers.

# References

ANDERSON, J. E., AND E. VAN WINCOOP (2004): "Trade costs," *Journal of Economic literature*, 42(3), 691–751.

ASTURIAS, J. (2018): "Endogenous Transportation Costs," Discussion paper, Georgetown University Qatar.

BALDWIN, R., AND J. HARRIGAN (2011): "Zeros, Quality, and Space: Trade Theory and Trade Evidence," *American Economic Journal: Microeconomics*, 3(2), 60–88.

BERNARD, A. B., S. J. REDDING, AND P. K. SCHOTT (2011): "Multiproduct Firms and Trade Liberalization," *The Quarterly Journal of Economics*, 126(3), 1271–1318.

BERNHOFEN, D., Z. EL-SAHLI, AND R. KNELLER (2016): "Estimating The Effects of the Container Revolution on World Trade," *Journal of International Economics*, 98, 36–50.

BLONIGEN, B. A., AND W. W. WILSON (2008): "Port efficiency and trade flows," *Review of International Economics*, 16(1), 21–36.

BRANCACCIO, G., M. KALOUPTSIDI, AND T. PAPAGEORGIOU (2017): "Geography, Search Frictions and Endogenous Trade Costs," Working Paper 23581, National Bureau of Economic Research.

BUSTOS, P. (2011): "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, 101(1), 304–40.

CHANEY, T. (2008): "Distorted gravity: the intensive and extensive margins of international trade," *The American Economic Review*, 98(4), 1707–1721.

CLARK, X., D. DOLLAR, AND A. MICCO (2004): "Port Efficiency, Maritime Transport Costs, and Bilateral Trade," *Journal of Development Economics*, 75(2), 417 – 450.

COŞAR, A. K., AND B. DEMIR (2016): "Domestic road infrastructure and international trade: Evidence from Turkey," *Journal of Development Economics*, 118, 232–244.

DISDIER, A.-C., AND K. HEAD (2008): "The puzzling persistence of the distance effect on bilateral trade," *The Review of Economics and Statistics*, 90(1), 37–48.

DONALDSON, D. (2018): "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 108(4-5), 899–934.

ECKEL, C., L. IACOVONE, B. JAVORCIK, AND J. P. NEARY (2015): "Multi-product firms at home and away: Cost- versus quality-based competence," *Journal of International Economics*, 95(2), 216 – 232.

ECKEL, C., AND J. P. NEARY (2010): "Multi-product Firms and Flexible Manufacturing in the Global Economy," *The Review of Economic Studies*, 77(1), 188–217.

FEENSTRA, R. C., AND J. ROMALIS (2014): "International Prices and Endogenous Quality," *The Quarterly Journal of Economics*, 129(2), 477–527.

GELBACH, J. B. (2016): "When Do Covariates Matter? And Which Ones, and How Much?," *Journal of Labor Economics*, 34(2), 509–543.

HARRIGAN, J. (2010): "Airplanes and Comparative Advantage," *Journal of International Economics*, 82(2), 181–194.

HEAD, K., AND T. MAYER (2013): "What separates us? Sources of resistance to globalization," *Canadian Journal of Economics/Revue canadienne d'Economique*, 46(4), 1196–1231.

HUMMELS, D. (2007): "Transportation Costs and International Trade in the Second Era of Globalization," *Journal of Economic Perspectives*, 21(3), 131–154.

HUMMELS, D., V. LUGOVSKYY, AND A. SKIBA (2009): "The Trade Reducing Effects of Market Power in International Shipping," *Journal of Development Economics*, 89(1), 84 – 97.

HUMMELS, D., AND G. SCHAUR (2013): "Time as a Trade Barrier," *American Economic Review*, 103(7), 2935–59.

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

IRARRAZABAL, A., A. MOXNES, AND L. D. OPROMOLLA (2015): "The Tip of the Iceberg: A Quantitative Framework for Estimating Trade Costs," *The Review of Economics and Statistics*, 97(4), 777–792.

KORINEK, J. (2011): "Clarifying Trade Costs in Maritime Transport," Trade policy paper, OECD Publishing.

KUGLER, M., AND E. VERHOOGEN (2012): "Prices, Plant Size, and Product Quality," *Review of Economic Studies*, 79(1), 307–339.

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

MELITZ, M., AND S. REDDING (2015): "New Trade Models, New Welfare Implications," *American Economic Review*, 105(3), 1105–46.

MELITZ, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

MICCO, A., AND T. SEREBRISKY (2006): "Competition regimes and air transport costs: The effects of open skies agreements," *Journal of International Economics*, 70(1), 25–51.

PASCALI, L. (2017): "The Wind of Change: Maritime Technology, Trade and Economic Development," *American Economic Review*, 107(9).

RAUCH, J. E. (1999): "Networks versus Markets in International Trade," *Journal of International Economics*, 48(1), 7–35.

RUA, G. (2014): "Diffusion of Containerization," Finance and Economics Discussion Series 2014-88, Board of Governors of the Federal Reserve System (U.S.).

SCHOTT, P. K. (2008): "The relative sophistication of Chinese exports," *Economic Policy*, 23, 5–49.

STOPFORD, M. (2009): *Maritime Economics*. Routledge, New York, NY, 3 edn.

WONG, W. F. (2018): "The Round Trip Effect: Endogenous Transport Costs and International Trade," Discussion paper, University of Oregon.

# APPENDIX

## A1    Derivation of the Revenue Function (equation 12)

Using the expression for revenues in equation (7), we can write revenue of a firm exporting to destination $d$ using transport technology $m$ as a function of the marginal firm exporting to the same destination using transport technology $m$:

$$\frac{r_d^m(a)}{r_d^m(\tilde{a}_d^m)} = \left(\frac{a}{\tilde{a}_d^m}\right)^{\theta(\sigma-1)}, \tag{20}$$

where $\tilde{a}_d^m$ denotes the productivity of the marginal firm using transport technology $m$. The revenue of the marginal exporter can be written as:

$$r_d^m(\tilde{a}_d^m) = \Theta_d \cdot (\tilde{a}_d^m)^{\theta(\sigma-1)} \cdot (t_d^m)^{-(\sigma-1)(1-\theta)}, \tag{21}$$

where $\Theta_d$ is a function of, among other, destination country $X_d$ (see equation 7). For the marginal firm choosing containerization over breakbulk, we have the following relationship between its revenue from containerized exports and its revenue from breakbulk exports:

$$\frac{r_d^c(\tilde{a}_d^c)}{r_d^b(\tilde{a}_d^c)} = \left(\frac{t_d^c}{t_d^b}\right)^{-(\sigma-1)(1-\theta)}. \tag{22}$$

Putting the two expressions together, we can write firm revenues as follows:

$$r_d(a) = \begin{cases} r_d^b(\tilde{a}_d^b) \left(\frac{a}{\tilde{a}_d^b}\right)^{\theta(\sigma-1)} \left(\frac{t_d^c}{t_d^b}\right)^{-(\sigma-1)(1-\theta)} & \text{if } a \geq \tilde{a}_d^c, \\ r_d^b(\tilde{a}_d^b) \left(\frac{a}{\tilde{a}_d^b}\right)^{\theta(\sigma-1)} & \text{if } a < \tilde{a}_d^c. \end{cases}$$

Using the specific trade cost function (11) and the indicator function $CONT$ denoting container usage, i.e., $CONT = 1$ for $a \geq \tilde{a}_d^c$, and taking logarithms yields the expression (12) in the text.

## A2    Derivation of the Container Share (equation 14)

Firm productivity is distributed Pareto with the CDF, $k$ is shape parameter:

$$G(a) = 1 - a^{-k}, \tag{23}$$

Share of containerized exports for a given sector-destination pair is given by:

$$\mu_d = \frac{\int_{\tilde{a}_d^c}^{\infty} r_d(a) dG(a)}{\int_{\tilde{a}_d^b}^{\infty} r_d(a) dG(a)}.$$

We can write revenues in terms of revenues of the marginal exporter using equations (20) and (22) to obtain:

$$
\mu_d = \frac{\int_{\tilde{a}_d^c}^{\infty} r_d(\tilde{a}_d^b)(\Delta t_d)^{-(\sigma-1)(1-\theta)}(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a)}{\int_{\tilde{a}_d^b}^{\tilde{a}_d^c} r_d(\tilde{a}_d^b)(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a) + \int_{\tilde{a}_d^c}^{\infty} r_d(\tilde{a}_d^b)(\Delta t_d)^{-(\sigma-1)(1-\theta)}(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a)}
$$

$$
= \frac{(\Delta t_d)^{-(\sigma-1)(1-\theta)}(\tilde{a}_d^c)^{\theta(\sigma-1)-k}}{(\tilde{a}_d^b)^{\theta(\sigma-1)-k} + [(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1](\tilde{a}_d^c)^{\theta(\sigma-1)-k}},
$$

where $\Delta t_d = t_d^c/t_d^b$.

Next, we use the profit functions in (8) and (9) to express marginal productivities in terms of trade costs:

$$
(\tilde{a}_d^b)^{\theta(\sigma-1)} = \frac{\chi\sigma}{(\chi+1)\Theta_d} \cdot (t_d^b)^{(\sigma-1)(1-\theta)} \cdot f_d^b
$$

$$
(\tilde{a}_d^c)^{\theta(\sigma-1)} = \frac{\chi\sigma}{(\chi+1)\Theta_d} \cdot \frac{1}{(t_d^c)^{-(\sigma-1)(1-\theta)} - (t_d^b)^{-(\sigma-1)(1-\theta)}} \cdot (f_d^c - f_d^b).
$$

Substituting these in the expression for $\mu_d$ above and simplifying yield the following:

$$
\mu_d = \frac{(\Delta t_d)^{-(\sigma-1)(1-\theta)}\left(\frac{\Delta f_d-1}{(\Delta t_d)^{-(\sigma-1)(1-\theta)}-1}\right)^{\frac{\theta(\sigma-1)-k}{\theta(\sigma-1)}}}{1+\left(\frac{\Delta f_d-1}{(\Delta t_d)^{-(\sigma-1)(1-\theta)}-1}\right)^{\frac{\theta(\sigma-1)-k}{\theta(\sigma-1)}} \cdot \left[(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1\right]},
$$

where $\Delta f_d = f_d^c/f_d^b$. Derivation of $\mu_d^{ext}$ follows similar steps.

## A3    Derivations for the Counterfactual Exercise

The first measure, $\Delta EXP_1$, compares the current level of exports to a destination to the counterfactual level that would obtain if container shipping was not available:

$$
\Delta EXP_1 = \frac{\int_{\tilde{a}_d^b}^{\tilde{a}_d^c} r_d(\tilde{a}_d^b)(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a) + \int_{\tilde{a}_d^c}^{\infty} r_d(\tilde{a}_d^b)(\Delta t_d)^{-(\sigma-1)(1-\theta)}(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a)}{\int_{\tilde{a}_d^b}^{\infty} r_d(\tilde{a}_d^b)(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a)}.
$$

The numerator was derived in Appendix A2, and the denominator is:

$$
r_d(\tilde{a}_d^b)(\tilde{a}_d^b)^{-\theta(\sigma-1)}(\tilde{a}_d^b)^{\theta(\sigma-1)-k}.
$$

So, using the expressions for $\tilde{a}_d^b$ and $\tilde{a}_d^c$ in Appendix A2, we obtain:

$$
\Delta EXP_1 = 1 + \left(\Delta t_d^{-(\sigma-1)(1-\theta)} - 1\right)\left(\frac{\Delta f_d - 1}{\Delta t_d^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{\theta(\sigma-1)-k}{\theta(\sigma-1)}}. \tag{24}
$$

The other measure, $\Delta EXP_2$, compares the counterfactual level that would obtain if the relative fixed cost container shipping would fall such that $\Delta f_d = \Delta t_d^{-(\sigma-1)(1-\theta)}$ to the current level of exports to a destination. Let $\widehat{a}_d^c$ denote the ability of the marginal exporter to destination $d$ when

$\Delta f_d = \Delta t_d^{-(\sigma-1)(1-\theta)}$. Then, $\Delta EXP_2$ is equal to:

$$\Delta EXP_2 \quad = \quad \frac{\int_{\hat{a}_d^c}^{\infty} r_d(\hat{a}_d^c)(a/\hat{a}_d^c)^{\theta(\sigma-1)}dG(a)}{\int_{\tilde{a}_d^b}^{\tilde{a}_d^c} r_d(\tilde{a}_d^b)(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a) + \int_{\tilde{a}_d^c}^{\infty} r_d(\tilde{a}_d^b)(\Delta t_d)^{-(\sigma-1)(1-\theta)}(a/\tilde{a}_d^b)^{\theta(\sigma-1)}dG(a)}.$$

The denominator was derived in Appendix A2. The numerator is equal to $r_d(\hat{a}_d^c)(\hat{a}_d^c)^{-k}$, where $r_d(\hat{a}_d^c) = \frac{\chi\sigma}{\chi+1}\Delta t_d^{-(\sigma-1)(1-\theta)}f_d^b$. Rearranging terms, we obtain:

$$\Delta EXP_2^d \quad = \quad \frac{\Delta t_d^{-(\sigma-1)(1-\theta)}}{1 + \left(\Delta t_d^{-(\sigma-1)(1-\theta)} - 1\right)\left(\frac{\Delta f_d - 1}{\Delta t_d^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{\theta(\sigma-1)-k)}{\theta(\sigma-1)}}}. \tag{25}$$

To see that these expressions do not depend on the assumed values of $(\sigma, \theta, k)$ or relative fixed costs, $\Delta f_d$, define $W_d = \left(\frac{\Delta f_d - 1}{(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1}\right)^{\frac{\theta(\sigma-1)-k}{\theta(\sigma-1)}}$ to re-write the share of containerized exports to a given destination as:

$$\mu_d \quad = \quad \frac{(\Delta t_d)^{-(\sigma-1)(1-\theta)}W_d}{1 + W_d \cdot \left[(\Delta t_d)^{-(\sigma-1)(1-\theta)} - 1\right]},$$

where $(\Delta t_d)^{-(\sigma-1)(1-\theta)} = \exp(\eta_1)dist_d^{\eta_2}$. We can express $W_d$ as a function of observables $(\mu_d, dist_d)$ and estimates of $(\eta_1, \eta_2)$:

$$W_d = \frac{\mu_d}{\exp(\eta_1)dist_d^{\eta_2} - \mu_d\left(\exp(\eta_1)dist_d^{\eta_2} - 1\right)}.$$

Next, re-write $\Delta EXP_1$ and $\Delta EXP_2$ as functions of $W_d$:

$$\Delta EXP_1 \quad = \quad 1 + \left(\exp(\eta_1)dist_d^{\eta_2} - 1\right)W_d,$$
$$\Delta EXP_2 \quad = \quad \frac{\exp(\eta_1)dist_d^{\eta_2}}{1 + \left(\exp(\eta_1)dist_d^{\eta_2} - 1\right)W_d}. \tag{26}$$

The expressions in (24) and (25) for the model with iceberg trade costs become:

$$\Delta EXP_1 \quad = \quad 1 + \left(\Delta t_d^{-(\sigma-1)} - 1\right)\left(\frac{\Delta f_d - 1}{\Delta t_d^{-(\sigma-1)} - 1}\right)^{\frac{\sigma-1-k}{\sigma-1}},$$
$$\Delta EXP_2 \quad = \quad \frac{\Delta t_d^{-(\sigma-1)}}{1 + \left(\Delta t_d^{-(\sigma-1)} - 1\right)\left(\frac{\Delta f_d - 1}{\Delta t_d^{-(\sigma-1)} - 1}\right)^{\frac{\sigma-1-k}{\sigma-1}}}.$$

It is easy to see that the expressions above can be re-written as in (26) using the equation for the share of containerized exports in (19).

## Table 1: **Summary Statistics for Turkish Export Data**

| | | **Panel A: All products** | | |
|---|---|---|---|---|
| | | Level of observation | | |
| | Firm | Product | Destination | Firm-product-destination |
| Fraction of zeros (no containerization) | 0.335 | 0.107 | 0.000 | 0.360 |
| Fraction of ones (full containerization) | 0.337 | 0.143 | 0.029 | 0.533 |
| Share of containerized exports | 0.524 | 0.556 | 0.598 | 0.527 |
| Share of containerized exports (excl. zeros & ones) | 0.571 | 0.551 | 0.586 | 0.568 |
| | | Number of | | |
| | Firms | Products | Destinations | Observations |
| Vessel exports | 27,241 | 5,557 | 139 | 220,993 |
| Non-containerized | 18,070 | 4,762 | 135 | 103,259 |
| Containerized | 18,112 | 4,961 | 139 | 141,454 |
| | Mean | | Median | |
| | Value (USD) | Quantity (Kg) | Value (USD) | Quantity (Kg) |
| Non-containerized | 378,407.7 | 378,274.7 | 28,673 | 4,201 |
| | (5,046,657) | (5,945,869) | | |
| Containerized | 241,333.4 | 183,440.4 | 29,253 | 6,789 |
| | (1,943,938) | (2,997,263) | | |
| | | **Panel B: Containerizable products only** | | |
| | | Level of observation | | |
| | Firm | Product | Destination | Firm-product-destination |
| Fraction of zeros (no containerization) | 0.337 | 0.100 | 0.000 | 0.361 |
| Fraction of ones (full containerization) | 0.338 | 0.134 | 0.050 | 0.533 |
| Share of containerized exports | 0.524 | 0.558 | 0.737 | 0.594 |
| Share of containerized exports (excl. zeros & ones) | 0.573 | 0.553 | 0.723 | 0.572 |
| | | Number of | | |
| | Firms | Products | Destinations | Observations |
| Containerizable vessel exports | 25,846 | 4,865 | 139 | 206,659 |
| Non-containerized | 17,112 | 4,212 | 132 | 96,602 |
| Containerized | 17,136 | 4,377 | 139 | 132,067 |
| | Mean | | Median | |
| | Value (USD) | Quantity (Kg) | Value (USD) | Quantity (Kg) |
| Non-containerized | 194,114.1 | 71,934.9 | 26,893 | 3,595 |
| | (1,232,342) | (808,124.3) | | |
| Containerized | 217,471.8 | 92,438.7 | 27,881 | 5,849 |
| | ( 1,635,886) | ( 785,887.8) | | |

*Notes:* This table presents the summary statistics for the annualized data. Panel A is based on all products, regardless of their containerizability. Panel B restricts the sample to containerizable products based on the classification suggested by Korinek (2011). The following products are excluded from the sample: dry and industrial bulk, tankers, and bulk shipped goods, goods shipped in tankers, motor vehicles, trailers and non-mechanically propelled vehicle n.e.s., aircraft, spacecraft, satellites, ships, boats and other floating structures. Standard errors are in parentheses.

Table 2: **Explaining the Variation in Containerization (Fact 1)**

|  | Direct effect | Coefficient of partial determination |
|---|---|---|
| **Panel A: Turkish exports** | | |
| Firm ($a$) | 0.557 | 0.509 |
| Product ($j$) | 0.180 | 0.066 |
| Destination ($d$) | 0.240 | 0.205 |
| Firm-product ($aj$) | 0.620 | 0.603 |
| Firm-destination ($ad$) | 0.909 | 0.896 |
| Product-destination ($jd$) | 0.442 | 0.381 |
| **Panel B: U.S. exports** | | |
| Port ($p$) | 0.020 | 0.015 |
| Product ($j$) | 0.054 | 0.049 |
| Destination ($d$) | 0.019 | 0.014 |
| Port-product ($pj$) | 0.133 | 0.128 |
| Port-destination ($pd$) | 0.076 | 0.071 |
| Product-destination ($jd$) | 0.150 | 0.146 |

*Notes:* First column reports adjusted $R^2$'s from regressing container shares $ContShr_{ajd}$ on individual (top panel) or paired (bottom panel) fixed effects. As an example, for firms, this would be the $R_a^2$ from the regression $ContShr_{ajd} = \mu_a$ where $\mu_a$ represents firm fixed effects. Second column reports the coefficient of partial determination capturing the proportion of variation that cannot be explained in a reduced model without the particular element among the fixed effects. As an example for firms ($k = a$), take the $R_{a,jd}^2$ from the regression $ContShr_{ajd} = \mu_a + \mu_{jd} + \epsilon_{ajd}$, where $\mu_{jd}$ represents product-destination pair fixed effects. Dropping the fixed effect for the dimension of interest $a$, the regression $ContShr_{ajd} = \mu_{jd} + \epsilon_{ajd}$ yields $R_{jd}^2$. The coefficient of partial determination for firms is then $(R_{a,jd}^2 - R_{jd}^2)/(1 - R_{jd}^2)$, that is the fraction of the unexplained variation captured when firm fixed effects are included. Destination refers to a destination-month pair.

Table 3: **Distance and Containerization (Fact 2)**

|  | (1) $ContShare_d$ | (2) $ContShare_d$ | (3) $ContShare_d$ | (4) $ContShare_d$ |
|---|---|---|---|---|
| $\ln dist_d$ | 0.088*** | 0.118*** | 0.088*** | 0.117*** |
|  | (0.023) | (0.030) | (0.024) | (0.030) |
| $\ln GDPpc_d$ | -0.016 | 0.010 | -0.016 | 0.010 |
|  | (0.022) | (0.010) | (0.022) | (0.014) |
| $MajorFTA_d$ | -0.161* | -0.030 | -0.157*** | -0.030 |
|  | (0.057) | (0.040) | (0.054) | (0.041) |
| Exporter | Turkey | U.S. | Turkey | U.S. |
| Method | OLS | OLS | Fractional logit | Fractional logit |
| $R^2$ | 0.217 | 0.175 | - | - |
| Observations | 103 | 103 | 103 | 103 |

*Notes:* The dependent variable is the share of containerized maritime exports in total maritime exports in 2013 for Turkey (column 1) and the U.S. (column 2). $\ln dist_d$ is the shortest sea distance to the destination country. $\ln GDPpc_d$ is (log) per capita GDP of the destination country. $MajorFTA_d$ is a dummy taking the value one for major free trade agreements, which is the EU for Turkey and NAFTA countries for the U.S. Fractional logit coefficients are average marginal effects. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 4: **Economies of Scale and Unit Values in Containerization (Fact 3)**

|  | All | Differentiated | Non-differentiated |
|---|---|---|---|
|  | (1) $CONT_{ajdm}$ | (2) $CONT_{ajdm}$ | (3) $CONT_{ajdm}$ |
| $\ln weight_{ajdm}$ | 0.0099*** | 0.0116*** | 0.0057*** |
|  | (0.0008) | (0.0009) | (0.0016) |
| $\ln UnitValue_{ajdm}$ | -0.0070*** | -0.0077*** | -0.0030 |
|  | (0.0017) | (0.0018) | (0.0040) |
| Observations | 711742 | 532277 | 179465 |
| $R^2$ | 0.690 | 0.697 | 0.682 |
| Firm-product-month FE | + | + | + |
| Destination-month FE | + | + | + |

*Notes:* The dependent variable $CONT_{ajdm}$ is a binary variable that takes the value one if there is a positive containerized flow at the firm-product-destination level in a given month. Column headings refer to the sample used to produce them. Product differentiation is based on the classification developed by Rauch (1999). $\ln weight_{ajdm}$ denotes the logarithm of the weight (kg) of the export flows. $\ln UnitValue_{ajdm}$ denotes the logarithm of the unit value, defined as value per quantity. Robust standard errors clustered at the product-destination-level in parentheses. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 5: **Firm Characteristics and Containerization (Fact 4)**

| | (1) $CONT_{ajdm}$ | (2) $CONT_{ajdm}$ | (3) $CONT_{ajdm}$ | (4) $CONT_{ajdm}$ | (5) $CONT_{ajdm}$ | (6) $CONT_{ajdm}$ |
|---|---|---|---|---|---|---|
| $\ln weight_{ajdm}$ | 0.0144*** (0.00189) | 0.0172*** (0.00287) | 0.0162*** (0.00284) | 0.0171*** (0.00293) | 0.0149*** (0.00284) | 0.0145*** (0.00280) |
| $\ln exports_a$ | 0.0169*** (0.00197) | | | | 0.0169*** (0.00329) | 0.0179*** (0.00355) |
| $\ln employment_a$ | | 0.0104** (0.00473) | | | | |
| $\ln sales_a$ | | | 0.0191*** (0.00371) | | | |
| $\ln(sales\ per\ worker)_a$ | | | | 0.0136*** (0.00461) | 0.00158 (0.00536) | 0.00288 (0.00605) |
| $NumProducts_{ad}$ | | | | | | -0.000133 (0.000237) |
| Observations | 711743 | 437208 | 437208 | 437208 | 437208 | 437208 |
| $R^2$ | 0.673 | 0.725 | 0.726 | 0.725 | 0.726 | 0.726 |
| Product-destination-month FE | + | + | + | + | + | + |

*Notes:* The dependent variable $CONT_{ajdm}$ is a binary variable that takes the value one if there is a positive containerized flow at the firm-product-destination level in a given month. $\ln weight_{ajdm}$ denotes the logarithm of the weight (kg) of the export flows. $\ln exports_a$ is the logarithm of the total value of exports of firm $i$, $\ln employment_a$ the logarithm of the average number of paid employees, $\ln sales_a$ the logarithm of total sales, and $\ln(sales\ per\ worker)_a$ the logarithm of total sales divided by the number of workers. Robust standard errors in parentheses are clustered at the firm-level. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 6: **Main Estimation: Inferring Containerization's Effects on Transport Costs from its Effect on Export Revenues**

| | Coefficient | (1) $\ln(r_{ajdm}/\overline{r}_j)$ | (2) $\ln(r_{ajdm}/\overline{r}_j)$ | (3) $\ln(r_{ajdm}/\overline{r}_j)$ | (4) $\ln(r_{ajdm}/\overline{r}_j)$ |
|---|---|---|---|---|---|
| $CONT_{ajdm}$ | $\eta_1$ | 0.296*** | -0.289*** | -0.275** | -0.529** |
| | | (0.0125) | (0.101) | (0.125) | (0.250) |
| | | | | | |
| $CONT_{ajdm} \cdot \ln dist_d$ | $\eta_2$ | | 0.0720*** | 0.0741*** | 0.114*** |
| | | | (0.0123) | (0.0152) | (0.0306) |
| Observations | | 244713 | 244713 | 244713 | 244713 |
| $R^2$ | | 0.401 | 0.401 | 0.565 | 0.768 |
| Sector-destination FE | | + | + | + | |
| Firm FE | | + | + | | |
| Firm-sector FE | | | | + | |
| Firm-sector-destination FE | | | | | + |

*Notes:* The dependent variable is the logarithm of the value of export revenue at the firm-product-destination-mode level, measured in terms of deviations from the respective product-level means. $CONT_{ajdm}$ is a binary variable that takes the value one if there is a positive containerized flow at the firm-product-destination level. Robust standard errors clustered at the product-destination level in parentheses. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 7: **Robustness Checks: Inferring containerization's Effects on Transport Costs from its Effect on Export Revenues**

| | Coefficient | (1) Containerizable products (OECD defn.) $\ln(r_{ajdm}/\overline{r}_j)$ | (2) Containerizable products (Bernhofen, El-Sahli, and Kneller (2016)) $\ln(r_{ajdm}/\overline{r}_j)$ | (3) Non-differentiated products $\ln(r_{ajdm}/\overline{r}_j)$ | (4) All products $\ln r_{ajdm}$ | (5) All products $\ln(r_{ajdm}/\overline{r}_j)$ |
|---|---|---|---|---|---|---|
| $CONT_{ajdm}$ | $\eta_1$ | -0.514*** (0.180) | -0.532*** (0.175) | -0.597* (0.342) | -0.647** (0.198) | |
| $CONT_{ajdm} \cdot \ln dist_d$ | $\eta_2$ | 0.116*** (0.0221) | 0.114*** (0.0214) | 0.114*** (0.0412) | 0.131*** (0.0243) | |
| $CONT_{ajdm} \cdot I\{dist_d \leq 1000\}$ | $\eta_2$ | | | | | 0.138** (0.0539) |
| $CONT_{ajdm} \cdot I\{1000 < dist_d \leq 4000\}$ | $\eta_2$ | | | | | 0.361*** (0.0285) |
| $CONT_{ajdm} \cdot I\{<4000 < dist_d \leq 8000\}$ | $\eta_2$ | | | | | 0.472*** (0.0408) |
| $CONT_{ajdm} \cdot I\{8000 < dist_d \leq 12000\}$ | $\eta_2$ | | | | | 0.535*** (0.0540) |
| $CONT_{ajdm} \cdot I\{dist_d > 12000\}$ | $\eta_2$ | | | | | 0.593*** (0.0683) |
| Observations | | 110542 | 117699 | 45376 | 244713 | 244713 |
| $R^2$ | | 0.587 | 0.587 | 0.589 | 0.765 | 0.587 |
| Firm-sector-destination FE | | + | + | + | + | + |
| Product-destination FE | | | | | + | |
| Firm-product FE | | | | | + | |

*Notes:* The dependent variable is the logarithm of the value of export revenue at the firm-product-destination-mode level, measured in terms of deviations from the respective product-level means. $CONT_{ajdm}$ is a binary variable that takes the value one if there is a positive containerized flow at the firm-product-destination level. Column (1) restricts the sample to containerizable products based on the classification suggested by Korinek (2011); the following products are excluded from the sample: dry and industrial bulk, tankers, and bulk shipped goods, goods shipped in tankers, motor vehicles, trailers and non-mechanically propelled vehicle n.e.s., aircraft, spacecraft, satellites, ships, boats and other floating structures. Column (2) also restricts the sample to containerizable products according to the classification employed by Bernhofen, El-Sahli, and Kneller (2016). Column (3) drops differentiated goods as classified by Rauch (1999). Column (3) adds product-destination and firm-product fixed effects to the specification presented in the last column of Table 6. Column (4) assumes a flexible functional form for transport costs defined in equation (11). $I\{\}$ are indicator variables which take on the value one if distance to destination satisfies the condition in parentheses, and zero otherwise. Robust standard errors clustered at the product-destination level in parentheses. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 8: **Transport Cost Estimation using U.S. Import Data**

| | Coefficient | (1) $\ln(CIF_{js}^m)$ | (2) $\ln(CIF_{js}^m)$ |
|---|---|---|---|
| $CONT_{jsm}$ | $\beta_1$ | 0.777*** | 0.487*** |
| | | (0.184) | (0.167) |
| $CONT_{jsm} \cdot \ln dist_s$ | $\beta_2$ | -0.103*** | -0.056*** |
| | | (0.020) | (0.018) |
| $\ln(Value_{jsm}/Weight_{jsm})$ | | | 0.578*** |
| | | | (0.003) |
| Observations | | 494598 | 494598 |
| $R^2$ | | 0.210 | 0.302 |
| Product FE | | + | + |
| Source country-port of entry FE | | + | + |

*Notes:* The dependent variable is the mode specific (breakbulk vs container) logarithm of freight and insurance charges per weight for 2013 U.S. import flows at the level of 8-digit HS-product $j$ and source country-port of entry $s$. $CONT_{jsm}$ is a dummy for containerized flows. $\ln dist_s$ is sea distance from the source country. Robust standard errors clustered at the product-source country level in parentheses. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Figure 1: **Distribution of Container Shares for Turkish Exports**
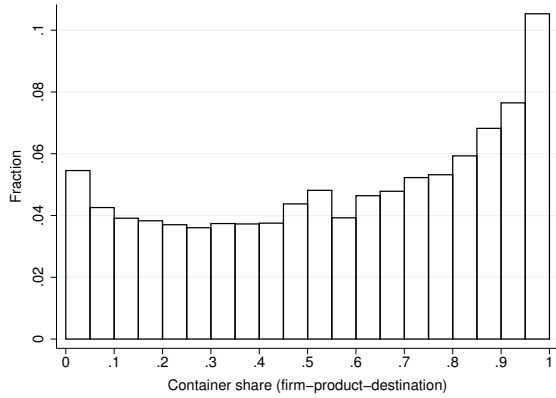


Panel A: Firm-Product-Destination
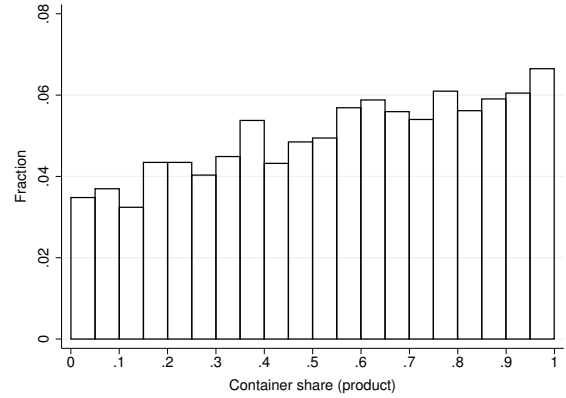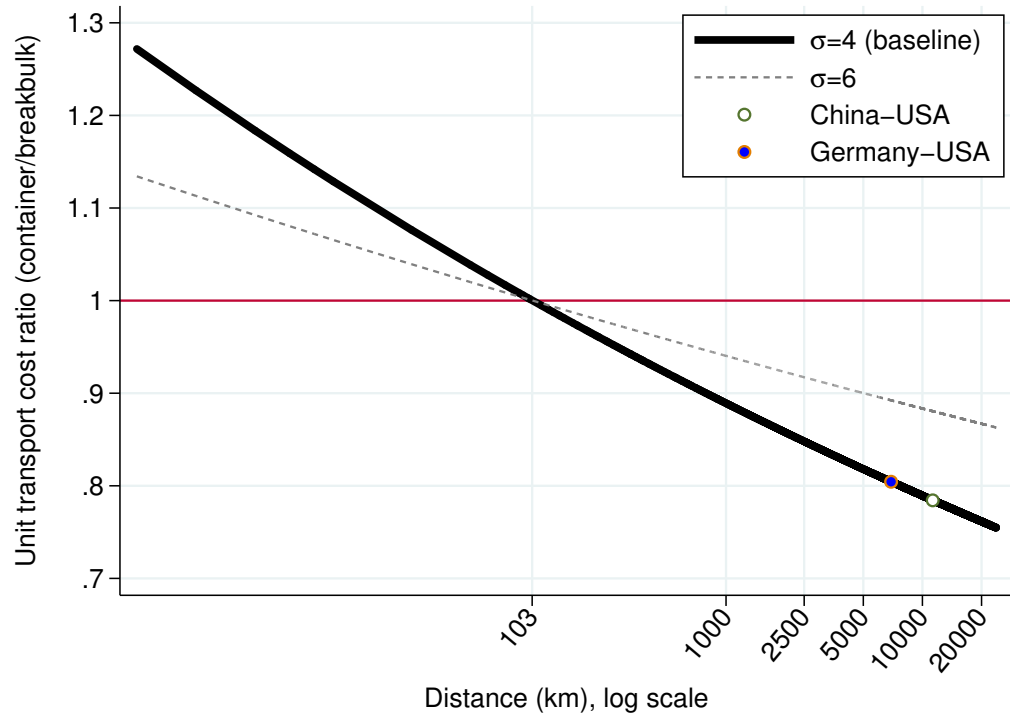
Panel B: Products
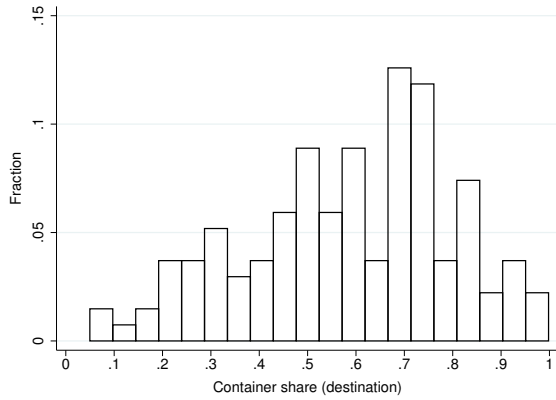
Panel C: Destinations
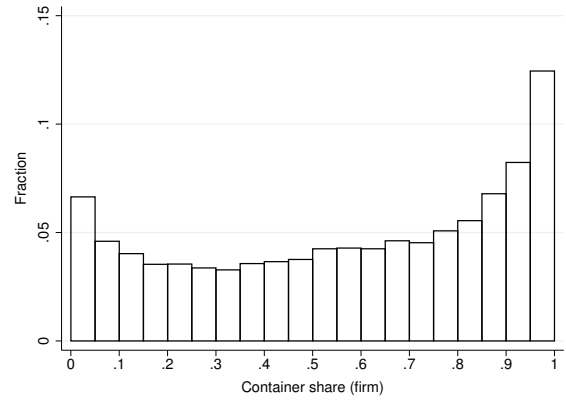
Panel D: Firms

*Notes:* Histogram of container shares, excluding mass points at 0 (no containerization at all) and 1 (full containerization). Panel A: 23,720 observations, Panel B: 4,166 8-digit HS products, Panel C: 135 destinations, Panel D: 8,941 firms.

Figure 2: **Relative Variable and Fixed Trade Costs (Additive Specification)**
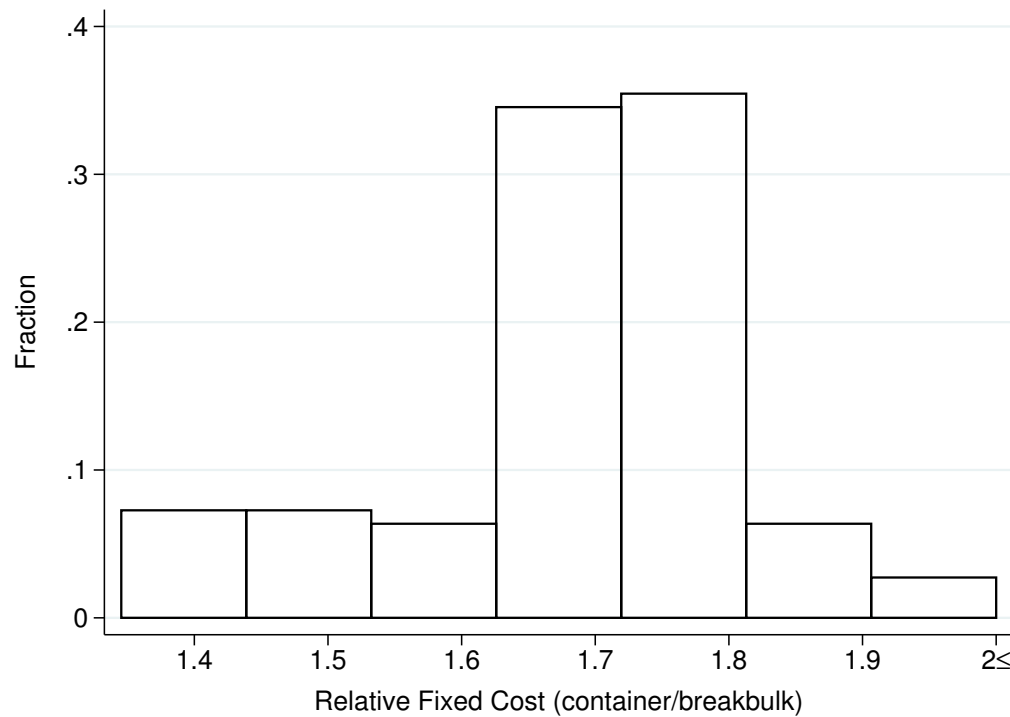
*Notes:* The upper panel shows the variable cost $\widehat{\Delta t_d} = t_d^c/t_d^b$ against distance. Red (blue) point marks USA-China (USA-Germany), and the horizontal line at 1 marks the breakeven distance. The bottom panel is the histogram of relative fixed cost values ($\widehat{\Delta f_d}$) calibrated to destination countries in the Turkish export data using the baseline value of $\sigma = 4$.

Figure 3: **Relative Variable and Fixed Trade Costs (Multiplicative Specification)**
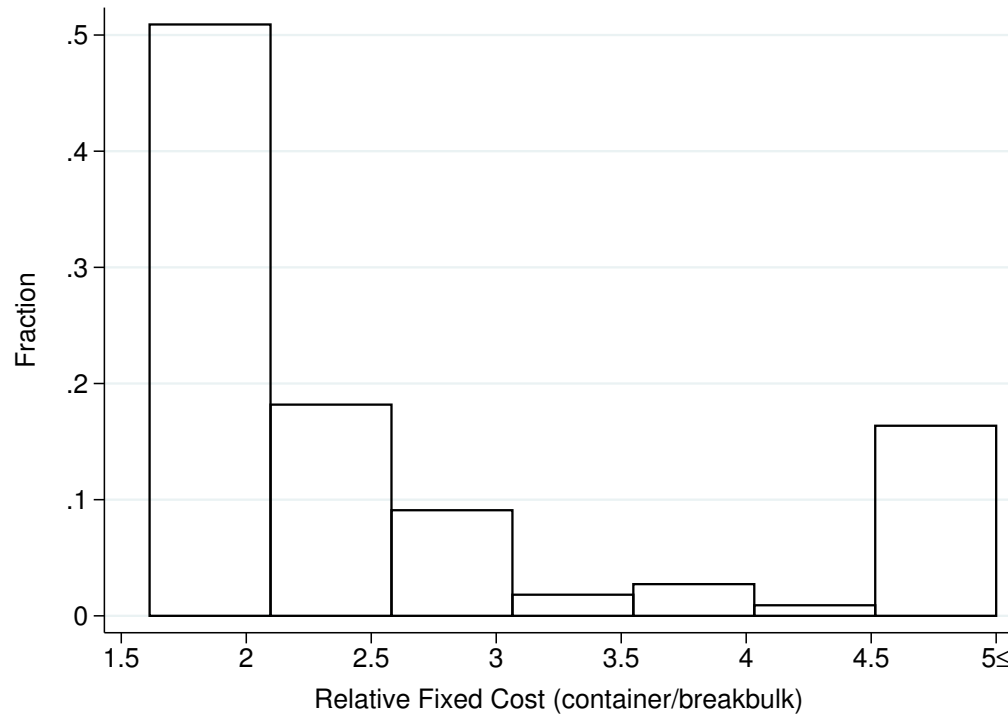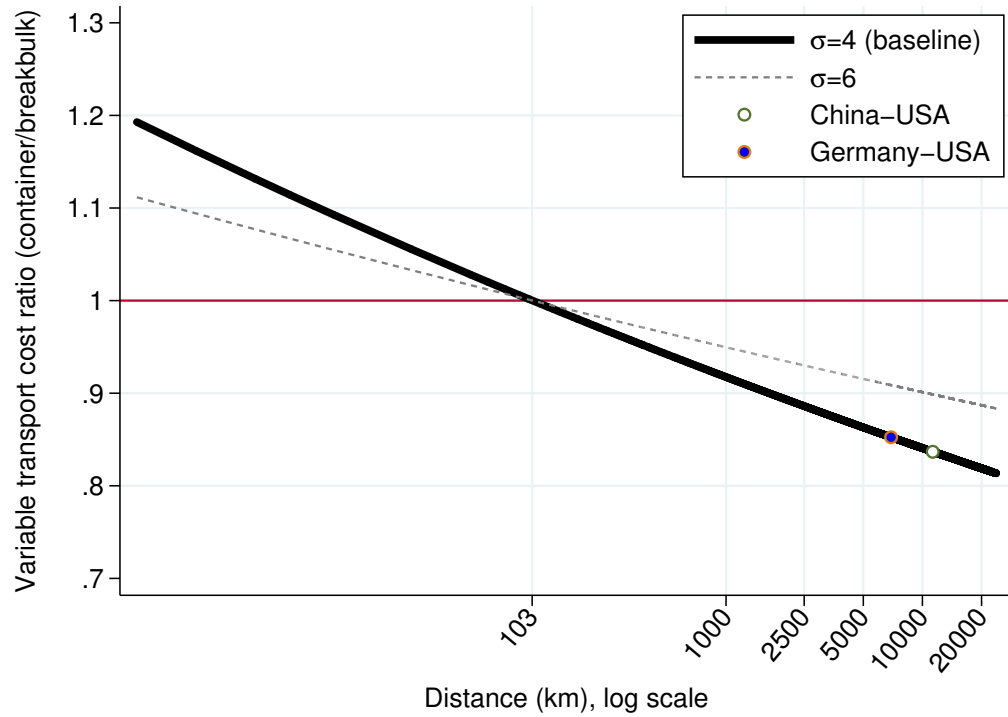


*Notes:* The upper panel shows the variable cost $\widehat{\Delta t_d} = t_d^c / t_d^b$ against distance. Red (blue) point marks USA-China (USA-Germany), and the horizontal line at 1 marks the breakeven distance. The bottom panel is the histogram of relative fixed cost values $(\widehat{\Delta f_d})$ calibrated to destination countries in the Turkish export data using the baseline value of $\sigma = 4$.

Figure 4: **Relative Cost of Containerized Charges for U.S. Imports**
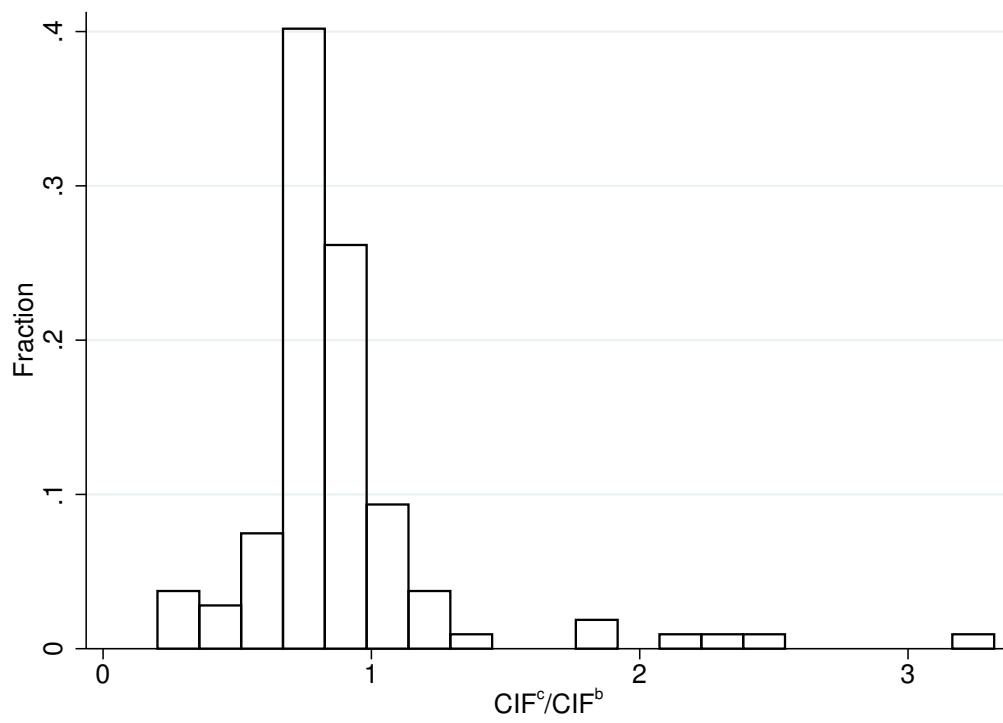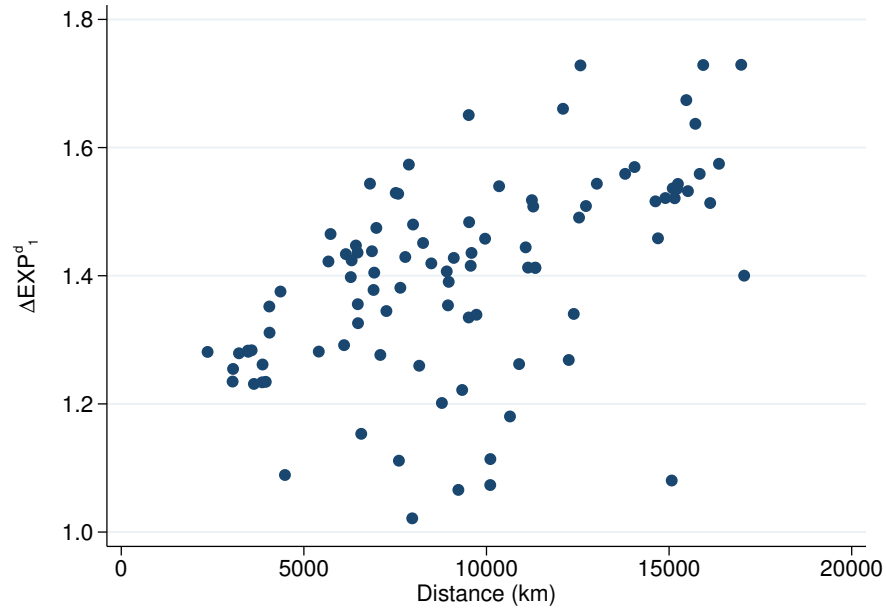


*Notes:* This figure plots insurance and freight charges per weight for containerized 2013 U.S. import flows relative to breakbulk across 8-digit HS products and source countries.

Figure 5: **Increase in Trade due to Containerization for the Median Sector**
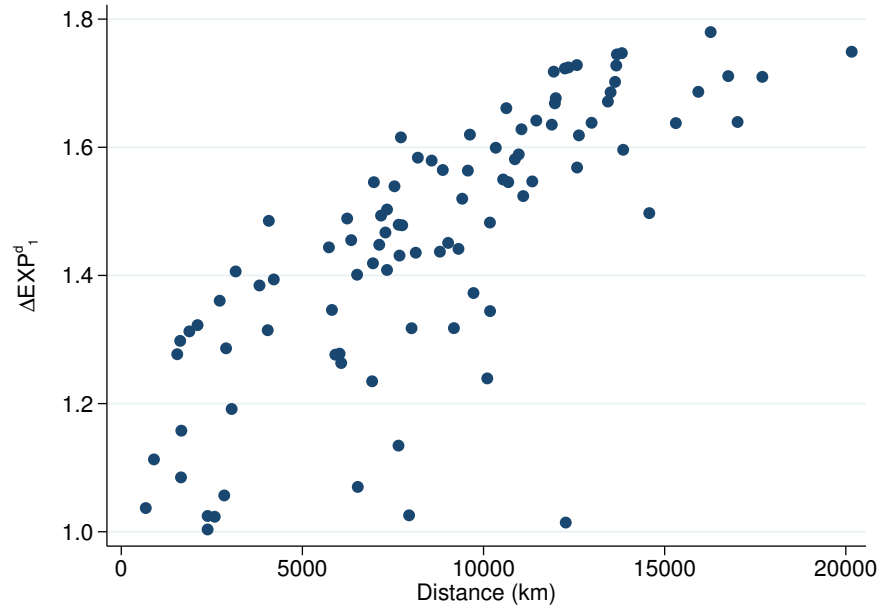


Panel A: Turkey



Panel B: U.S.

*Notes:* This figure plots $\Delta EXP_1^d$, the ratio of current level of destination-specific exports for the median sector in terms of container usage to the counterfactual level that would obtain if container shipping was not available. Each dot is an export destination from Turkey (top panel) and the U.S. (bottom panel) as the source country.