Beyond the Iceberg Hypothesis: Opening the Black Box of Transport Costs

Guillaume Daudin[*] Jérôme Héricourt[†] Lise Patureau[‡]

July 21, 2016

July 2016

Abstract

Following Samuelson (1954), standard models of international trade have usually relied on modelling trade costs as an ad valorem tax equivalent. However, many common empirical facts support the existence of additive costs. This paper measures the size of additive costs in transport, using SITC 3and 4 digit cif-fob unit values over 1974-2013 taken from US imports data. We estimate the two components of transport costs, by transport mode (air or ocean). We find that additive costs are 2.85% of fob unit values for ocean transport costs (and ad-valorem ones 3.22%). These values are respectively equal to 1.9% and 2.5% for air transport costs. We show that taking additive costs into account improves the fit of the modelling of transport costs. The time dimension of our data allows us to characterize the evolution of transport costs. After correcting for composition effects, we find that all types of transport costs have been roughly constant from 1974 to 1984 and then steadily decreased by 40% over the period 1984-2013. Yet, this steady decline hides shifts in the relative importance of additive and ad-valorem. E.g. most of the early decline in air transport costs can be explained by the ad-valorem component. By contrast, this component nearly doubled in the 2000s.

JEL classification: F14, N70, R40

Keywords: Transport costs estimates, transport costs determinants, non-linear econometrics, period 1974-2013, additive transport costs

[*]Université Paris Dauphine, LEDa & PSL, Chercheur associé - Sciences Po (OFCE)

[†]Université de Lille - LEM-CNRS (UMR 9221), Université de Bretagne Occidentale-ICI (EA 2652) and CEPII

[‡]Université Paris Dauphine, LEDa & PSL

1 Introduction

Trade costs remain central in international economic analysis. Defined as the costs associated with the exchange of goods across national borders, they are usually split into transaction costs (information costs, contract enforcement costs, costs associated with the use of different currencies...), policy costs (tariff and non-tariff costs), time costs (time to ship goods) and transport costs *per se*. Many believe that the latter have dramatically decreased with technological advance in transportation, infrastructure development and new communication technologies (see Lafourcade and Thisse, 2011). Glaeser and Kohlhase (2004) find that, over the twentieth century, the costs of moving goods have declined by over 90% in real terms. However, Hummels (2007) shows that the bulk of price declines in transportation comes from air shipping, where average cost per ton-kilometer shipped dropped by 92% between 1955 and 2004; concerning ocean shipping, which represents the major part of world trade value, decline in trade prices are much less obvious, even if the rise in containerization lowered shipping costs from 3 to 13%. Studies overviewed by Behar and Venables (2011) are consistent with those results: The fall in measured transport costs has been relatively small. According to these authors, this is attributable to the growing importance of fuel costs and technical progress, which improved speed and reliability rather than decreased costs.

Trade costs are considered as a major obstacle to international economic integration and international trade flows. According to the estimates by Jacks *et al.* (2008), trade cost declines explain around 55% of the pre-World War I trade boom and 33% of the post-World War II trade boom, while the abrupt rise in trade costs explains the inter-war trade collapse. After 1950, average trade costs fell by 16%, notably through the reduction of policy trade costs promoted through the GATT (WTO starting in 1995) multilateral agreements. Based on panel data, Novy (2013) thus finds that U.S. trade costs with major trading partners declined on average by about 40 percent between 1970 and 2000. Yet, several papers (mostly based on empirical estimates of the gravity equation) have shown that trade costs (typically captured through distance) still remain a major obstacle to trade (e.g. Head and Mayer, 2004 and Disdier and Head, 2008). Anderson and Van Wincoop (2004) estimate that average trade costs for industrialized countries put up to 170% markups over production costs, divided into 55% distribution costs and 74% international costs. Within this latter dimension, 44% are border-related trade barriers and 21% transportation costs. Lafourcade and Thisse (2011) also find that the share of transport costs in the consumer price of manufactured goods remains high, while Behar and Venables (2011) obtain that the elasticity of trade with respect to freight costs is sizeable, by around -3. If much of trade-policy barriers have been removed over the second half of the 20th century, these findings suggest that the transport costs component of the overall trade costs remain large and deserve attention. This is accordingly the focus of the paper.

Following Samuelson (1954), standard models of international trade have usually relied on modelling trade costs as an *ad valorem* tax equivalent (ie, as a constant percentage of the producer price per unit traded, the "iceberg cost" hypothesis). However, many common empirical facts support the existence of additive costs. As documented by Irarrazabal *et al.* (2015), pricing structure in shipping, additive tariffs, distribution costs... often exhibit (at least partly) an additive structure. The structure (additive vs multiplicative) of transport costs is far from being anecdotal, as the literature has long pointed out its role in shaping the pattern of trade flows. The Alchian and Allen conjecture

(Alchian and Allen, 1964), which points out that the relative price of two varieties of some good will depend on the level of trade costs, does rely on the existence of additive costs: The relative demand for more expensive/higher quality product goods should increase with trade cost ("shipping the good apples out"). Martin (2012) gives a strong empirical support to this conjecture: Based on a very disaggregated firm-product-level database of French exporters, he finds that firms charge higher fob unit values on exports to more remote countries, whereas the iceberg hypothesis would imply the opposite. Hummels and Skiba (2004) also find some strong evidence in favor of the Alchian-Allen conjecture: The elasticity of freight rates with respect to price is estimated to be well below unity, in contradiction with the iceberg assumption. Also, their estimates implied that doubling freight costs increases average fob export prices by 80-141 percent, consistent with high quality goods being sold in markets with high freight costs. **je ne sais pas quoi faire de la reference a lashkaripour. Complique d'en parler, car on ne peut pas le contredire car on n'a pas les donnees qui le permettent. Ne pas en parler? C'est un peu limite.** In contrast with most of the litterature, **?** finds supporting evidence for the iceberg assumption by taking into account the fact that more expensive goods are systematically heavier and hence more costly to transport. His study is restricted to 60% of US imports that are enumerated by items in the statistics. Furethermore, while the positive correlation between weight and price seems reasonable for goods from the second industrial revolution like cars, it is dubious in the case of ITC goods which importance has been rising since 1994 (the end point of Lashkaripour's study). The existence of additive costs may also explain a large number of zeros in bilateral trade flows, and more generally, the granularity of trade flows. Relying on Spanish and US transaction-level trade data, Hornok and Koren (2015) find that additive trade costs of are associated with less frequent and larger shipments, i.e. more "lumpiness", in international trade; in other words, exporters wait to fill completely a container before sending it abroad, to decrease as much as possible the number of shipments.

Beyond the positive aspect of understanding trade patterns, several recent papers also point out the normative implications of additive trade costs. Sorensen, 2014 extends Melitz (2003)'s seminal model of international trade by including additive trade costs, in addition to the iceberg component. A key analytical result is that the welfare gain from a reduction in trade barriers is higher for a decrease in additive costs than a decrease in multiplicative costs. Calibrating on Norwegian firm-level data for 2004, Irarrazabal *et al.* (2015) find that an additive import tariff reduces welfare and trade by more than an identically-sized multiplicative tariff. While these results suggest that important welfare gains can be achieved by reducing additive trade costs, not much can be done in quantifying such gains though, by lack of an empirical characterization of the additive component of trade costs.[1] One objective of the paper is to palliate this gap.

Our paper contributes to the literature by providing results on the size of transport costs over time, explicitly distinguishing between multiplicative and additive parts. More precisely, we provide estimates of the relative importance of additive and multiplicative transport costs over several decades. To do so, we update the detailed US customs, sector-level data from the US Imports of Merchandise used by Hummels (2007), to cover the period from 1974 to 2013. Closely related to this paper is the work by Irarrazabal *et al.* (2015), which develop a structural framework for inferring additive

[1] The one exception being Irarrazabal *et al.* (2015), upon which we come back later.

trade costs from firm-level trade data. Based on Norwegian firm-level data, their results suggest that additive costs are on average 30-45% relative to the median consumer price, and that they are strongly correlated with standard proxies for trade costs (like e. g., distance). However, while our data requires that we only consider transportation costs, our approach departs from theirs in several key aspects.

First, our theoretically agnostic approach provides a fairly simple framework for assessing *both* multiplicative and additive parts of transportation costs, which may prove to represent a non-negligible advantage for calibrating related models. We thus obtain that the mean values over 1974-2012 of iceberg costs are equal to 2.5% in air and 3.2% in ocean shipping, whereas the additive component amounts to 1.9% and 2.85% of the fob price in air and ocean shipping respectively. To our best knowledge, our paper is the first to provide such an extensive quantitative assessment of the magnitude of both multiplicative and additive costs in total transport costs. Second, we provide, through standard measures of "goodness-of-fit", an empirical assessment of what standard international trade models lose by skipping additive transport costs. Quantitatively, the omission of the additive term leads to overestimate the iceberg component by roughly a factor 2. On average over the whole period, biased estimate for iceberg is 5% for Air and 6% for Vessel, while unbiased estimate is respectively 2.5% and 3.2%. Third, the time dimension of our data allows us to characterize the evolution of transport costs, by transport mode (air or vessel) and for each specific (multiplicative and additive) cost over a forty years time span. After excluding the composition effects, we obtain that transport costs *per se* have roughly decreased by 40% over the period 1974-2012. Decomposing by transport mode and by type of transport costs allows to go deeper into this picture. First, as in Hummels (2007), we find that the transport costs decrease is more pronounced in air than in ocean shipping, starting in 1984. However, while ocean shipping costs display a regular decreasing pattern until 2012, air transport costs rather get stabilized over 2005-2012. Second, while both the additive and multiplicative components of ocean transport costs have roughly decreased by 40% over the period, air transport costs exhibit more contrasted trends. If the reduction of air shipping costs can most be attributed to the decrease in its multiplicative component over 1974-2005, the opposite scheme prevails afterwards, the reduction in per-unit air transport costs being compensated by an increase in the ad-valorem component.

Section 2 explains the data sources and the empirical methodology retained in the paper. Sections 3 and 4 report our results. In Section 3, we characterize the role of the additive component of transport costs. After reporting the mean values over the period (by transport mode), we show the improved performance in including additive trade costs in the measure of transport costs. This being established, Section 4 characterizes the trends in each component of transport costs (by transport mode) over the period 1974-2012. Section 5 concludes.

2 Data Sources and Empirical Methodology

2.1 A measure of Transportation Costs

As in Hummels (2007) (among others), our measure of transportation costs consists in exploiting the difference between commodity-level export and import prices. We first use values, quantities and freight costs to recover free-on-board (FOB) and cost-insurance-fret (CIF) prices, by goods, country of origin and transportation mode. More precisely, the (unit) FOB price is computed as the total customs value divided by the shipping weight; in other words, it is the price for the good net of transportation costs. The CIF price is then computed as the sum of the customs value and freight charges, once again divided by the shipping weight. Our dependant variable is finally computed as the ratio of the CIF price divided by the FOB price. Strictly higher than 1, the variable provides therefore with a measure of transport costs as a proportion of the good's price, an *ad valorem* equivalent.

The database we use to construct our measure of transport comes from US annual Imports of Merchandise provided by the Census bureau[2], spanning from 1974 to 2013. Using this dataset has (at least) three main advantages. First, this dataset has been used by Hummels (2007), which enables us to compare our results to his findings. However, we complement Hummels' (2007) findings as we extend the time coverage to 2013 (while Hummels (2007) stops in 2004). As we show in Section 4, extending the time period over he recent years delivers interesting insights regarding the trends in air shipping costs. Second, and importantly, this dataset delivers a strong statistical reliability arising from a single, trustworthy customs origin. Based on customs declarations, this dataset inventories all imports (both values and quantities) by origin to the United states at the HS 10-digit highly disaggregated level, with a concordance code to the SITC 5-digit coding system. In addition, the database reports information regarding freight expenditures and transportation mode (Ocean Vessel and Air). The first will be crucial to compute transport costs (see below), the second will allow us enlightening substantial differences in the dynamics of transport costs across transportation mode. Third, using this dataset allows us to have the import price of the good (CIF price), next to the export price (FOB price). This is highly valuable, as we can estimate both the *levels* of the iceberg trade costs and of the additive trade costs. This differentiates us from Irarrazabal *et al.* (2015), which can only estimate the ratio of additive costs as a share of the total consumer price (the only price they can encover).

It is also true that using this dataset has drawbacks. First, our measure of the cif-fob price gap only covers transportation costs by nature, thereby being mute about the others dimensions of international trade costs, unlike Irarrazabal *et al.* (2015). Further, in terms of transport costs *per se*, this measure being based on freight costs, omits the other dimension of transport costs related to the time value of goods in transit. According to Anderson and Van Wincoop (2004)), the 21% markup over production costs coming from transport costs includes both directly measured freight costs and 9% tax equivalent of the time value of goods in transit. **en faire un contre-argument positif**

Even if the data of transportation costs is available at a more disaggregated level, we use sectorial price data at the 3-digit level, primarily for technical reasons. As detailed below, the use of a nonlinear estimator triggers computational limitations that do not make them a likely option, especially when covering a long period of time. Yet, we ensure the robustness of these results by conducting the estimations at the 4-digit level (for some selected years). Comparing different levels of aggregation is useful to check differences and the presence of biases precisely due to aggregation. Depending on the considered year, this leaves us with around 200 (3-digits) and 600-700 (4-digits) products, from approximately 200 countries of origin.

[2]More information available at: `http://www.census.gov/foreign-trade/reference/products/catalog/fl_imp.txt`

2.2 Empirical specification

The estimated equation Our purpose is to provide estimates over time of the sizes of multiplicative and additive costs among total transport costs. To do so, we start from a very simple equation (similar to Irarrazabal *et al.* (2015) or Martin (2012), among others):

$$p = \tau\widetilde{p} + t \tag{1}$$

This equation expressed the consumer (import, or cif) price p as a function of the producer (export, or fob) price $\widetilde{p}$ given both per-unit (t) and ad-valorem (τ) transport costs. As usual in the literature, the so-called "iceberg" trade costs are denoted τ (with $\tau = 1$ meaning no iceberg trade costs), while additive trade costs are labeled t (with $t = 0$ implying no additive trade costs). We estimate this equation for each year over the period 1974-2013, and for each of the two transportation modes reported (air or vessel), on a sectoral-origin country basis. Let us denote i the origin country, and k, the sector (or product). Transforming the above equation (1) as ratio, we thus get the equation to be estimated as given by:[3]

$$\frac{p_{ik}}{\widetilde{p}_{ik}} - 1 = \tau_{ik} - 1 + \frac{t_{ik}}{\widetilde{p}_{ik}} \tag{2}$$

Estimation Strategy We follow Irarrazabal *et al.* (2015) by considering that *1)* both multiplicative and additive costs are separable between the origin country (i) and the product (k) dimensions, and *2))* in a multiplicative way for the former and an additive way for the latter. In other words, τ_{ik} and t_{ik} from Equation (2) become:

$$\tau_{ik} = \tau_i \times \tau_k \tag{3}$$

$$t_{ik} = t_i + t_k \tag{4}$$

As a result, our underlying theoretical equation is specified as:

$$\frac{p_{ik}}{\widetilde{p}_{ik}} - 1 = \tau_i \times \tau_k - 1 + \frac{t_i + t_k}{\widetilde{p}_{ik}} \tag{5}$$

The ratio $\frac{p_{ik}}{\widetilde{p}_{ik}}$ has a "one-lower bond", since by construction, the *cif* price p cannot be lower than the *fob* price: $p > \widetilde{p}_{ik}$. Taking into account this constraint in the estimation requires to impose a multiplicative structure for the error term, according to:

$$\frac{p_{ik}}{\widetilde{p}_{ik}} - 1 = \left(\tau_i \times \tau_k - 1 + \frac{t_i + t_k}{\widetilde{p}_{ik}}\right) \times \varepsilon_{ik}$$

Taking in log, this finally drives us to estimate the following equation:

$$\ln\left(\frac{p_{ik}}{\widetilde{p}_{ik}} - 1\right) = \ln\left(\tau_i \times \tau_k + \frac{t_i + t_k}{\widetilde{p}_{ik}} - 1\right) + \epsilon_{ik} \tag{6}$$

[3] Also keeping in mind that the estimated equation is transport-mode and year specific (air or ocean shipping). We skip the year and transport-mode dimensions here to alleviate notations.

where τ_i, τ_k, t_i and t_k are the parameters to be estimated, i.e., fixed effects specific to each origin country i and sector k, and $\epsilon_{ik} = \ln(\varepsilon_{ik})$.

The shape of our main equation 6 is such that estimations cannot be performed using standard linear estimators. Therefore, all estimates are performed using non linear least squares. The basis of the method is to approximate the model by a linear one and to refine the parameters by successive iterations. The intuitive criterion for convergence is that the sum of squares does not decrease from one iteration to the next. In our case, due to computational limitations implied by the size of our dataset, we implement 100 iterations and set the convergence criterion for successive parameter estimates and for the residual sum of squares at 0.01. Finally, to eliminate the potential influence of outliers, we excluded observations in the 5 percent from the upper and lower tails of the distribution in the regression variables, and all our three measures of trade costs are bounded by 0 as minimal value. These cut-offs are aimed at eliminating reporting or coding errors.

Further, one key objective of the paper is to characterize the importance of additive costs relatively to iceberg costs. Put differently, what traditional models of international trade lose by ignoring additive costs? A natural way to answer this question is to perform estimations of equation (6) constraining t to be equal to zero, and compare the fitting properties and the explanatory power of the restricted and complete models. This is done by computing several standard diagnostic statistics (R^2 and statistics based on the likelihood function). Accordingly, for each year and transport mode, we estimate two equations, depending on additive transport costs being included (Equation (7) or not (Equation (8)):

$$\ln\left(\frac{p_{ik}}{\widetilde{p}_{ik}} - 1\right) = \ln\left(\underbrace{\tau_i \times \tau_k}_{\tau_{ik}^{ice}} - 1 + \underbrace{\frac{t_i + t_k}{\widetilde{p}_{ik}}}_{t_{ik}^{add}}\right) + \epsilon_{ik} \tag{7}$$

$$\ln\left(\frac{p_{ik}}{\widetilde{p}_{ik}} - 1\right) = \ln\left(\underbrace{\tau_i \times \tau_k}_{\tau_{ik}^{nlI}} - 1\right) + \epsilon_{ik}^{nlI} \tag{8}$$

After estimating Equation (8), we can re-built a measure of each component $\widehat{\tau}_{ik}^{ice} = \widehat{\tau}_i \times \widehat{\tau}_k$ and $\widehat{t}_{ik}^{add} = \widehat{t}_i + \widehat{t}_k$, that is country-product specific, by year and transport mode. When assuming iceberg costs only (Equation (8), we proceed similarly to get $\widehat{\tau}_{ik}^{nlI} = \widehat{\tau}_i \times \widehat{\tau}_k$.[4] Still following Irarrazabal *et al.* (2015), we take the average over the product-country dimension, using the values of each trade flow (ik-specific) over total yearly trade as a weighting scheme. We thus recover a "synthetic estimate" of each type of transport cost $\widehat{\tau}$ and $\widehat{t}$, for each year and transportation mode. These results are reported in Section 3.

[4]In this case, notice that the equation could be estimated relying on a non-linear form. To preserve comparability of the results, we keep the same non-linear estimation method in both cases though.

3 Decomposing Transport Costs: The importance of the additive component

The objective of this section is twofold. First, we characterize the magnitudes of transport costs over time (by transport mode), distinguishing whether the additive component t_ik is included or not in the estimated equation (6). Second, we quantitatively assess the importance of the per-unit component of in overall transport costs through the means of goodness-of-fit measures.

3.1 Decomposing transport costs over 1974-2013

Our first contribution to the literature is to provide estimates for the size of both the multiplicative and the additive components of transport costs. Tables 1 and 2 report a summary of our results. Precisely, they display the mean and median values of each type of trade costs (multiplicative estimated alone, estimated along with additive costs and the additive component), expressed in percentage of the fob price, as well as the associated standard deviation, averaged over the period 1974-2013 and for selected years.

Table 1: Air (3 digits): Mean values over the period and selected years

	Year						
	1974	1980	1990	2000	2010	2013	**Mean stat**
With only Iceberg Trade Costs							
Mean	6.9	5.4	5.0	3.6	4.2	3.4	**5.1**
Median	5.4	3.8	4.4	2.5	3.4	2.9	**4.2**
Standard Error	0.052	0.049	0.039	0.033	0.037	0.024	**0.042**
With Additive & Iceberg Trade Costs							
Additive term							
Mean	2.6	2.0	1.8	1.3	1.1	1.0	**1.9**
Median	1.1	0.5	0.8	0.5	0.4	0.5	**0.7**
Standard Error	0.040	0.041	0.033	0.028	0.024	0.020	**0.076**
Iceberg term							
Mean	3.6	2.3	2.4	1.7	2.6	1.7	**2.5**
Median	2.7	1.6	1.6	1.2	2.2	1.7	**1.8**
Standard Error	0.032	0.025	0.021	0.016	0.023	0.012	**0.023**
# observations	14955	16118	24958	35027	40279	39351	

Statistics are obtained weighting each observation by its value in transport (mode-dependent). Mean and median values of all type of transport costs are expressed in percentage of the fob price.

Tables 1 and 2 yield the following results. First, the magnitude of overall transport costs is sizeable, either considering the average value over the period or even in the most recent years. Over 1974-2013, they correspond to a mean increase of the export price by 5.8% for ocean shipping, and by 5?1% for air shipping. **comparer avec la litterature?**

3.2 Asessing the importance of per-unit trade costs

In this section, we investigate the properties of our two models (with and without an additive term), to assess what loss of explanatory power is implied by the omission of the additive term. We start with a simple visual inspection of Figure 1, where are reported the yearly median *ad valorem* estimates

Table 2: Vessel (3 digits): Mean values over the period and selected years

	Year						
	1974	1980	1990	2000	2010	2013	**Mean stat**
With only Iceberg Trade Costs							
Mean	9.8	6.5	5.7	5.1	4.0	3.6	**5.8**
Median	9.6	5.5	4.6	4.9	3.6	3.3	**5.1**
Standard Error	0.053	0.040	0.032	0.028	0.020	0.018	**0.032**
With Additive & Iceberg Trade Costs							
Additive term							
Mean	5.1	3.4	2.7	2.8	2.5	1.5	**2.9**
Median	2.9	2.3	1.7	2.2	1.9	0.8	**1.9**
Standard Error	0.085	0.046	0.040	0.043	0.025	0.020	**0.041**
Iceberg term							
Mean	5.4	3.1	3.3	2.5	1.9	2.2	**3.2**
Median	4.9	2.4	2.8	2.1	1.8	1.8	**2.8**
Standard Error	0.041	0.023	0.022	0.021	0.018	0.018	**0.028**
# obs	19007	17356	28383	36090	37748	38473	

Statistics are obtained weighting each observation by its value in transport (mode-dependent). Mean and median values of all type of transport costs are expressed in percentage of the fob price.

of the iceberg costs ($\widehat{\tau}_i$) estimated in equation 6[5]. The dotted line plots the estimated iceberg cost ($\widehat{\tau}^{nlI}$) when the additive term t is restricted to 0 (Equation (8) being estimated), and the plain dark line represents the iceberg cost in the model including a non-null additive term (Equation (7) being estimated). In both cases, multiplicative costs are reported minus one, for a better reading comfort and comparability with Figure 2.

In Figure 2, we report the yearly median value for each component of transport cost (expressed in terms of the fob price, by transport mode)) after running estimation of Equation (7).

[5]Qualitatively similar results are obtained when the mean values are reported instead. More details of those results available upon request to the authors.

Figure 1: Multiplicative Costs (-1), Median Value 3 digits

(a) Air (b) Vessel

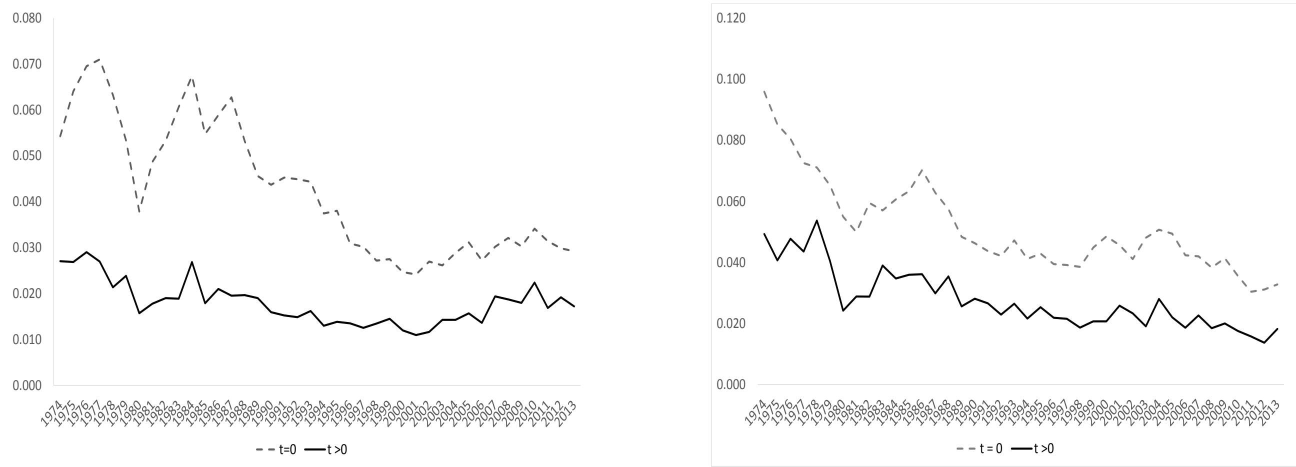
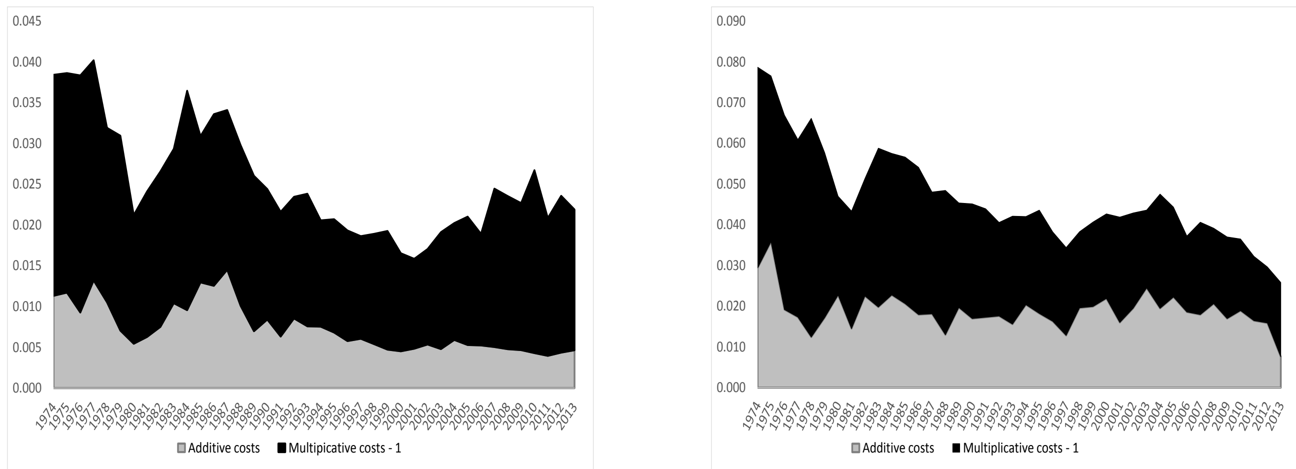


Figure 2: Multiplicative vs. Additive Trade Costs, Median Value 3 digits

(a) Air (b) Vessel



For both transport modes, we can infer from Figure 1 that the omission of the additive term seriously biases the multiplicative term upward. For Air, the size of the bias seems to decrease over time, while it seems pretty stationary for Vessel. In any case, the omission of the additive term leads to overestimate the iceberg component by roughly a factor 2.

In order to deliver a more systematic diagnosis, we use several standard measures of fit. Of course R^2 comes naturally to mind, but its use is far from being straightforward when evaluating non-linear estimates.[6]. Therefore, we provide as a complement the Standard Error of Regression (SER) which represents the average distance that the observed values fall from the regression line. Smaller values are better because it indicates that the observations are closer to the fitted line. We also report the log-likelihood function, and two measures derived, the Akaike Information Criterion (AIC) and the log-likelihood (LL) ratio test. A decrease in the log-likelihood function points to a better quality-of-fit. However, the likelihood function systematically decreases with the number of parameters included; the AIC criterion allows for correcting this overfitting by including a penalty in the computation of the statistic, so that AIC stat = 2 × number of parameters − 2 × Likelihood. Once again, the preferred model is the one with the minimum AIC value. Finally, the log-likelihood ratio test statistic compares systematically the likelihood of the Unrestricted model (UR, including an additive term, see equation 6) and the Restricted one (R, i.e. equation 6 with $t = 0$). The null tested is that the two models are statistically equivalent. Results are reported in Tables 3 and 4.

Unsurprisingly, it appears that the inclusion of the additive term leads to an improvement of the quality of fit, whatever the considered criterion or transport mode. On average over the whole period, R^2 doubles when restricting to Air Transport, and increases by 50% for Vessel. Similar qualitative conclusions arise from the comparisons of SERs, but may be more reliable on the quantitative side, since the use of R-squared on non-linear estimates may be qualified. Regarding the other criteria, improvements allowed by the inclusion of the additive term are roughly of the same extent across transport modes. Both AIC and Log-Likelihood statistics decrease with the inclusion of the additive term, and the LL test unambiguously reject the null of statistical equivalence of the two models. This is true whatever the considered year.

However, going into the examination of the levels of these goodness-of-fit statistics reveals intriguing differences. Multiplicative costs appear more important for the Vessel sector, where they account for 39% of the variance of the CIF/FOB ratio, versus 31% for the Air sector; the standard error of the regression based only on the iceberg term is consistently lower for Vessel than for Air. Besides, the dynamics over time also reveals intriguing differences. Quality of fit decreases strongly after 2000, for both transport modes. For Air, it seems that the additive term explains less and less variance over time: in 2000, the inclusion of the additive term allows adding more than 30% of explanatory power, and decreases the SER by 14 percentage points ; in 2010 and 2013, it is hardly more than 10% for the R^2, and 7 percentage points for the SER. The picture is quite different for Vessel. For that transport mode, the inclusion of the additive term increases the R^2 by a percentage very similar whatever the considered year (between 12 and 18%); a similar conclusion can be drawn for the SER, which decreases by 8-11 percentage points across the years. It appears that the deteriorating performance of the model in terms of goodness-of-fit from 2000 comes mainly from the decreasing performance of the iceberg term. In order to enlighten the roots of these differences, the following section elaborates further on the quantitative dynamics of each type of cost.

For comparison purposes, we provide in the appendix similar results for some years based on a more disaggregated dataset at the product level (4-digits).[7]. If anything the quality of fitting appears

[6]R-squared is based on the underlying assumption that the adjusted model is a linear one. In a non-linear context, R-squared is therefore inappropriate, strictly speaking. However, if the error distribution is approximately normal, a standard metric like R-squared remains informative on the quality of adjustment.)

[7]Since we use a non-linear estimator, estimations are highly computer-intensive. This explains why we decided to focus on the 3-digit product classification, which requires less time-consuming calculations.

slightly higher when estimations are based on the 4-digits classification. This is especially true for the model restricting trade cost to their iceberg dimension, whatever the transport mode considered. When the additive part is taken into account however, the difference becomes very small, whatever the considered criterion.

Using a more disaggregated classification unsurprisingly adds some statistical precision, but not to an extent which disqualifies the use of slightly more aggregated data.

3.2.1 Interpretation

We now more specifically investigate the long-run behavior of the estimated iceberg and additive parameters. Tables 1 and 2 report, respectively for the Air and Vessel sector, our estimates for Iceberg trade costs, first from equation (upper panel), then from equation 6 where they are estimated jointly with an additive term, which estimate is also reported (bottom panel). Remember that estimation is performed on a yearly basis, consequently, for the sake of clarity, we do not report estimates for each year, but for the beginning and the end of our sample, as well as starting year for each decade. we report in Tables 7 and 8 in the Appendix the estimated for coefficients for products disaggregated at the 4-digits level. In any case, these estimates are very close, and do not show any major difference with the baseline ones.

Unsurprisingly, the trend in transportation costs is clearly decreasing whatever the mode (air or vessel) or measure (with or without additive term) we consider. Starting from 1974, the decrease appears more important for sea than air transportation: transportation costs are divided by three for the first mode, and two for the second one. That said, it is widely know the two oil shocks inflated artificially transportation costs that specific year. If we start the analysis from 1980, the decrease is roughly 1.5-2 percentage points for Air, and 2-3 percentage points for Vessel - depending if we consider mean or median estimates. These figures are consistent with the ones by Hummels (2007). However, our methodology allows distinguishing in this trend, what comes from multiplicative and additive costs. To make a long story short, it appears that the additive costs represent the bulk of the decrease for Air, whereas the multiplicative component is the main responsible of the decreasing trend for Vessel, except for the years 2012-2103, witnessing an abrupt decrease in additive costs. To give a few average numbers, starting from 1980 and ending in 2013, the iceberg term decreases by 0.6 percentage point (pp), and the additive term is reduced by 1 pp for Air; for Ocean shipping, the reduction is roughly 2 pp for additive costs, and 1 pp for multiplicative costs.

4 Decomposing Transport Costs: Characterizing the trends

The evolution of transport costs depends both on the evolution of per product and per partner costs and on the evolution of the composition of trade. To isolate the evolution of per product and per partner cost, we estimate the following equations.

$$\begin{aligned}
\tau_{ikt} &= \exp\left(\sum_{i\neq\text{ARG}} \alpha_i.\mathbb{1}_i\right).\exp\left(\sum_k \beta_k.\mathbb{1}_k\right).\exp\left(\sum_{t\neq 1974} \gamma_t.\mathbb{1}_t\right).\exp\left(\epsilon_{ikt}\right) \\
\Leftrightarrow \ln(\tau_{ikt}) &= \sum_{i\neq\text{ARG}} \alpha_i.\mathbb{1}_i + \sum_k \beta_k.\mathbb{1}_k + \sum_{t\neq 1974} \gamma_t.\mathbb{1}_t + \epsilon_{ikt} \qquad (9)
\end{aligned}$$

$$\begin{aligned}
t_{ikt} &= \left(\prod_{i\neq\text{ARG}} \alpha_i.\mathbb{1}_i + \prod_k \beta_k.\mathbb{1}_k\right).\exp\left(\sum_{t\neq 1974} \gamma_t.\mathbb{1}_t\right).\exp\left(\epsilon_{ikt}\right) \\
\Leftrightarrow \ln(t_{ijt}) &= \ln\left(\prod_{i\neq\text{ARG}} \alpha_i.\mathbb{1}_i + \prod_k \beta_k.\mathbb{1}_k\right) + \sum_{t\neq 1974} \gamma_t.\mathbb{1}_t + \epsilon_{ijt} \qquad (10)
\end{aligned}$$

Both equations are estimated using the non-linear least square procedure in Stata. In the iceberg and multiplicative case, we report Γ_t, defined as:

$$\Gamma_t = 100.\frac{\bar{\tau}_{1974}.\exp(\gamma_t) - 1}{\bar{\tau}_{1974} - 1} \qquad (11)$$

Figure reports the time effect as the Γ_t in the iceberg and multiplicative case and $100.\exp(\gamma_t)$ in the additive case (assuming $\gamma_{1974} = 0$).

5 Conclusion

References

ALCHIAN, A. A. and ALLEN, W. R. (1964), *University Economics*, Belmont, CA: Wadsworth Publishing Company.

ANDERSON, J. and VAN WINCOOP, E. (2004), "Trade Costs", *Journal of Economic Literature*, vol. 42 n^o 3: pp. 691–751.

BEHAR, A. and VENABLES, A. J. (2011), "Transport costs and International Trade", in A. DE PALMA, E. Q. R. V., R. LINDSEY (editor), *A Handbook of Transport Economics*, Cheltenham: Edward Elgar Publishing Ltd.

DISDIER, A.-C. and HEAD, K. (2008), "The Puzzling Persistence of the Distance Effect on Bilateral Trade", *The Review of Economics and Statistics*, vol. 90 n^o 1: pp. 37–48.

GLAESER, E. L. and KOHLHASE, J. E. (2004), "Cities, regions and the decline of transport costs", *Papers in Regional Science*, vol. 83: pp. 197–228.

HEAD, K. and MAYER, T. (2004), "The empirics of agglomeration and trade", in J.V. HENDERSON, J. T. (editor), *Handbook of Regional ad Urban Economics, Volume IV*, Amsterdam: North-Holland.

HORNOK, C. and KOREN, M. (2015), "Per-Shipment Costs and the Lumpiness of International Trade", *Review of Economics and Statistics*, vol. 97 n^o 2: pp. 525–530.

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

HUMMELS, D. and SKIBA, S. (2004), "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture", *Journal of Political Economy*, vol. 112.

IRARRAZABAL, A., MOXNES, A. and OPROMOLLA, L. D. (2015), "The Tip of the Iceberg: A Quantitative Framework for Estimating Trade Costs", *The Review of Economics and Statistics*, vol. forthcoming.

JACKS, D. S., MEISSNER, C. M. and NOVY, D. (2008), "Trade Costs, 1870-2000", *American Economic Review, Papers and Proceedings*, vol. 98 n^o 2: pp. 529–534.

LAFOURCADE, M. and THISSE, J.-F. (2011), "New Economic Geography: the Role of Transport Costs", in A. DE PALMA, E. Q. R. V., R. LINDSEY (editor), *A Handbook of Transport Economics*, Cheltenham: Edward Elgar Publishing Ltd.

MARTIN, J. (2012), "Markups, Quality, and Transport Costs", *European Economic Review*, vol. 56: pp. 777–791.

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

NOVY, D. (2013), "Gravity Redux: Measuring International Trade Costs with Panel Data", *Economic Inquiry*, vol. 51 n^o 1: pp. 101–121.

Samuelson, P. (1954), "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments", *Economic Journal*, vol. 64: p. 264289.

Sorensen, A. (2014), "Additive versus multiplicative trade costs and the gains from trade", *Canadian Journal of Economics*, vol. 47 n^o 3: pp. 1032–1046.

Data Appendix

.1 Fob-cif prices

The Customs value is the value of imports as appraised by the U.S. Customs and Border Protection in accordance with the legal requirements of the Tariff Act of 1930, as amended. This value is generally defined as the price actually paid or payable for merchandise when sold for exportation to the United States, excluding U.S. import duties, freight, insurance, and other charges incurred in bringing the merchandise to the United States. The term "price actually paid or payable" means the total payment (whether direct or indirect, and exclusive of any costs, charges, or expenses incurred for transportation, insurance, and related services incident to the international shipment of the merchandise from the country of exportation to the place of importation in the United States) made, or to be made, for imported merchandise by the buyer to, or for the benefit, of the seller. In the case of transactions between related parties, the relationship between buyer and seller should not influence the Customs value.

In those instances where assistance was furnished to a foreign manufacturer for use in producing an article which is imported into the United States, the value of the assistance is required to be included in the value reported for the merchandise. Such "assists" include both tangible and intangible assistance, such as machinery, tools, dies and molds, blue prints, copyrights, research and development, and engineering and consulting services. If the value of these "assists" is identified and separately reported, it is subtracted from the value during statistical processing. However, where it is not possible to isolate the value of "assists", they are included. In these cases the unit values may be increased due to the inclusion of such "assists". Import Charges

The import charges represent the aggregate cost of all freight, insurance, and other charges (excluding U.S. import duties) incurred in bringing the merchandise from alongside the carrier at the port of exportation in the country of exportation and placing it alongside the carrier at the first port of entry in the United States. In the case of overland shipments originating in Canada or Mexico, such costs include freight, insurance, and all other charges, costs and expenses incurred in bringing the merchandise from the point of origin (where the merchandise begins its journey to the United States) in Canada or Mexico to the first port of entry. C.I.F. Import Value

The C.I.F. (cost, insurance, and freight) value represents the landed value of the merchandise at the first port of arrival in the United States. It is computed by adding "Import Charges" to the "Customs Value" (see definitions above) and therefore excludes U.S. import duties.

Table 3: Air: Measures of Goodness-of-fit

	Year						
	1974	1980	1990	2000	2010	2013	**Mean stat**
$\mathbf{R^2}$							
Term I only	0.30	0.27	0.25	0.32	0.42	0.31	**0.31**
Terms A & I	0.59	0.65	0.63	0.64	0.51	0.42	**0.60**
SER							
Term I only	0.79	0.86	0.81	0.84	0.86	0.92	**0.85**
Terms A & I	0.67	0.71	0.67	0.70	0.79	0.85	**0.73**
Log-likelihood							
Term I only	-17530.49	-20253.49	-29977.79	-43341.27	-50746.83	-52690.29	**-34888.64**
Terms A & I	-15125.65	-17263.20	-25393.46	-36788.44	-47277.53	-49419.70	**-30508.29**
AIC criteria							
Term I only	35674.98	41170.98	60715.58	87492.55	102297.66	106130.58	**70498.08**
Terms A & I	31387.29	35738.39	52098.91	74954.88	95887.05	100155.41	**62284.99**
Test LL							
2×(ll(UR) -ll(R))	4809.68	5980.59	9168.67	13105.67	6938.61	6541.17	**8760.69**
# restrictions	355	369	393	426	426.00	422	**401.93**
p-value	0.000	0.000	0.000	0.000	0.000	0.000	

Note: I = Iceberg; A = Additive; SER = Standard Error of regression; AIC = Akaike Information Criterion. Statistics are obtained weighting each observation by its value in transport (mode-dependent). Term A expressed in fraction of fob price. R^2 between the log of predicted ratio and the log of the observed ratio. The number # of restrictions is equal to the number of parameters estimated, i.e., the number of partner countries plus the number of products.

Table 4: Vessel, Measures of Goodness-of-fit

	Year						
	1974	1980	1990	2000	2010	2013	**Mean stat**
$\mathbf{R^2}$							
Term I only	0.45	0.41	0.46	0.40	0.35	0.34	**0.39**
Terms A & I	0.61	0.58	0.59	0.57	0.49	0.46	**0.56**
SER							
Term I only	0.58	0.62	0.59	0.65	0.74	0.76	**0.66**
Terms A & I	0.48	0.53	0.51	0.55	0.66	0.68	**0.57**
Log-likelihood							
Term I only	-16287.40	-16129.13	-25169.31	-35263.95	-41998.95	-43692.93	**-28534.30**
Terms A& I	-12985.76	-13353.65	-21171.37	-29490.96	-37418.66	-39751.86	**-24151.31**
AIC criteria							
Term I only	33328.81	33010.27	51142.62	71365.89	84789.89	88191.87	**57848.60**
Terms A & I	27331.52	28067.31	43664.74	60475.91	76161.33	80873.72	**49683.32**
Test LL							
2×(ll(UR) -ll(R))	6603.28	5550.96	7995.88	11545.98	9160.56	7882.15	**8765.97**
# restrictions	393	395	411	436	424	427	**417.05**
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: I = Iceberg; A = Additive; SER = Standard Error of Regression; AIC = Akaike Information Criterion. Statistics are obtained weighting each observation by its value in transport (mode-dependent). Term A expressed in fraction of fob price. R^2 between the log of predicted ratio and the log of the observed ratio. The number # of restrictions is equal to the number of parameters estimated, i.e., the number of partner countries plus the number of products.

Table 5: Air: Measures of Goodness-of-fit, 4-digits

	Year						
	1974	1981	1989	2001	2009	2013	**Mean stat**
R2							
Term I only	0.48	0.49	0.50	0.50	0.45	0.35	**0.47**
Terms A & I	0.63	0.66	0.65	0.66	0.54	0.45	**0.63**
SER							
Term I only					0.88	0.93	**0.89**
Terms A & I					0.80	0.86	**0.79**
Log-likelihood							
Term I only	-17505.55	-21813.46	-30960.56	-44067.62	-49375.57	-53197.87	**-34744.40**
Terms A& I	-14895.81	-18589.91	-26553.53	-37297.93	-45747.57	-49899.14	**-30243.95**
AIC criteria							
Term I only	36243.10	44966.91	63417.12	89747.24	100317.13	107963.73	**70940.07**
Terms A & I	31873.63	39495.82	55777.05	77439.85	94059.14	102224.28	**62955.73**
Test LL							
2×(ll(UR) -ll(R))	5219.47	6447.09	8814.06	13539.39	7255.99	6597.45	**9000.89**
# restrictions	640	698	778	833	824	818	**755.73**
p-value	0.000	0.000	0.000	0.000	0.000	0.000	

Note: I = Iceberg; A = Additive; SER = Standard Error of regression; AIC = Akaike Information Criterion. Statistics are obtained weighting each observation by its value in transport (mode-dependent). Term A expressed in fraction of fob price. R^2 between the log of predicted ratio and the log of the observed ratio. The number # of restrictions is equal to the number of parameters estimated, i.e., the number of partner countries plus the number of products.

A Results with a Higher Level of Products Disaggregation

Table 6: Vessel: Measures of Goodness-of-fit, 4-digits

	Year						
	1974	1981	1989	2001	2009	2013	**Mean stat**
$\mathbf{R^2}$							
Term I only	0.50	0.45	0.47	0.41	0.37	0.35	**0.44**
Terms A & I	0.66	0.62	0.62	0.58	0.51	0.46	**0.59**
SER							
Term I only					0.79	0.82	**0.77**
Terms A & I					0.69	0.75	**0.68**
Log-likelihood							
Term I only	-16460.10	-16951.61	-26771.44	-39008.34	-43888.90	-47161.62	**-29883.62**
Terms A& I	-12743.65	-13546.92	-21752.77	-33280.96	-39078.86	-43399.22	**-25303.92**
AIC criteria							
Term I only	34464.19	35491.21	55272.87	79800.67	89459.80	95987.23	**61425.60**
Terms A & I	28271.29	29877.84	46595.55	69743.91	81155.73	89692.44	**53573.29**
Test LL							
2×(ll(UR) -ll(R))	12385.80	11226.75	17354.65	20113.52	16608.16	12589.59	15704.63
# restrictions	797	814	881	910	886	874	860
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: I = Iceberg; A = Additive; SER = Standard Error of regression; AIC = Akaike Information Criterion. Statistics are obtained weighting each observation by its value in transport (mode-dependent). Term A expressed in fraction of fob price. R^2 between the log of predicted ratio and the log of the observed ratio. The number # of restrictions is equal to the number of parameters estimated, i.e., the number of partner countries plus the number of products.

Table 7: Air: Dynamics of Trade Costs over Time, 4-digit

	Year						
	1974	1981	1989	2001	2009	2013	**Mean Stat**
With only Iceberg Trade Costs							
Mean	1.066	1.058	1.052	1.033	1.037	1.032	**1.049**
Median	1.052	1.044	1.041	1.021	1.027	1.026	**1.037**
Standard Error	0.056	0.054	0.046	0.040	0.036	0.025	**0.045**
With Additive & Iceberg Trade Costs							
Additive term							
Mean	0.026	0.021	0.017	0.012	0.012	0.010	**0.019**
Median	0.012	0.006	0.006	0.005	0.004	0.004	**0.008**
Standard Error	0.039	0.042	0.033	0.027	0.029	0.019	**0.034**
Iceberg term							
Mean	1.035	1.026	1.031	1.015	1.021	1.016	**1.024**
Median	1.025	1.017	1.019	1.010	1.017	1.014	**1.016**
Standard Error	0.036	0.028	0.030	0.021	0.024	0.015	**0.026**
# obs	14944	16844	25307	35005	38475	39460	

Statistics are obtained weighting each observation by its value in transport (mode-dependent). Additive term expressed in fraction of fob price.

Table 8: Vessel: Dynamics of Trade Costs over Time, 4-digit

	Year						
	1974	1981	1989	2001	2009	2013	**Mean Stat**
With only Iceberg Trade Costs							
Mean	1.098	1.061	1.058	1.051	1.042	1.036	**1.060**
Median	1.094	1.051	1.048	1.045	1.038	1.031	**1.052**
Standard Error	0.060	0.038	0.036	0.030	0.023	0.020	**0.036**
With Additive & Iceberg Trade Costs							
Additive term							
Mean	0.046	0.026	0.031	0.024	0.021	0.015	**0.028**
Median	0.029	0.013	0.019	0.015	0.013	0.008	**0.017**
Standard Error	0.068	0.044	0.037	0.035	0.031	0.023	**0.039**
Iceberg term							
Mean	1.054	1.034	1.028	1.028	1.024	1.021	**1.033**
Median	1.049	1.030	1.024	1.025	1.026	1.018	**1.028**
Standard Error	0.043	0.026	0.025	0.021	0.016	0.013	**0.025**
# obs	19196	17916	29387	36677	37643	38820	

Statistics are obtained weighting each observation by its value in transport (mode-dependent). Additive term expressed in fraction of fob price.