

# INTERNATIONAL TRANSPORT COSTS: NEW FINDINGS FROM MODELING ADDITIVE COSTS\*

Guillaume DAUDIN<sup>†</sup>    Jérôme HÉRICOURT<sup>‡</sup>    Lise PATUREAU<sup>§</sup>

December 2018

## Abstract

This paper investigates the pattern of international transport costs over time, using information contained in the US imports flows from 1974 to 2013. First, we document the importance of the per-unit (additive) component of transport costs. We find that additive costs are quantitatively sizable, representing between one-third and one-half of overall transport costs. Second, we identify the respective roles of the reduction in “pure” transport costs and trade composition effects in the downward trend of international transport costs, in the same spirit as ?. Unlike him, we find that trade composition effects do not matter much and, when they do, they tend to amplify (rather than reduce) the decrease in pure transport costs. Importantly, this difference of results can be attributed to the new method of modeling the per-unit component of transport costs we offer. In both aspects, our results point to the importance of the additive component in accounting for international transport costs.

*JEL classification:* F14, N70, R40

*Keywords:* Transport costs estimates, non-linear econometrics, period 1974-2013, additive costs, trade composition effects

---

\*We are grateful to Holger Grg, Sébastien Jean and Vincent Vicard for their insightful comments. The paper has greatly benefited from remarks from the participants at ETSG 17<sup>th</sup> (Helsinki), DEGIT XXII (Paris), seminars at the universities of California-Irvine, Paris-Dauphine, Paris-I and Nancy and the Kiel Institute for the World Economy. Any remaining errors are ours. This paper was partly funded by a university grant from Lille-I University. It features an online appendix containing additional results, available on the authors’ websites.

<sup>†</sup>University Paris-Dauphine & PSL University, IRD, LEDa, UMR 225, DIAL, 75016 PARIS, FRANCE  
Sciences Po, Observatoire Français des Conjonctures Économiques (OFCE), 75014 PARIS, FRANCE  
; email: [guillaume.daudin@dauphine.psl.eu](mailto:guillaume.daudin@dauphine.psl.eu)

<sup>‡</sup>Université de Lille - LEM-CNRS UMR 9221, France & CEPPI, France; email: [jerome.hericourt@univ-lille.fr](mailto:jerome.hericourt@univ-lille.fr)

<sup>§</sup>Corresponding author. University Paris-Dauphine & PSL University, LEDa, 75016 PARIS, FRANCE;  
email: [lise.patureau@dauphine.psl.eu](mailto:lise.patureau@dauphine.psl.eu)

# 1 Introduction

Trade costs are central to international economic analysis. In particular, they are known to be a major obstacle to international economic integration and international trade flows. Based on panel data, ? finds that US trade costs with major trading partners declined on average by about 40% between 1970 and 2000. Trade costs declined faster for more developed countries (?). Still, several papers (mostly based on empirical estimates of the gravity equation) have shown that trade costs remain a major obstacle to trade (? or ?, to name a few). Using data over 1989-2000, ? estimate that average international trade costs represent a 74% markup over production costs for industrialized countries.

Defined as the costs associated with the exchange of goods across national borders, trade costs are usually split between transaction costs (information costs, contract enforcement costs, costs associated with the use of different currencies, etc.), policy costs (tariffs and non-tariff costs), time costs (time to ship goods) and transport costs *per se*. In this vein, ? find that around 30% of international trade costs are attributable to transport costs. Equivalently, international transportation costs represent a 21% markup over production costs. Further, the sizable elasticity of trade with respect to freight costs obtained by ? (around -3) testifies to the impact of transport costs on trade flows. With many trade policy barriers removed over the second half of the 20th century, this suggests that the transport costs component remains large and deserves attention. International transport costs accordingly stand at the heart of this paper.

This focus is shared with part of the international trade literature that studies the patterns of transport costs over time, such as ? and ?. Many argue that transport costs have substantially decreased with technological advance in transportation (?), infrastructure development and new communication technologies (?). However, there is no consensus. Distance as an incompressible cost is still a very important determinant of trade flows (?); further, progress in transportation techniques (which reduces the constant-dollar cost of transport) should be compared with technical progress in production (which reduces the constant-dollar cost of production) to get a clear picture of the evolution of the transport costs burden. In accordance with this view, ? points out that transport costs as a proportion of export prices have declined more slowly than is commonly assumed.

According to ?, this lower-than-expected decline of transport costs is partly explained by trade composition effects (i.e. changes over time in the baskets of imported goods and/or the origin countries). Based on US imports over 1974-2004, he finds that total transport costs have decreased in both air and maritime shipment, but less so than “pure” transport costs (i.e. *ceteris paribus* for given goods, mode, and trade routes) because trade composition effects have partly offset the pure transport costs decrease, thereby attenuating the overall downward trend in transport costs.<sup>1</sup> In this paper, we dig deeper into this issue. As a major contribution, and in contrast to ?, we find that trade composition effects play a minor role in accounting for the trend patterns of transport costs. Rather,

---

<sup>1</sup>Throughout the paper, we refer to “pure” or “*per se*” transport costs changes as the changes over time in the transport costs for a given product/country partner, ie excluding trade composition effects.

the evolution of transport costs *per se* constitutes the main driving force behind the overall transport costs decrease, in particular for air transport. In maritime shipment, trade composition effects matter more, but they amplify (rather than offset) the reduction in pure transport costs. In this respect, we find that pure transport costs declined even less since the 1970s than argued by ?.

Proving this point drives us to provide a careful modeling of transport costs. Following ?, standard models of international trade have usually modeled trade costs as an ad-valorem tax equivalent (i.e. as a constant percentage of the producer price per unit traded, part of the “iceberg cost” hypothesis).<sup>2</sup> However, a recent strand of the literature points to the existence of an additive component, that is, a cost per unit traded (see, among others, ?, or ?). Based on US sectoral data, our own estimates also bring out the importance of additive costs in international transport costs. We obtain this by providing an empirical decomposition of the structure of transport costs over time, in which we explicitly distinguish between ad-valorem and additive costs. To do so, we exploit information contained in the Census Bureau’s database of US imports flows over 1974-2013. This database contains a measure of international transport costs as the difference between the import and the export prices, on a year-product-partner country basis. This database is the same as that used by ?, extended over more recent years (2005-2013).

Our results may be summarized in two main new findings. First, additive transport costs are quantitatively sizable. On average over 1974-2013, the additive cost is estimated to be 1.8% and 2.9% of the export price in air and vessel transport respectively (at the 3 digit classification level). This is slightly lower than our estimates for the ad-valorem component (2.5% and 3.2% for air/vessel respectively on average over the period). Put differently, additive costs represent between one-third and almost one-half of the overall transport costs. To the best of our knowledge, our paper is the first to provide such an extensive quantitative measure of both ad-valorem and additive costs in total transport costs. This represents a valuable insight for calibrating related international trade or business cycle models. We also provide 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 ad-valorem component by roughly a factor 2. Further, the goodness of fit is systematically better when the additive component is included in the regression, even when taking into account the removed degree of freedom.

Second, we find that the main source of the downward trend in overall transport costs, by roughly 50% in air shipping and 60% in vessel shipping over 1974-2013, comes mainly from the reduction of transport cost *per se*. Trade composition effects have only limited influence, and when they matter (in maritime transport), composition effects amplify the reduction of pure transport costs. Otherwise stated, pure transport costs have declined less than observed costs. This stands in contrast to the conclusion obtained by ? on the same database (until 2004). Our result thus deepens the paradox of the less-than-

---

<sup>2</sup>Strictly speaking, the “iceberg cost” includes both the ad-valorem dimension of trade cost and the fact that these costs are paid in terms of the good that is traded. As this last element is irrelevant in our case, we use the terms “iceberg” and “ad-valorem” interchangeably, as is commonly done in the literature.

expected decline of transport costs. Importantly, we show that the difference of results can be attributed to the new method of modeling the additive component that we offer. Allowing the share of additive costs to vary across time, products and country partners (rather than being constant as in ?), turn out to be key in the decomposition of underlying sources of the decrease of overall transport costs observed over the period.

In both aspects, our results point to the importance of the additive component in accounting for international transport costs. In this respect, it contributes to the literature that challenges the dominant role of iceberg costs in international trade (even in the case of ice transport, see ?). As shown by ?, the relative price of two varieties of some goods will depend on the level of additive trade costs. In this context, the relative demand for more expensive/higher-quality product goods increases with trade costs (“shipping the good apples out”). This is consistent with the findings by ?, who estimate the elasticity of freight rates with respect to price to be well below unity. Also, their estimates imply that doubling freight costs increases average free alongside ship (fas) export prices by 80% to 141%, consistent with high-quality goods being sold in markets with high freight costs. Recent empirical studies based on micro-level data also provide strong empirical support for the role of additive costs (i.e. cost per unit exported) in international trade costs. Based on a firm-product-level database of French exporters, ? finds that firms charge higher unit values for exports to more remote countries, supporting the importance of additive costs. Using transaction-level trade data from Colombia, ? estimates that the transport cost elasticity varies greatly across industries, with a value of 0.8 for the average industry, in contrast with the iceberg specification.<sup>3</sup>

Closely related to our paper is the work by ?, which develops a structural framework for inferring relative additive trade costs from firm-level trade data. Implementing their methodology on Norwegian firm-level export data for the year 2004, they find that, for the median shipment, additive costs (in Norwegian crowns) amount to 6% of the export price multiplied by the ad-valorem cost (also expressed in Norwegian crowns). Our results share in common with ? the important role of the additive component of international trade costs. Yet, our paper complements their findings in two main respects. First, while our study covers a narrower set of trade costs focused on transport costs, their data and their empirical approach only allows the identification of the ratio of the additive cost to the export price multiplied by the ad-valorem cost; by contrast, our estimation strategy enables us to uncover separately both values of the ad-valorem and the additive costs. From this, we can rebuild the ratio in similar terms to them, thereby gaining in generality in this respect. While ? find that, for an export price multiplied by the ad-valorem cost

---

<sup>3</sup>Beyond the positive aspect, several recent papers also point out the normative implications of additive trade costs. ? extends ?’s seminal model of international trade by including additive trade costs, in addition to the ad-valorem 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 ad-valorem costs, due to the alteration of relative prices in a heterogeneous-firms trade framework. This is confirmed by ?. While these results suggest that important welfare gains can be achieved by reducing additive trade costs, not much progress has been made in quantifying such gains. One potential reason is the lack of an empirical characterization of the additive component of trade costs (the one exception being ?). One contribution of this paper is to fill that gap.

of 100 USD, 6 USD are paid in additive trade costs, our results point to a value paid in additive transport costs of 2.8 USD for maritime shipping and 1.5 USD for air transport for 2004. This comparison of results suggests that between a third and a half of the additive trade costs are attributable to the transport sector (depending on the transport mode). Second, we exploit exhaustive information about the imports flows of the US over a wide time-span from 1974 to 2013. In this respect, our results deliver a broader view of the magnitude of additive costs in international trade over time. In particular, we show that the modeling of additive costs is of critical importance in determining the underlying sources of the trends patterns of international transport costs observed since 1974.<sup>4</sup>

The paper is built as follows. First, we estimate the values of international transport costs annually throughout the period 1974-2013 by transport mode, explicitly distinguishing between the additive and the ad-valorem components. Section 2 reports the mean values of the transport costs estimated over the period (by transport mode). Relying on these estimation results, we then provide a decomposition of the transport costs trend patterns, between what arises from changes in the trade composition (by product and/or partner country), and what is attributable to changes in transport costs *per se*. Section 3 is devoted to this analysis, while Section 4 checks the robustness of our results to the separability assumption and the weighting scheme. Section 5 concludes.

## 2 Estimating international transport costs

### 2.1 Data

Our analysis of transportation costs consists of exploiting the difference between commodity-level export and import prices, as in ?. The database we use to construct our measure of transport costs comes from the US annual “Imports of Merchandise” provided by the US Census Bureau, spanning 1974 to 2013 (details on the database provided in Appendix A). We first use customs values, quantities and freight costs to recover FAS and cost-insurance-freight (CIF) prices, by good, country of origin and transportation mode.<sup>5</sup> More precisely, the (unit) FAS price is computed as the total “customs value” in the US trade statistics divided by the shipping weight; in other words, it is the price for one kg of the goods 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 FAS price. By construction

---

<sup>4</sup>On top of the previously cited papers about additive costs, our paper also relates to ?, who estimate the size and shape of per-shipment costs based on Swiss data, and to ? or ??, who point out the role of per-shipment costs (among which, administrative costs) in generating some “lumpiness” in international trade transactions.

<sup>5</sup>The related literature commonly refers to the FOB (free on board) price rather than the FAS price. The FAS price means that the seller must transport the goods all the way to the dock, close enough to be reached by the crane of the ship it will be transported in. It is also the seller’s responsibility to clear the goods for export. The FOB price means that the seller is obligated to bring the goods all the way to the port, clear the goods for export, and see that they are loaded onto the ship nominated by the buyer. Once the goods clear the railing of the vessel, the buyer assumes the risk. Note that this term is used exclusively for maritime and inland waterway transport. While both terms are closely related, the US Foreign Trade Statistics reports FAS prices.

higher than 1, the variable provides a measure of transport costs as a proportion of the goods' price, an ad-valorem equivalent. This is quite a standard and widespread strategy, as emphasized by ?.

Obviously, this dataset has limitations. First, it restricts our analysis to the study of international *transport* costs, as our measure of the CIF-FAS price gap only covers freight, insurance and handling costs. It is thus silent about the other dimensions of international trade costs. Second, in terms of transport costs *per se*, our measure omits the “indirect” costs related to the time value of goods on their way to their export market (including holding cost for the goods in transit, inventory cost due to buffering the variability of delivery dates, preparation costs associated with shipment size, etc.). In this respect, this dataset embraces only a partial view of international transport barriers. However, it is also clear that these direct transport costs do represent a sizable share of trade costs. According to ?, the 21% markup over production costs coming from transport costs includes both directly measured freight costs (11%) and “indirect” costs that amount to a 9% tax equivalent. Further, the evidence summarized in ? points to the persisting importance of direct transport costs, especially compared to other trade barriers. They remain more important than, e.g., policy barriers (8% tax equivalent), language barrier (7%) or information cost (6%).<sup>6</sup>

Exploiting this dataset also has three advantages. First, it comes from a single, homogenous and trustworthy customs source. This limits the measurement error bias. Based on customs declarations, the US Imports Database inventories all imports (both values and quantities), by country of origin, to the United States at the HS 10-digit level, with a concordance code to the SITC 5-digit coding system. In addition, the database reports information regarding freight, insurance and handling expenditures by transportation mode, land/ocean (or “vessel”) and air. We will make use of this to highlight potential differences in the dynamics of transport costs across transportation mode. Second, this dataset provides both the import price and the export price for the same good (for a given origin country). This is highly valuable, as it gives us a direct measure of international transport costs (at the country/sector level), from which we can estimate separately the *levels* of both the iceberg costs and of the additive costs, in contrast to e.g. ?. Third, this dataset is available over a long time span, which enables us to analyze the trend patterns of international transport costs over the period 1974-2013.<sup>7</sup>

As detailed below, the use of a non-linear estimator triggers computational limitations that limit the level of possible detail, especially when covering a long period of time. We estimate international transport costs at the 3-digit classification level, even if data series on the CIF and FAS prices are available at the 5-digit level. However, we ensure the robustness of these results by conducting the estimations at the 4-digit level for some

---

<sup>6</sup>In his survey, ? mentions several papers which all indicate that transport costs pose a barrier similar in size to, or larger than tariffs. In the same vein, ? highlight the importance of infrastructures for trade costs in general, through their impact on transport costs.

<sup>7</sup>Note that, as mentioned by ?, our transport costs measures are based on actual trade flows, i.e. ignoring those flows that did not happen because of presumably prohibitive transport costs. In light of this, the estimated values that we obtain can be viewed as the lower bounds of the “potential” transport costs.

selected years.<sup>8</sup> Depending on the year, this leaves us with around 200 sectors at the 3-digit level, from around 200 countries of origin.

## 2.2 Estimating transport costs: empirical specification

**The estimated equation** Our purpose is to provide time-varying estimates of the size of ad-valorem transport costs and additive transport costs. To do so, we start from the equation that expresses the price  $p$  (per kg) of goods paid by the importer (import, or CIF price) as a function of the producer price  $\tilde{p}$  (per kg) (export, or FAS price), given both additive (or per kg) transport costs ( $t$ ) and ad-valorem ( $\tau$ ) transport costs, according to:

$$p = \tau \tilde{p} + t \quad (1)$$

As is usual in the literature, the iceberg trade costs are denoted  $\tau$  (with  $\tau \geq 1$ ,  $\tau = 1$  meaning no ad-valorem trade costs), while additive trade costs are labeled  $t$  (with  $t \geq 0$ ,  $t = 0$  implying no additive costs). Let us denote  $i$  the origin country, and  $k$ , the product at the 5-digit level. Transforming the above equation (1) into a ratio, and removing 1 on both sides, in order to get the percentage change in prices induced by transportation (and get rid of the kg unit), we get the following baseline specification underlying our estimation (skipping the year and transport-mode dimensions in the notations for reading convenience):

$$\frac{p_{ik}}{\tilde{p}_{ik}} - 1 = \tau_{ik} - 1 + \frac{t_{ik}}{\tilde{p}_{ik}} \quad (2)$$

**Estimation strategy** We follow ? by considering that  $i$ ) both ad-valorem and additive costs are separable between the origin country ( $i$ ) and the product ( $k$ ) dimensions, and  $ii$ ) this separability is in a multiplicative way for the former and an additive way for the latter. In other words,  $\tau_{ik}$  and  $t_{ik}$  from Equation (2) are written as:<sup>9</sup>

$$\tau_{ik} = \tau_i \times \tau_k \quad (3)$$

$$t_{ik} = t_i + t_k \quad (4)$$

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

---

<sup>8</sup>The selected years for the 4-digit level estimations are: 1974, 1977, 1981, 1985, 1989, 1993, 1997, 2001, 2005, 2009, 2013. Comparing different levels of aggregation is useful to check differences and the presence of biases precisely due to aggregation. However, we obtain no substantial difference between the estimation results conducted at the 3 and 4-digit levels. Estimation results at the 4-digit classification level are reported in Appendix C.

<sup>9</sup>Notice that, given the magnitude or order of transport costs, assuming an additive or a multiplicative form for country/product fixed effects does not make a substantial difference since, for small values (as we generally obtain), we have  $\tau_i \times \tau_k - 1 \simeq (\tau_i - 1) + (\tau_k - 1)$  and  $t_i + t_k \simeq (1 + t_i) \times (1 + t_k) - 1$ . We also provide a robustness analysis to the separability assumption. Given the much higher number of fixed effects to estimate, we run this robustness check on a sub-sample of goods/origin countries. See Section 4 for a detailed presentation of this robustness check.

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

The ratio  $\frac{p_{ik}}{\tilde{p}_{ik}}$  has a lower bound of one, since by construction, the CIF price  $p$  cannot be lower than the FAS price ( $p_{ik} > \tilde{p}_{ik}$ ). Taking into account this constraint in the estimate implies that the error term should be always positive. We ensure that this constraint is fulfilled by specifying the error term as follows:

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

where  $\epsilon_{ik}$  follows a normal law centered on 0. Considered in logs, the above equation becomes:

$$\ln \left( \frac{p_{ik}}{\tilde{p}_{ik}} - 1 \right) = \ln \left( \tau_i \times \tau_k + \frac{t_i + t_k}{\tilde{p}_{ik}} - 1 \right) + \epsilon_{ik} \quad (5)$$

Equation (5) is non-linear. It cannot be estimated using standard linear estimators. All estimates are thus performed using non-linear least squares.<sup>10</sup> Yet, the use of a non-linear estimator triggers computational limitations that limit the level of possible detail, especially when covering a long period of time. Confronted with this trade-off, we estimate international transport costs at the 3-digit level as our benchmark classification, even though data series are available at the 5-digit level ( $k$ ). This amounts to making the additional assumption, that all 5-digit products  $k$  in a 3-digit sector  $s$  share the same structure of costs.<sup>11,12</sup>

This drives us to estimate a modified version of Equation (5), specified as:

$$\ln \left( \frac{p_{ik}}{\tilde{p}_{ik}} - 1 \right) = \ln \left( \tau_i \times \tau_{s(k)} + \frac{t_i + t_{s(k)}}{\tilde{p}_{ik}} - 1 \right) + \epsilon_{ik} \quad (6)$$

where  $\tau_i$ ,  $\tau_{s(k)}$ ,  $t_i$  and  $t_{s(k)}$  are the parameters to be estimated, i.e., fixed effects specific to each origin country  $i$  and sector  $s$  (at the 3-digit classification level), and  $\epsilon_{ik}$  the residual

<sup>10</sup>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 of residuals does not increase from one iteration to the next. See ? for more details.

<sup>11</sup>One way to gauge the relevance of this assumption is to provide a decomposition variance exercise on the observed CIF-FAS price. As reported in Appendix B.3, the share of the observed variance that is accounted for by the between-sectors ( $s$ ) variance is roughly similar to the between-product ( $k$ ) variance.

<sup>12</sup>One may object that we could preserve the estimation at the 5-digit level by running an OLS estimation on the equation taken in level, i.e. on the basis of Equation (2), specifying the error term additively. This would not solve the problem, however, for three main complementary reasons. First, at the 5-digit level the number of fixed effects to include in the estimation would be more than 430,000 (with 216 countries and 2,029 products), making the estimation computationally extremely burdensome, even in OLS. Imposing Equations (3) and (4) to reduce the number of fixed effects would drive us back to the non-linearity issue for fixed effects. Second, making the error term (specified as following a normal law centered on 0) enter Equation (2) additively implies negative values for some of the estimated residuals  $\hat{\epsilon}_{ik}$ , which is inconsistent with the constraint that  $\frac{p_{ik}}{\tilde{p}_{ik}} > 1$ . Third, estimating the equation in level does not eliminate the fact that it is non-linear by nature, as long as there are additive costs to estimate (see Equation (2) for  $t_{ik} \neq 0$ ).



centered on 0.<sup>13</sup> To eliminate the potential influence of outliers, we exclude 5% of the upper and lower tails of the distribution in the regression variables. These cut-offs are aimed at eliminating reporting or coding errors. We estimate Equation (6) for each year over the period 1974-2013, for each of the transportation modes reported (air or vessel), on a sectoral-origin country basis ( $i, s$ ). Depending on the year considered, this leaves us with around 800 fixed effects to estimate by transport mode at the 3-digit level.

As mentioned in the Introduction, one contribution of the paper is to provide a careful estimation of international transport costs, which decomposes into both an additive and an ad-valorem component. In this respect, our paper relates to the recent literature about the importance of additive costs in accounting for international trade costs (e.g., among others). We contribute to this issue by adopting the following strategy. For each year and transport mode, we estimate two models. In Model (a), additive costs are excluded, transport costs being modeled as iceberg costs only (Equation (7)). In this case, the estimated equation simplifies as:

$$\ln \left( \frac{p_{ik}}{\tilde{p}_{ik}} - 1 \right) = \ln (\tau_i \times \tau_{s(k)} - 1) + \epsilon_{ik}^{ice} \quad (7)$$

Under Model (b), transport costs are decomposed in the two additive and ad-valorem dimensions (Equation (6)). From this, we compare the fitting properties and the explanatory power of both models to evaluate the importance of modeling the additive component.<sup>14</sup>

One may be concerned that the specification of Equations (6) or (7) might be subject to an endogeneity bias, as the price set by the exporter may vary depending on the transport cost burden. Studies on the pricing-to-market behavior of firms (see e.g., ?) show that we cannot exclude that the export price set by the firm ( $\tilde{p}_{ik}$ ) is partly endogenous to the size of transport costs (for instance, the exporting firm absorbing (part of) the transport costs by reducing the FAS price). This is not an issue here, as we are not interested in causal inference. Rather, our aim is to provide an accounting breakdown of transport costs between the additive and the multiplicative components.

Otherwise stated, conditional to an export price and a transport mode, the gap between the export price declared to the custom data and the import price recorded by the US

---

<sup>13</sup>Note that, strictly speaking, the exact equation that we estimate is the following:

$$\ln \left( \frac{p_{ik}}{\tilde{p}_{ik}} - 1 \right) = \ln \left( \sum_i \alpha_i^\tau \mathbb{1}_i \times \sum_{s(k)} \alpha_{s(k)}^\tau \mathbb{1}_{s(k)} + \frac{\sum_i \alpha_i^t \mathbb{1}_i + \sum_{s(k)} \alpha_{s(k)}^t \mathbb{1}_{s(k)}}{\tilde{p}_{ik}} - 1 \right) + \epsilon_{ik}$$

After proceeding to the estimation, we uncover the estimated values of the transport costs components  $\tau_i$ ,  $\tau_{s(k)}$ ,  $t_i$ ,  $t_{s(k)}$  from the obtained coefficients  $\alpha_z^x$  for  $x = \tau, t$  and  $z = i, s(k)$  (by year and transport mode). For reading convenience, we adopt a lighter expression of this equation by resorting to the notations of  $\tau_i, \tau_{s(k)}$  and  $t_i, t_{s(k)}$  as a short-cut.

<sup>14</sup>One may object that a comprehensive study of the structure of transport costs should also include the third model with only additive costs. This has driven us to estimate this model as well, in which case the estimated equation is written according to:  $\ln \left( \frac{p_{ik}}{\tilde{p}_{ik}} - 1 \right) = \ln \left( \frac{t_i + t_{s(k)}}{\tilde{p}_{ik}} \right) + \epsilon_{ik}^{add}$ . The main result that emerges is that the model with additive costs only is dominated (in terms of quality of fit properties) by the model with multiplicative costs only (Equation (7)), which is itself dominated by the complete model (Equation (6)), anticipating further results. More details of these results are given in the Online Appendix.

administration is beyond the firm’s control. In this respect, we are confident that our estimation strategy is immune from endogeneity problems.

After estimating Equation (6), we can rebuild a measure of each component,  $\hat{\tau}_{is(k)}^{adv} = \hat{\tau}_i \times \hat{\tau}_{s(k)}$  for the ad-valorem cost and  $\hat{t}_{is(k)} = \hat{t}_i + \hat{t}_{s(k)}$  for the additive cost, that are country-sector specific, by year and transport mode. When assuming iceberg costs only (Equation (7)), we proceed similarly to get  $\hat{\tau}_{is(k)}^{ice} = \hat{\tau}_i \times \hat{\tau}_{s(k)}$ . In this case, notice that the equation could be estimated relying on a linear form. To preserve comparability of the results, we keep the same non-linear estimation method in both cases. Similarly, as ?, we take the average over the sector-country dimension, using the values of each trade flow (*is*-specific) over total yearly trade as a weighting scheme. We thus recover a “synthetic estimate” of each type of transport cost:  $\hat{\tau}^{ice}$  for Model (a),  $\hat{\tau}^{adv}$  and  $\hat{t}$  for Model (b), for each year and transportation mode. These results are reported in Section 2.3.

### 2.3 The importance of the additive component

**Quantification** The first step of the analysis consists of providing a quantification of the magnitude of transport costs over time (by transport mode), distinguishing whether the additive component  $t_{ik}$  is excluded or included in the estimated equation (Equation (6) or (7)). This allows us to provide estimates for the size of both the ad-valorem and the additive components of transport costs, and assess whether additive costs represent a sizable component of the latter. These results constitute our first original contribution to the literature. We also provide diagnostic tests about the goodness-of-fit measures, which confirm the empirical relevance of adding the additive component in the estimation (see results in Appendix B.2).

Table 1 reports a summary of our results. It displays the mean and median values of each type of transport costs (either ad-valorem estimated alone or estimated along with additive costs), as well as the associated standard deviation, averaged over the period 1974-2013, for estimation driven both at the 3- and 4-digit sectorial level for different specifications and data. The top panel reports the results for a specification based on ad-valorem costs estimated alone (Model (a), based on Equation (7)), while the intermediate panel presents estimates for a specification involving both ad-valorem and multiplicative components (Model (b), based on Equation (6)). Finally, the bottom panel reports the same set of descriptive statistics, but for the actual CIF/FAS ratio in our data.<sup>15,16</sup>

Table 1 calls for two types of comments. The first type is of statistical order. As displayed in Table 1, estimated mean values are systematically higher than medians, by around one percentage point. This result is not surprising, given that our key variable is by nature bounded to 0; therefore, outliers can only be positive, pushing the mean upward

<sup>15</sup>In Appendix B, we report similar results for a sample of years, for both transport modes, at the 3- and 4-digit classification level. Results for all years (available at the 3-digit level) are reported in the Online Appendix, available on the authors’ webpages.

<sup>16</sup>We present the estimation results for air transport at the 3-digit level, removing the year 1989, as the results reveal the presence of strong outliers that bias the estimates of transport costs upwards this particular year. Overall results (over the whole period) are not substantially affected if this year is included. This can be seen in the results detailed year by year in the Online Appendix, available on the authors’ webpages.

Table 1: Transport costs estimates: summary

| Mean value over 1974-2013   |          |          |          |        |
|---|----------|----------|----------|--------|
| # digit   | 3 digits |          | 4 digits |        |
| Mode  | Vessel   | Air (**) | Vessel   | Air    |
| <b>Model (a) - With only Ad-Valorem Transport Costs (<math>\hat{\tau}^{ice} - 1</math>, in %)</b> |          |          |          |        |
| Mean  | 5.8      | 5.1      | 6.0      | 4.9    |
| Median  | 5.1      | 4.2      | 5.2      | 3.7    |
| <b>Model (b) - With Additive &amp; Ad-Valorem Transport Costs</b>                                 |          |          |          |        |
| <i>Ad-valorem term (<math>\hat{\tau}^{adv} - 1</math>, in %)</i>                                  |          |          |          |        |
| Mean  | 3.2      | 2.5      | 3.3      | 2.4    |
| Median  | 2.8      | 1.8      | 2.8      | 1.6    |
| <i>Additive term</i>  |          |          |          |        |
| <i>In % of the export price (<math>\hat{t}/\tilde{p}</math>, in %)</i>                            |          |          |          |        |
| Mean  | 2.9      | 1.8      | 2.8      | 1.9    |
| Median  | 1.9      | 0.7      | 1.7      | 0.8    |
| <i>In USD per kg traded (<math>\hat{t}</math>)</i>  |          |          |          |        |
| Mean  | 0.09     | 1.12     | 0.08     | 2.16   |
| Median  | 0.07     | 1.04     | 0.06     | 1.06   |
| <b>Data</b>   |          |          |          |        |
| <i>Transport costs (<math>p/\tilde{p} - 1</math>, in %)</i>                                       |          |          |          |        |
| Mean  | 5.3      | 5.0      | 5.6      | 3.9    |
| Median  | 4.3      | 2.0      | 4.4      | 1.9    |
| <i>Export price (<math>\tilde{p}</math>), in USD</i>  |          |          |          |        |
| Mean  | 16.0     | 6488.4   | 9.6      | 6643.6 |
| Median  | 4.1      | 142.5    | 4.1      | 142.2  |
| # obs.  | 29279    | 28207    | 29317    | 27680  |
| # origin country  | 188      | 191      | 188      | 189    |
| # products  | 230      | 211      | 666      | 567    |

Notes: Statistics are obtained weighting each observation by its value relative to total trade flows. The additive term is expressed in fraction of FAS price. For the 4-digit classification, statistics for observed data have been calculated for the same set of years as used for estimation, i.e. 1974, 1981, 1989, 2001, 2009, 2013. (\*\*): 1989 omitted in 3-digit estimation for air.

compared to the median. One can also note that estimated total costs are systematically higher than observed costs, by 0.5 to 1 percentage point. This is first due to the fact that observed data are *de facto* trade-weighted (i.e. by the weight of each good in total trade) while our estimates are based on simple arithmetic means within sector and country. Furthermore, Table 1 reports values expressed in percentages (for the observed values,  $\frac{p_{ik}}{\bar{p}_{ik}} - 1$ , in %), whereas all our estimates are based on log-linearization. Therefore, what matters on the statistical ground is that the mean of our predicted values (based on logarithms) must be equal to the mean of the log-linearized data ( $\ln \left( \frac{p_{ik}}{\bar{p}_{ik}} - 1 \right)$ ), which does not necessarily imply equality between values. Descriptive statistics showing that both observed and predicted values do match in logarithms are available upon request to the authors.

**Interpretation** Coming now to interpretation, two main results emerge from Table 1. First, it provides estimation results about the size of the overall transport costs. Table 1 shows that transport costs amount to around a 6% markup in ocean shipping, and 5% markup in air shipping on average over 1974-2013.<sup>17</sup> If this seems modest in comparison with the 11% of markup obtained by ? over a set of industrialized countries, it stands in line with the results obtained by ? for the US economy.<sup>18</sup>

Second, and most importantly, Table 1 provides new quantitative evidence about the size of the additive component of transportation costs. On average over the whole period, for sectors defined at the 3-digit level, additive costs amount to 10 cents (0.1 USD) per kilogram exported in maritime transport, and 1.12 USD per kilogram exported in air transport, indicating that freight rates are much higher in the latter case. Expressed in terms of the export price, the additive costs amount to 2.9% and 1.8% for ocean shipping and air transport respectively. This is slightly lower than the estimates for the ad-valorem component, which is estimated at 3.2% and 2.5% for vessel and air shipping respectively (on average over the period). Put differently, the additive component amounts to 48% of total transport costs in vessel, and 42% in Air. This shows that the per-kg cost dimension is quantitatively sizable. For both transport modes, the omission of the additive term seriously biases the ad-valorem term upward, by roughly 2. Additionally, Appendix B.2 reports quality-of-fit diagnostic tests that confirm the importance of the additive cost component. The goodness of fit is systematically better when the additive component is included in the regression, even when taking into account the removed degree of freedom. An important aspect of these results is also to provide quantitative estimates for the values of both additive and ad-valorem transport costs, which can be very useful for more theoretical approaches needing calibration.

<sup>17</sup>Expressed as a markup over production costs, transport costs are equal to  $\frac{p-\bar{p}}{\bar{p}}$ . Under Model (a) (ad-valorem costs only), making use of the figures reported in Table 1 gives a markup of 5.8% in vessel, and 5.1% in air shipping (considering the mean value over the period). Under Model (b), total transport costs are equal to  $\frac{p-\bar{p}}{\bar{p}} = (\tau - 1) + \frac{t}{\bar{p}}$ . With  $\tau - 1 = 0.032$  and  $\frac{t}{\bar{p}} = 0.029$  in vessel (as reported in Table 1), we get a value equal to 6.1% in vessel. A similar calculus gives 4.3% in air transport.

<sup>18</sup>Beyond country coverage and time period, another plausible explanation for the difference with ?, lays in the difference of empirical methodology. ?'s estimates are based on standard gravity equations, while ours come from a non-linear estimation of a "structural" decomposition of the difference between export and import prices.

Our estimation results on the size of additive transport costs can be compared to ?, the most recent evidence on this question. Based on Norwegian trade data for 2004, ? estimate that additive costs (in Norwegian crowns) amount to 6% of export price multiplied by the ad-valorem cost (expressed in Norwegian crowns), i.e.  $\frac{t}{\tau p} = 6\%$  expressed in our terminology (with  $\tau > 1$ , and retaining their estimated weighted mean value, which is the one directly comparable to our own results). One important difference with respect to their work is that we can provide separate estimates for each cost component (additive and ad-valorem). From this, we can rebuild the ratio in similar terms to them, allowing us to gain in generality in this respect. Making use of our estimates for 2004 (see Tables B.1 and B.2 in Appendix B), we thus obtain a ratio  $\frac{t}{\tau p} = 2.8\%$  in ocean shipping and 1.5% in air transport. This may seem surprisingly low, in contrast to the 6% obtained by ?. This can be accounted for by recalling the difference in the type of dataset - hence, of costs, embraced in each case. While the database of ? allows studying trade costs in general, our database covers a subset made of the monetary international transport costs, as we start from the gap between the import and the export prices (on top of differences due to the country coverage of the two databases). In this respect, our two papers can be used in a complementary fashion to infer the proportion of additive trade costs that are due to international shipment. ?'s results imply that, for an export price multiplied by the ad-valorem cost of 100 USD, 6 USD are paid in additive trade costs. For the same export price multiplied by the ad-valorem cost of 100 USD, our results yield the value of 2.8 (1.5) USD paid in additive transport costs for maritime (air) transport.<sup>19</sup> This comparison of results suggests that between 25% (air) and 47% (vessel) of additive trade costs are attributable to international transport. On top of our results summarized in Table 1, this points to evidence that the additive component represents a quantitative sizable dimension of international transport costs.

**Time dynamics** Given the time coverage of our database, one may then wonder whether this important part of additive costs we put forward, is a long-lasting, i.e. a structural phenomenon. We investigate this point by reporting the share of additive costs in total transport costs on a yearly basis, in both air and vessel transport, based on the weighted-mean values  $\hat{t}/\hat{p}$  and  $\hat{\tau}$  (in similar terms as those reported in Tables B.1 and B.2. The corresponding values reported in detailed tables are available in the Online Appendix), as displayed in Figure ???. Each yearly value is obtained as:  $\frac{\frac{\hat{t}}{\hat{p}}}{\hat{\tau}-1+\frac{\hat{t}}{\hat{p}}}$ , expressed in percent of the export price  $\hat{p}$  (at the 3-digit level).

In both transport modes, the additive component appears of sizable importance throughout the period. Confirming the results obtained in Table 1, this suggests that additive costs are neither negligible in magnitude nor an erratic phenomenon. On the contrary, they represent a sizable and structural dimension of international transport costs. If the

<sup>19</sup>Also notice that, in both papers, starting from the estimated value of the ratio  $\frac{t}{\tau p} = x$ , we can rebuild the proportion of additive costs in terms of the import price  $\frac{t}{\tau p+t} = \frac{x}{1+x}$ . Given our estimates for 2004 ( $x = 0.028$  and  $0.015$  in vessel and air transport respectively), we get a ratio  $\frac{t}{\tau p+t}$  equal to 2.7% and 1.4% in each transport mode respectively.

share of additive costs in total costs roughly amounts to 48% (42%) in vessel (air) transport on average, one can further notice that it varies substantially over time, from a minimum of 30% to a maximum of 60% depending on the year considered. In the next section, we further investigate the importance of the (varying pattern of) additive costs, as regards the determinants of transport costs time trends.

### 3 Transport costs time trends: the role of the additive component

In this section, we investigate the trend patterns of international transport costs over time, by exploiting the time dimension of our database. This focus is shared with some of the international trade literature, such as ?, ? and ?. Specifically, in the same spirit as ? we investigate the underlying sources behind the downward trend of international transport costs, by identifying the respective roles of the reduction in “pure” transport costs and changes in trade composition effects.

#### 3.1 A first look at the trend patterns

As a first step, Figure ?? displays the evolution of total transport costs over the period, summing the additive and the ad-valorem estimated costs, expressed in percentage of the FAS price,<sup>20</sup> by transport mode for each year between 1974 and 2013.

#### **JH: ATTENTION!!! TYPO changer dans la Figure 2**

As shown in Figure ??, both air and vessel shipping exhibit a downward trend in overall transport costs since 1974, by -2.1% per year for mean air transport costs and -2.0% per year for mean vessel transport costs, implying a 50% decrease in air and a 60% decrease in vessel over the period.<sup>21</sup> Using US data, ? found that overall transport cost declined from 6% to 4% of the import value between 1974 and 1996. For the same years, we obtain a total decrease from 6.9% to 4.2% in terms of the export price, on average for air shipping, and from 9.8% to 4.8% for ocean shipping. Overall, our results display trends that are close to those reported by ?.

Before making any definite statement, however, it is worth emphasizing that the time trend of international transport costs depends on both the evolution of per product and per partner (“pure”) transport costs and the evolution of the composition of trade. Total transport costs may thus have decreased over time because the share of neighboring

---

<sup>20</sup>Specifically, we use the yearly values of the estimated transport cost components averaged over the country/product dimensions (at the 3-digit level), corresponding to the numbers reported for each year in detail in the Online Appendix. Each yearly value is obtained as:  $\hat{\tau} - 1 + \frac{\hat{\tau}}{p}$ .

<sup>21</sup>One may be puzzled by the high magnitude of estimates for the beginning of the period (until 1980 approximately) for vessel transport. ? finds similar outcomes on tramp prices indexes, and suggests the oil shock as a likely culprit, in a context where technological progress was quicker in aviation than in vessel, allowing a better dampening of oil shocks on air freight rates. In a related manner, one may worry that the strong decrease in transport costs documented in Figure ?? springs from high oil-shock-related transport costs in 1974. However, computing the time trends from 1980 does not dramatically change the picture. The yearly trend from 1980 is -2% for mean air transport costs and -1.6% for mean vessel transport costs. We thus choose to exploit the whole time dimension of our database by taking 1974 as starting date of our time-trend analysis.

countries in total US trade or the share of goods cheaper to transport has increased, independently of any change in transport costs *per se*. As emphasized by ? and ?, it is hence necessary to eliminate the composition effects of trade flows to isolate the evolution of pure international transport costs. Specifically, Hummels’s (2007) results underlie the view that trade composition effects partially compensate the reduction in pure transport costs, thereby attenuating the downward trend in overall (observed) transport costs since 1974. Our own results tend to challenge this view, as we further show.

**A word on Hummels’ methodology** At this point, let us make a brief summary of Hummels’ methodology. He starts from the following specification of the ratio of destination to origin prices modelled as  $p_{ikt} = \tilde{p}_{ikt} + \frac{f_{ikt}}{\tilde{p}_{ikt}}$ , with  $f_{ikt}$  the shipping charge per kg shipped (for a given product  $k$  imported from country  $i$  in year  $t$ ). Further, and as in ?, ? writes the per kg shipping charge as  $f_{ikt} = \tilde{p}_{ikt}^{1-\beta} X_{ikt}$ , where  $X_{ikt}$  represents other costs shifters (distance, port quality, etc.) Accordingly, his transport cost measure  $TC_{ikt}$  writes is written as:

$$TC_{ikt} \equiv \frac{p_{ikt} - \tilde{p}_{ikt}}{\tilde{p}_{ikt}} = \tilde{p}_{ikt}^{-\beta} X_{ikt} \quad (8)$$

Notice that  $\beta$  can be interpreted as the elasticity of transport costs to the export price (in absolute value). As such, a value of  $\beta = 0$  corresponds to the standard case of “iceberg” costs, in which transport costs are purely ad valorem. At the other extreme,  $\beta = 1$  means that all transport costs are additive. Accordingly, the estimated value of  $\beta$  can be interpreted as the share of additive costs in total transport costs. Equation (8) is the baseline specification from which ? decomposes the changes in transport costs over time between trade composition effects and changes in “pure” transport costs, as described in more detail in Appendix D.2. In doing so, ? implicitly assumes  $\beta$  constant across the triplet origin country origin/product/year. Put differently, this means that the share of additive costs in total costs does not vary across these dimensions.

**Challenging the  $\beta$ ’s constancy assumption** As Figure ?? (based on the yearly estimates of transport costs averaged over the country/sector dimension) suggests, the share of additive costs is rather varying over time, in both transport modes. We investigate this point deeper by reporting the histogram of the distribution of the share of additive costs in total costs. To do so, we start from Equation (2) with the trade costs measure on the left-hand side, from which we deduce the formulae to get the elasticity of transport costs to the export price, i.e.  $\beta$ , according to:

$$\begin{aligned} \beta_{ikt} &\equiv \frac{\partial TC_{ikt}}{\partial \tilde{p}_{ikt}} \frac{\tilde{p}_{ikt}}{TC_{ikt}} \\ &= \frac{\hat{t}_{ikt}/\tilde{p}_{ikt}}{\hat{t}_{ikt}/\tilde{p}_{ikt} + \hat{\tau}_{ikt} - 1}, \end{aligned}$$

making use of our first-stage estimates for both the additive and the ad-valorem components ( $\hat{t}_{ikt}$  and  $\hat{\tau}_{ikt}$ ). We can then rebuild the value of the elasticity on a per product-

year-country basis. The histogram of the distribution of the estimated  $\beta_{ikt}$  is reported in Figure ??, in panels (a) and (b) for air and maritime transport respectively.

As displayed in Figure ??, for both transport modes the distribution of  $\beta_{ikt}$  over the triplet (year, product, origin country) is smoothly distributed over the interval  $[0, 1]$ , with a mode of the distribution standing around 0.3-0.4 depending on the transport mode, consistently with our previous findings. This stands in sharp contrast with the assumption made by ? or ?, which would imply a distribution concentrated on a single point. Rather, this suggests that the elasticity of transport costs to the export price, or equivalently the share of additive costs in total costs, is varying across time, product and country partner.<sup>22</sup> In the next section, we shed light on the role of this result in the analysis of the underlying sources of the overall transport costs trend patterns, between trade composition effects and pure transport costs.

### 3.2 Empirical specification

Based on the above results, we provide a decomposition of the transport costs trends between the trade composition effects and the “pure” transport costs time trends, which explicitly takes into account the varying share of additive costs over time, country partner and product.<sup>23</sup> An important contribution of the paper is to show the crucial role of this (empirically relevant) assumption, as it notably modifies the conclusions relative to the role of trade composition effects. We start with the presentation of our estimation strategy before turning to the results.

#### 3.2.1 Estimation strategy

The estimation carried out in Section 2.3 provides us with the additive and ad-valorem measures of international transport costs (Equation (6)), which vary over time, product and origin country (i.e.,  $\hat{t}_{ikt}$  and  $\hat{\tau}_{ikt}$ ). Starting from these values, for each additive and multiplicative cost component (by transport mode) we extract the pure transport cost measure by the mean of a time fixed effect. Precisely, we extract the changes over time in the pure transport cost dimension by assuming a composition of trade flows by country partner and product that is constant throughout the period, and equal to the one observed in 1974. We conduct the same analysis on an overall transport costs measure, built by combining the two estimated components (additive and iceberg) in a unified measure of transport costs.<sup>24</sup> The objective is to obtain six time series, all built as indices with the reference value 100 in 1974: three time series for the unfitted transport costs measures  $\{\Gamma_t^{add,raw}, \Gamma_t^{adv,raw}, \Gamma_t^{tc,raw}\}$  (additive, ad-valorem and total cost respectively), and three series for the fitted values  $\{\Gamma_t^{add}, \Gamma_t^{adv}, \Gamma_t^{tc}\}$  (additive, ad-valorem and total cost respectively).

<sup>22</sup>We have checked that it is also the case when we look at the distribution over the range of origin countries or products, for a given year. These results are not reported for sake of space, but are available upon request to the authors.

<sup>23</sup>See Appendix D.2 for a detailed comparison of Hummel’s (2007) methodology and ours.

<sup>24</sup>Note that the unfitted total transport costs are virtually the same as those reported in Figure ??, but reported from another perspective (basis 100 in year 1974).



One advantage of this method is that it yields measures of transport costs (fitted and unfitted) that are easily comparable between transport modes and transport cost components. For each cost component (additive and ad-valorem) as well as for the total transport cost, comparing the unfitted measure and the fitted measure (composition effects excluded) makes it possible to identify if the decrease observed over the period in the unfitted series is due to trade composition effects (for instance, changes in the country partners, in the type/quality of products traded), or if it is the pure transport costs (for instance, insurance or handling costs) that have reduced over time. We now describe the method to extract these series (with more details provided in Appendix D).

### 3.2.2 Estimation method for each additive and multiplicative component

We begin by describing the estimation method to extract the fitted and unfitted series for both the additive and the multiplicative components, before turning to the ‘total costs’ series.

**Obtaining the “pure” transport costs component series** - The empirical strategy to find the fitted transport cost measure (i.e. extracting from trade composition effect) can be described as a two-stage process. First, we decompose the estimated measure in the three product/country/time dimensions, using fixed effects. For the estimated ad-valorem component, we estimate the following equation:

$$\ln(\widehat{\tau}_{ikt}) = \delta + \underbrace{\sum_{i \neq \text{ARG}} \alpha_i \cdot \mathbb{1}_i}_{(a)} + \underbrace{\sum_{s(k) \neq 011} \beta_{s(k)} \cdot \mathbb{1}_{s(k)}}_{(b)} + \underbrace{\sum_{t \neq 1974} \gamma_t \cdot \mathbb{1}_t + \epsilon_{ikt}}_{(c)} \quad (9)$$

where  $\mathbb{1}_i$  and  $\mathbb{1}_{s(k)}$  represent country and sector fixed effects.<sup>25</sup> Equation (9) is estimated using OLS, with a weighting scheme based on the value of each flow in the total value of flows for the relevant year. As for the additive component, given that the sector fixed effect and the country fixed effect are additive rather than multiplicative by construction, we estimate the following equation using non-linear least squares:<sup>26</sup>

$$\ln(\widehat{t}_{ikt}) = \ln \left( \delta + \underbrace{\sum_{i \neq \text{ARG}} \alpha_i \cdot \mathbb{1}_i}_{(a)} + \underbrace{\sum_{s(k)} \beta_{s(k)} \cdot \mathbb{1}_{s(k)}}_{(b)} \right) + \underbrace{\sum_{t \neq 1974} \gamma_t \cdot \mathbb{1}_t + \epsilon_{ikt}}_{(c)} \quad (10)$$

As displayed in Equations (9) and (10), we decompose the estimated transport cost com-

<sup>25</sup>Throughout the exercise, we consider Argentina, the sector 011 and the first year of our dataset 1974, as references for the country, product and year dummies.

<sup>26</sup>For sake of notational simplicity, we do not distinguish the coefficients associated with the fixed effects between Equations (9) and (10), even if they are specific to the type of transport costs considered (e.g. the series of  $\gamma_t$  differs from one estimation to the other). Note that we apply the same weighting scheme as for the OLS regression (based on the relative value of the flow).

ponent in three elements: the country dimension (Term (a)), the product dimension (Term (b)) and the pure transport costs time trend (Term (c)). Note that Equations (9) and (10) preserve our specification of the ad-valorem and additive costs of Equations (3) and (4), as we consider that the iceberg cost is the product of the country of origin and the goods dimension, while the additive cost is the sum of the two dimensions. Both equations are estimated by transport mode.

In this exercise, we are interested in isolating the change in the time dimension of each transport cost component. This constitutes the second stage of our procedure. As for the ad-valorem component defined in  $[1; +\infty]$ , from the estimation of Equation (9), we build the variable  $\Gamma_t^{adv}$ , for each year  $t \geq 1974$ , according to:

$$\Gamma_t^{adv} = 100 \cdot \frac{\bar{\tau}_{1974} \cdot \exp(\gamma_t) - 1}{\bar{\tau}_{1974} - 1} \quad (11)$$

with  $\bar{\tau}_{1974} = \exp(\delta + \sum_i \alpha_i + \sum_s \beta_s)$  the mean ad-valorem transport cost in 1974 (see details in Appendix D). In plain words, we measure how these costs have changed over time by blocking the composition of trade flows by product and country partners to the one observed in 1974 (the beginning of our sample).

As for the additive cost defined in  $[0; +\infty]$ , we build the variable  $\Gamma_t^{add}$ , the reference year being 1974 (ie, with  $\gamma_{1974} = 0$ ) according to:

$$\Gamma_t^{add} = 100 \cdot \exp(\gamma_t) \quad (12)$$

As a result, the two series  $\Gamma_t^{adv}$  and  $\Gamma_t^{add}$  have a straightforward interpretation in percentage changes from the initial value of 100 for  $t = 1974$ .

**Obtaining the unfitted measures** - The objective is to find the unfitted transport cost component (additive and multiplicative) built as an index with reference value 100 in 1974, starting from the estimated values previously obtained ( $\hat{\tau}_t$ ,  $\hat{t}_t$ , for each year  $t$  from 1974 to 2013). To do so, we apply the simple formula to get the following indices, for the ad-valorem and the additive cost components respectively:

$$\Gamma_t^{adv,raw} = 100 \times \frac{\hat{\tau}_t}{\hat{\tau}_{1974}}, \quad \Gamma_t^{add,raw} = 100 \times \frac{\hat{t}_t}{\hat{t}_{1974}}$$

### 3.2.3 Estimation method for the total transport cost measures

We also build two measures of the “total” transport costs, combining our estimates of the two additive and ad-valorem components, for both the unfitted series and the pure transport cost series (i.e. composition effects excluded). As for each additive and ad-valorem component, the objective is to get a measure of total transport cost changes built as an index starting from the value 100 in 1974. Even if obeying the same logic, we proceed slightly differently for the unfitted and the fitted measures, as we now explain.

**Obtaining the unfitted total transport cost index** - For each transport mode, we build the total transport cost series based on Equation (2) according to:

$$\widehat{tc}_t^{raw} = \widehat{\tau}_t^{adv} - 1 + \frac{\widehat{t}_t}{\widetilde{p}_t}$$

where  $\widehat{\tau}_t^{adv}$  and  $\widehat{t}_t$  are the average values over the country/sector dimension estimated (conditional on a year and a transport mode), as explained in Section 2. Recall that  $\widehat{\tau}_t^{adv} - 1$  measures the ad-valorem transport cost component and  $\frac{\widehat{t}_t}{\widetilde{p}_t}$  the additive component, both expressed as percentage of the FAS price. We then transform this value in an index with base year 1974, applying a similar formula as above (by transport mode):

$$\Gamma_t^{tc,raw} = 100 \frac{\widehat{tc}_t^{raw} - 1}{\widehat{tc}_{1974}^{raw} - 1}$$

**Obtaining the fitted total transport cost index** - The same logic as above applies to constructing the fitted measure of total transport cost (i.e. composition effect excluded), with one notable exception: Fitted ad-valorem and additive components have not been estimated in value, but extracted and built as indices. In a first step then, we rebuild the fitted measures of each transport cost component in value starting from these indices. For the ad-valorem component, this can be obtained by rewriting Equation (13):

$$\widehat{\tau}_t^{pure,adv} = \frac{\Gamma_t^{adv} (\bar{\tau}_{1974} - 1)}{100} + 1$$

with  $\widehat{\tau}_t^{pure,adv} \equiv \bar{\tau}_{1974} \cdot \exp(\gamma_t)$  the yearly value of the pure (fitted) ad-valorem cost. A similar reasoning starting from Equation (15) gives the fitted value for the additive cost component as:

$$\widehat{t}_t^{pure} = \frac{\Gamma_t^{add}}{100}$$

with  $\widehat{t}_t^{pure} \equiv \exp(\gamma_t)$  the yearly value of the pure (fitted) additive cost.

We then deduce the fitted value of the overall cost according to:

$$\widehat{tc}_t^{pure} = \widehat{\tau}_t^{pure,adv} + \widehat{t}_t^{pure}$$

Last, we transform this (fitted) measure of total transport cost in an index with base year 1974, applying a similar formula as above (by transport mode):

$$\Gamma_t^{tc} = 100 \frac{\widehat{tc}_t^{pure} - 1}{\widehat{tc}_{1974}^{pure} - 1}$$

### 3.3 Characterizing the time trends in transport costs

Figure ?? reports the results for all types of goods.<sup>27</sup> In panels (a), (b) and (c) we report the time changes of the ad-valorem costs, the additive costs and the total costs respectively, for air transport (starting from the reference value 100 in 1974). Panels (d), (e) and (f) report the results for vessel. In each panel, we report the evolution of transport costs for both the unfitted (plain blue line) and the fitted (dotted red line) measures.

Three main results emerge from Figure ?. First, in accordance with Figure ??, we find that international transport costs have substantially decreased over the period, for both transport modes (Panels (c) and (f)). International transport costs were reduced by 50% between 1974 and 2013 in air shipping, and by 60% in maritime shipping. This stands in line with the related literature (see ?; ?). Second, the magnitude of the decrease is roughly of the same order for both for the ad-valorem and the additive components (Panels (a), (b), (d) and (e)). Further, the transport cost reduction is much smoother in maritime transport, while air transport shows more volatility in the trend pattern, in particular in the 1980s and the 2000s. Third, and most importantly, we find that composition effects do not play a major role in accounting for the time trend of overall transport costs. Inspecting the panels of Figure ?? for air transport (panels (a) to (c)), we do not find much evidence of a substantial trend difference between the unfitted transport costs measure and the pure transport costs. For all three series (additive, ad-valorem and overall transport costs), the dotted and the plain line follow closely, almost every year throughout the period. Air transport costs were reduced by 50% between 1974 and 2013, which is mainly attributable to a reduction in the transport costs “per se”. Composition effects are stronger in vessel transport, for all three series (Figure ??, panels (d) to (f)): they amplify the reduction in “pure” transport costs at the beginning of the period, in the 1970s (the dotted line is below the plain line, and the gap is increasing). Afterwards, the dotted line remains below the plain line, but the gap between the two remains roughly constant across years, indicating that the amplification effect does not strengthen anymore: the two trends are similar. Considering the raw series (plain line), maritime transport costs have decreased by 60% over the period, which can be decomposed into a 50% decrease in transport costs “per se” (dotted line), and a 10% reduction that comes from composition effects (the difference of the two). This is particularly the case for the multiplicative component (panel (d)).

### 3.4 Time trends in transport costs and the modeling of the additive component

The last results stand in sharp contrast with ?, who finds that trade composition effects do matter, as they partially offset the reduction in the pure transport costs for both air and maritime shipping. If anything, we find the opposite result here: Trade composition effects do not matter much in accounting for the downward trend of observed international transport costs, and when they do, they tend to amplify (rather than reduce) the downward

---

<sup>27</sup>In the Online Appendix, we report the results at a more disaggregated level, distinguishing between primary and manufacturing goods.

pattern. This drives us to investigate this difference further. As we explain in more detail in Appendix D, our empirical strategy differs from ? in one main dimension. Our characterization of the time trends in transport costs starts from our estimates of both the additive and the ad-valorem components (obtained in Section 2.3), as well as for the overall transport cost (rather than the actual CIF-FAS price gap). Consequently, our methodology lets the ratio between the additive and the ad-valorem components of transport costs vary over the three time-product-partner country dimensions. This turns out to be of primary importance in the disentangling, in the time trend of transport costs, of what comes from the trade composition effects and what comes from changes in the pure transport cost dimension, as we show below.

To establish this point clearly, we replicate the method adopted by ? exposed above on our database (which is the same as his until 2004). The results are reported in Figure ???. The dotted line labeled “expenditure/import value” represents the unfitted measure of transport costs ( $TC_t$  in the above terminology) and the plain line labeled “fitted ad-valorem rate” is the measure of “pure” transport costs, i.e. composition effects excluded ( $\widehat{TC}_t$ ). In Panels (a) (for air) and (b) (for vessel), we report the transport costs measures (fitted and unfitted) expressed as a percentage of the export price, as in ?.<sup>28</sup> To ease the interpretation of the trends, we express them as indices with the reference value 100 in 1974, in Panels (c) and (d) for air and vessel respectively.

These results stand in sharp contrast with the ones obtained with our methodology and reported in Figure ??. Applying ?’ (?) method, we find that the composition effects tend to partly offset the decrease in transport costs in both air and vessel shipping, as the downward trend is of higher magnitude for the fitted rate (plain line) than the unfitted rate (dotted line), especially for vessel (Panel (d)). Yet, this result is overturned when we allow for more flexibility in the role of the additive component, as indicated in Figure ??, Panels (c) and (f).

Assuming a varying share of the additive component over time, product and country partner indeed modifies the decomposition of the trend reduction of transport costs between the one attributable to trade composition effects and the reduction in the “pure” transport costs. In both air and vessel transports, we thus find that this last dimension is the main driver of the reduction in international transport costs observed over time, in particular in air transport. Complementing the findings of Section 2.3, these results highlight the importance of integrating the additive dimension of international transport costs, with the aim here of characterizing their time trends.

## 4 Robustness analysis

### 4.1 Robustness to the separability assumption

In this section, we provide a robustness analysis to the separability assumption of country-good fixed effects, which considers (following ?) that both the ad-valorem and the additive

---

<sup>28</sup>Unsurprisingly, this accords with ?’ (?) results, see his Figures 5 (for air) and 6 (for vessel).

costs are separable between the origin country ( $i$ ) and the product ( $s(k)$ ) dimensions. That is, we relax Equations (3) and (4) to model  $t_{is(k)}$  and  $t_{is(k)}$  (on a yearly/transport mode basis). Yet, this comes at the cost of having to estimate a much larger number of fixed effects (for approx. 200 countries of origin and 200 sectors at the 3-digit level, this means estimating 40,000 versus 400 fixed effects, per year and per transport mode), making the estimation (all the more given our non-linear setting) intractable in practice. This drives us to conduct the analysis on a sub-sample of trade flows, upon which two estimations are conducted (for every year and each transport mode): with the separability assumption (Equations (3) and (4) hold, our baseline case) and without. For each year and transport mode, we select the countries ( $i$ ) and the products (at the 3-digit level,  $s(k)$ ) that constitute 80% of the total value of trade flows. Table 2 shows a summary of the results, considering the average estimated values over the period (by transport mode).

Table 2: Robustness to the separability assumption (average over the period) - restricted sample

| Mean value over 1974-2013                                   |         |                |          |                |
|---|---------|----------------|----------|----------------|
| 3 digits  |         |                |          |                |
| Mode  | Air     |                | Vessel   |                |
| Separability  | No      | Yes (baseline) | No       | Yes (baseline) |
| <i>Additive term (<math>\hat{t}/\tilde{p}</math>, in %)</i> |         |                |          |                |
| Mean  | 1.76    | 1.96           | 2.61     | 2.99           |
| Median  | 0.65    | 0.84           | 1.83     | 2.23           |
| <i>Ad-valorem term (<math>\hat{\tau}</math>, in %)</i>      |         |                |          |                |
| Mean  | 1.04    | 0.94           | 2.43     | 2.40           |
| Median  | 0.56    | 0.83           | 2.06     | 2.04           |
| <i>Share of additive costs in total costs (in %)</i>        |         |                |          |                |
| Mean  | 62.8    | 67.5           | 51.8     | 55.4           |
| Median  | 54.1    | 50.5           | 47.0     | 52.2           |
| <i>Diagnostic test</i>                                      |         |                |          |                |
| AIC criteria  | 4657.60 | 4654.23        | 10744.83 | 10690.80       |
| # observations  | 2381    | 2381           | 2798     | 2798           |
| # origin country  | 13.4    | 13.4           | 19.8     | 19.8           |
| # products  | 25.5    | 25.5           | 53       | 53             |

Notes: Estimations are performed on a restricted sample, based on the following rule: for each year and transport mode, we select the countries ( $i$ ) and the products (at the 3-digit level,  $s(k)$ ) that constitute 80% of the total value of trade flows. The additive term is expressed in fraction of FAS price. (\*): 1989 omitted in 3-digit estimation for air.

Three main comments can be made. First, the subset of large trade flows we are considering (around 8% of the observations of our initial sample, but 80% of the total value of aggregated trade) involve a higher share of additive costs: between 52% and 67% of total costs on average, versus 40% to 50% on the whole sample. Second, both models provide a similar quality of fit of the regression (once we take into account the number of right-hand-side variables) as the AIC criteria display very close values in both models (conditional on transport mode). Third, the estimated values remain of similar magnitude under both separability and non-separability assumptions, even if one can note that under the separability assumption, the share of additive costs seems slightly biased upward, at

least for averages over the period. The picture is reversed for medians.

Before jumping to the conclusion that our results are robust to the separability assumption, however we investigate this comparison deeper on a year-to-year basis, as shown in Figure ???. Specifically, Figure ??? displays the results for air transport in panel (a) and for maritime transport in panel (b). In each case, the estimated values (as well as the fitted regression line) under both separability (baseline) and non-separability are reported, for the additive term (upper panel) and the ad-valorem term (lower panel).

In line with Table 2, the results shown in Figure ??? confirm that the separability assumption (retained in our baseline estimation) tends to overestimate the value of the additive cost component (and underestimate the value of the ad-valorem cost). However, they also show that the difference is quantitatively small. Further, and most importantly, whatever the transport mode and for both types of transport costs, the trend patterns of international transport costs, whether estimated under the separability assumption or not, are very similar. Along with the quality-of-fit diagnosis, this confirms the robustness of our estimation results to this assumption.

As an ultimate check, we investigate whether the varying pattern over time/country of origin/product pointed out in Section 2.3 is also robust to the separability assumption. With this aim, we replicate Figure ??? on the reduced sample for both specifications of separable (baseline) and non-separable fixed costs. Figure ??? shows the histogram of the distribution of the elasticity of transport costs to the export price ( $\beta_{ikt}$ , for air transport (top panel) and for maritime transport (bottom panel) under the baseline specification (Panels (a), (c))) and the non-separability scenario (Panels (b), (c)).

Similarly, as for the whole sample (Figure ???), for both transport modes the distribution of  $\beta_{ikt}$  over the triplet (year, country of origin, product) appears smoothly distributed over the interval  $[0, 1]$ , whatever the assumption regarding the specification of fixed effects. Relaxing the assumption that fixed effects are separable between the two country and product dimensions does not alter qualitatively our conclusion, that the elasticity of transport costs to the export price (or the share of additive costs in total transport costs), is strongly varying over time, countries of origin and products. As a consequence, this also provides an indirect robustness check to the results regarding the trend decomposition provided in Section 3, which establish that the decline of “pure” transport costs is a main reason for the decrease of overall transport costs observed over time.

## 4.2 Robustness to the weighting scheme

As discussed in Appendix D, one source of difference between our methodology and that of ? concerns the weighting scheme used to obtain the evolution of the pure transport costs over time. In this section, we provide a robustness test to this difference of treatment.

When aggregating the trade cost measure over the product/country ( $i, k$ ) dimension, ? takes the unweighted average value over the  $i, k$  dimension, which implicitly attributes a weight equal to 1 to each flow. Our benchmark specification (at the root of Figure ???) proceeds differently, as each flow is weighted by its relative value in total trade flows observed in 1974. In this section, we check whether this difference of treatment is responsible or

not for the difference in results relative to the role of the trade composition effects, which we rather attribute to the modeling of the additive component.

To this aim, we rebuild the fitted values for each of our transport cost measures (ad-valorem, additive and overall) applying the ? weighting scheme (i.e. taking the unweighed average value over the  $i, k$  dimension), which we then express as indices, with the reference value 100 in 1974. The results are reported in Figure ??.

The comparison of Figure ?? (applying ?’s weighting scheme) and Figure ?? (applying our weighting scheme) prompts two comments. First, differences do show up, in particular for the multiplicative component of air transport. This suggests that the weighting scheme is not innocuous in the time trend decomposition exercise. However, Figure ?? also shows that the composition effects have contributed to strengthen the decrease in the “ceteris paribus” transport costs in air transport (in Figure ??, panel (a), overall transport costs have decreased more than the fitted component), in accordance with the conclusions drawn from Figure ??, rather than partly offsetting this decline, as argued by ?. This confirms that it is the functional form and precisely the modeling of the additive component, rather than the weighting scheme, that is key in understanding the underlying determinants of the overall transport costs time trends, as pointed out in Section 3.

## 5 Conclusion

This paper empirically studies the magnitude of additive costs in international transport costs, by exploiting the differences between the import and the export prices. Using SITC 3 and 4- digit CIF-FAS unit values taken from the US import database over 1974-2013, we estimate the two components (additive and ad-valorem) of transport costs, by transport mode (air or ocean). Two main findings emerge from our results. First, we provide a quantitative measure of both the additive and the ad-valorem transport costs. We find that additive costs are quantitatively sizable: They amount to 2.8% of the export price unit values for ocean shipping, and 1.8% for air transport. This is slightly lower than the estimated values for the ad-valorem costs, around 2.5% (vessel) and 3.2% (air) as mean values over 1974-2013. Further, modeling the additive component significantly improves the quality of fit of the regression. Second, we show the importance of integrating the additive component in accounting for the time trend of international transport costs. Allowing for a varying share of additive costs in the product/country/time dimension, we find that the decrease of international transport costs observed in the data is mostly attributable to a reduction in the *pure* transport costs, with trade pattern composition effects playing a small role. If anything, trade composition effects have contributed to amplifying the reduction in the pure transport costs in maritime transport. In both aspects, our results highlight the importance of the additive component in accounting for international transport costs.

Our results could be extended in two main ways. On the empirical side, one may want to go deeper into the structural determinants of (pure) transport costs - i.e. identify the respective roles of handling costs, insurance and freight at the root of the gap between



export and import prices. On the theoretical side, our results can be used to explore the role of additive costs in shaping international trade flows in an international trade theory perspective and in affecting the international transmission of business cycles. This is left for further research.

## A Data Appendix

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 this respect, the “custom value” corresponds to the FAS price (“free-alongside” price) delivered by the seller. More information on this database is available at: [http://www.census.gov/foreign-trade/reference/products/catalog/fl\\_imp.txt](http://www.census.gov/foreign-trade/reference/products/catalog/fl_imp.txt).

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.

The cif (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.

## B Estimation at the 3-digit classification level

### B.1 Transport costs estimates: More detailed results

In this section, we report more detailed results for the estimates for international transport costs, by transport mode on a yearly basis, when either additive costs are included in the estimation (Equation (6)) or not (Equation (7)), under our benchmark sectoral classification level (3 digit). Precisely, we complement the results displayed in Table 1 by reporting the estimates of international transport costs for a sample of years over 1974-2013, when the degree of classification retained ( $s$ ) is at the 3-digit classification level. Table B.1 reports the results for air transport. The results for vessel transport are displayed in Table B.2.

Table B.1: Air: Transport costs estimates, 3-digits (selected years)

| Year   | 1974   | 1980   | 1990   | 2000   | 2004   | 2010   | 2013   |
|--|--------|--------|--------|--------|--------|--------|--------|
| <b>Model (a) - With only Ad-Valorem TC (<math>\hat{\tau}^{ice} - 1</math>, in %)</b> |        |        |        |        |        |        |        |
| Mean   | 6.9    | 5.4    | 5.0    | 3.6    | 4.0    | 4.2    | 3.4    |
| Median   | 5.4    | 3.8    | 4.4    | 2.5    | 2.9    | 3.4    | 2.9    |
| <b>Model (b) - With Additive &amp; Ad-Valorem TC</b>                                 |        |        |        |        |        |        |        |
| <i>Ad-valorem term (<math>\hat{\tau}^{adv} - 1</math>, in %)</i>                     |        |        |        |        |        |        |        |
| Mean   | 3.6    | 2.3    | 2.4    | 1.7    | 1.9    | 2.6    | 1.7    |
| Median   | 2.7    | 1.6    | 1.6    | 1.2    | 1.4    | 2.2    | 1.7    |
| <i>Additive term (<math>\hat{t}/\tilde{p}</math>, in %)</i>                          |        |        |        |        |        |        |        |
| Mean   | 2.6    | 2.0    | 1.8    | 1.3    | 1.5    | 1.1    | 1.0    |
| Median   | 1.1    | 0.5    | 0.8    | 0.5    | 0.6    | 0.4    | 0.5    |
| # observations   | 14,955 | 16,118 | 24,958 | 35,027 | 36,990 | 40,279 | 39,351 |

Notes: TC = Transport Costs. Statistics are obtained weighting each observation by its share in trade (mode-dependent). Additive term expressed in fraction of FAS price.

Table B.2: Vessel: Transport costs estimates, 3 digit (selected years)

| Year   | 1974   | 1980   | 1990   | 2000   | 2004   | 2010   | 2013   |
|--|--------|--------|--------|--------|--------|--------|--------|
| <b>Model (a) - With only Ad-Valorem TC (<math>\hat{\tau}^{ice} - 1</math>, in %)</b> |        |        |        |        |        |        |        |
| Mean   | 9.8    | 6.5    | 5.7    | 5.1    | 5.4    | 4.0    | 3.6    |
| Median   | 9.6    | 5.5    | 4.6    | 4.9    | 5.1    | 3.6    | 3.3    |
| <b>Model (b) - With Additive &amp; Ad-Valorem TC</b>                                 |        |        |        |        |        |        |        |
| <i>Ad-valorem term (<math>\hat{\tau}^{adv} - 1</math>, in %)</i>                     |        |        |        |        |        |        |        |
| Mean   | 5.4    | 3.1    | 3.3    | 2.5    | 2.7    | 1.9    | 2.2    |
| Median   | 4.9    | 2.4    | 2.8    | 2.1    | 2.8    | 1.8    | 1.8    |
| <i>Additive term (<math>\hat{t}^{add}/\tilde{p}</math>, in %)</i>                    |        |        |        |        |        |        |        |
| Mean   | 5.1    | 3.4    | 2.7    | 2.8    | 2.9    | 2.5    | 1.5    |
| Median   | 2.9    | 2.3    | 1.7    | 2.2    | 1.9    | 1.9    | 0.8    |
| # observations   | 19,007 | 17,356 | 28,383 | 36,090 | 37,757 | 37,748 | 38,473 |

Notes: TC = Transport Costs. Statistics are obtained weighting each observation by its share in trade (mode-dependent). Additive term expressed in fraction of FAS price.

## B.2 Assessing the importance of additive transport costs

In this section, we explore the performances of each type of model (with and without additive costs) in fitting the observed CIF-FAS prices gap, in order to deliver a more systematic diagnosis about the importance of additive costs. To do so, we rely on several standard measures of fit. The first indicator is through comparing  $R^2$ . However, its use is far from being straightforward when evaluating non-linear estimates.  $R^2$  is based on the underlying assumption that the adjusted model is a linear one. In a non-linear context,  $R^2$  is strictly speaking inappropriate. However, if the error distribution is approximately normal, a standard metric like  $R^2$  remains informative on the quality of adjustment. This drives us to complement the goodness of fit diagnosis with three alternative measures. We provide the Standard Error of Regression (SER), which represents the average distance that the observed values fall from the regression line. The smaller the SER value, the better the quality of fit, as 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.<sup>29</sup> 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 the additive term, i.e. Equation (6)) and the Restricted one ( $R$ , i.e. Equation (7)). The null tested is that the two models are statistically equivalent. Results are reported in Tables B.3 and B.4, for air and vessel respectively, at the 3-digit level.

Table B.3: Air: Measures of Goodness-of-fit (3 digits)

| Year                  | 1974     | 1980     | 1990     | 2000     | 2010     | 2013     | Mean stat |
|-----------------------|----------|----------|----------|----------|----------|----------|-----------|
| $R^2$                 |          |          |          |          |          |          |           |
| Model (A)             | 0.30     | 0.27     | 0.25     | 0.32     | 0.42     | 0.34     | 0.31      |
| Model (B)             | 0.59     | 0.65     | 0.63     | 0.64     | 0.51     | 0.46     | 0.60      |
| <b>SER</b>            |          |          |          |          |          |          |           |
| Model (A)             | 0.79     | 0.86     | 0.81     | 0.84     | 0.86     | 0.92     | 0.85      |
| Model (B)             | 0.67     | 0.71     | 0.67     | 0.70     | 0.79     | 0.85     | 0.73      |
| <b>AIC criteria</b>   |          |          |          |          |          |          |           |
| Model (A)             | 35675.0  | 41171.0  | 60715.6  | 87492.6  | 102297.7 | 88191.9  | 70498.1   |
| Model (B)             | 31387.3  | 35738.4  | 52098.9  | 74954.9  | 95887.1  | 80873.7  | 62285.0   |
| <b>Log-likelihood</b> |          |          |          |          |          |          |           |
| Model (A)             | -17530.5 | -20253.5 | -29977.8 | -43341.3 | -50746.8 | -43692.9 | -34888.6  |
| Model (B)             | -15125.6 | -17263.2 | -25393.5 | -36788.4 | -47277.5 | -39751.9 | -30508.3  |
| LL ratio              | 4809.7   | 5980.6   | 9168.7   | 13105.7  | 6938.6   | 7882.1   | 8760.69   |
| nb of restrictions    | 355      | 369      | 393      | 426      | 426      | 427      | 402       |
| p-value               | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00      |

Notes: Model (A) = with only ad-valorem transport costs. Model (B) = with additive & ad-valorem transport costs. SER = Standard Error of regression; AIC = Akaike Information Criterion.  $R^2$  between the log of predicted ratio and the log of the observed ratio. For the LL ratio test, the number of restrictions is equal to the number of parameters estimated, i.e. the number of partner countries plus the number of products. The mean statistics calculated as the average value over all years.

<sup>29</sup>Precisely, the AIC stat is equal to  $2 \times \text{number of parameters} - 2 \times \text{Likelihood}$ , the number of parameters being given by the number of restrictions.

Table B.4: Vessel: Measures of Goodness-of-fit (3 digits)

| Year                  | 1974     | 1980     | 1990     | 2000     | 2010     | 2013     | Mean stat |
|-----------------------|----------|----------|----------|----------|----------|----------|-----------|
| $R^2$                 |          |          |          |          |          |          |           |
| Model (A)             | 0.45     | 0.42     | 0.46     | 0.40     | 0.35     | 0.34     | 0.39      |
| Model (B)             | 0.61     | 0.58     | 0.59     | 0.57     | 0.49     | 0.46     | 0.56      |
| <b>SER</b>            |          |          |          |          |          |          |           |
| Model (A)             | 0.58     | 0.62     | 0.59     | 0.65     | 0.74     | 0.76     | 0.66      |
| Model (B)             | 0.48     | 0.53     | 0.51     | 0.55     | 0.66     | 0.68     | 0.57      |
| <b>AIC criteria</b>   |          |          |          |          |          |          |           |
| Model (A)             | 33328.8  | 33010.3  | 51142.6  | 71365.9  | 84789.9  | 88191.9  | 57848.6   |
| Model (B)             | 27331.5  | 28067.3  | 43664.7  | 60475.9  | 76161.3  | 80873.7  | 49682.3   |
| <b>Log-likelihood</b> |          |          |          |          |          |          |           |
| Model (A)             | -16287.4 | -16129.1 | -25169.3 | -35263.9 | -41998.9 | -43692.9 | -28534.3  |
| Model (B)             | -12985.8 | -13353.7 | -21171.4 | -29491.0 | -37418.7 | -39751.9 | -24151.3  |
| LL ratio              | 6603.28  | 5550.96  | 7995.88  | 11545.98 | 9160.56  | 7882.15  | 8766.0    |
| nb of restrictions    | 393      | 395      | 411      | 436      | 424      | 427      | 417       |
| p-value               | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00      |

Notes: Model (A) = with only ad-valorem transport costs. Model (B) = with additive & ad-valorem transport costs. SER = Standard Error of regression; AIC = Akaike Information Criterion.  $R^2$  between the log of predicted ratio and the log of the observed ratio. For the LL ratio test, the number of restrictions is equal to the number of parameters estimated, i.e. the number of partner countries plus the number of products. The mean statistics calculated as the average value over all years.

Tables B.3 and B.4 lead to the same conclusion: The inclusion of the additive term leads to an improvement of the quality of fit, whatever the considered criterion or the transport mode. On average over the whole period, the  $R^2$  doubles when per-kg costs are included for air, and increases by 50% for vessel. Similar qualitative conclusions arise from the comparisons of the standard errors of the regression (SER). 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 log-likelihood test unambiguously rejects the null of statistical equivalence of the two models. These results holds whatever the considered year.

For comparison purposes, we provide a similar goodness-of-fit exercise at the 4-digit product level (4-digits), reported in in Appendix C, Tables C.3 and C.4. If anything, the quality of fit appears slightly higher when estimations are based on the 4-digit classification. This is especially true for the model restricting transport cost to their ad-valorem dimension, whatever the transport mode considered. When the additive part is taken into account however, the difference in goodness of fit between the 3- and the 4-digit classification level becomes very small, whatever the considered criterion. In other words, if using a more disaggregated classification unsurprisingly adds some statistical precision, this is not to an extent that would disqualify the use of slightly more aggregated data. Further, the same conclusion established at the 3-digit level regarding the significant role of the additive component in fitting international transport costs emerges at the 4-digit level.

### B.3 Variance decomposition exercise

In this section, we provide a variance decomposition exercise on the observed CIF-FAS price gap. Precisely, we determine the share of the observed variance in the ratio  $\ln(\frac{p_{ik}}{p_{ik}} - 1)$

that comes from *i*) the between-product variance (at the 5-digit level, *k*), *ii*) the between-sector variance (at the 3-digit level, *s*). The variance decomposition expression at the product level *k* (5-digit level) is obtained by applying the following formula (by year and transport mode):

$$\underbrace{\sum_{k=1}^K \sum_{i=1}^I (x_{ik} - \bar{x}_g)^2}_{\text{Total variability}} = \underbrace{\sum_k \sum_i (x_{ik} - \bar{x}_k)^2}_{\text{Within-product variability}} + \underbrace{\sum_k (\bar{x}_k - \bar{x}_g)^2}_{\text{Between-product variability}}$$

with  $x$  the observed CIF-FAS price gap and  $\bar{x}_g, \bar{x}_k$  average values defined as:

$$\bar{x}_g \equiv \frac{1}{n} \sum_{k=1}^K \sum_{i=1}^I x_{ik}, \quad \bar{x}_k \equiv \frac{1}{K} \sum_{k=1}^K x_{ik}$$

with  $n$  the total number of observations,  $K$  the total number of products and  $I$  the total number of country partners. We apply the same variance decomposition exercise at the sector level, in which case the sector  $s$  index (at the 3-digit level) replaces the  $k$  index (at the 5-digit level). This gives us an alternative way to ensure the robustness of the estimation results to the degree of classification retained to estimate international transport costs. We also determine the share of the observed variance that can be attributed to the between-country variance, adapting the variance decomposition formula written above accordingly. Results are reported in Figure ??.

Two interesting results emerge from Figure ?. First, the share of the CIF-FAS price gap variance that comes from the variance between products (5-digit level) is of same magnitude of order at the variance between sectors at the 3-digit level. Both account for between 30 and 40% of the total variance in air transport, depending on the years considered. This is also the case for vessel transport, even if the difference between the between-product variance share and the between-sector share is more pronounced (30% for the between-sector vs 40% for the between-product variance at the beginning of the period). This delivers an indirect robustness check to the degree of classification we have retained to estimate international transport costs. Second, the variance of the CIF-FAS price gap that can be attributed to the product (or sector) dimension is much larger than the between-country variance. This holds throughout the period and for both transport modes. This suggests that what primarily matters in international transport costs is mostly attributable to the product *per se*, rather than to the country where it comes from.

## C Estimation at the 4-digit level

In this section, we report the estimation results when we retain the 4-digit classification level ( $s=4$ -digit).

### C.1 Transport cost estimates

Tables C.1 and C.2 report the estimates of both models (with and without additive costs) in air and Ocean transport respectively.

Table C.1: Air: Transport costs estimates, Selected years, 4-digit

| Year   | 1974  | 1981  | 1989  | 2001  | 2009  | 2013  |
|--|-------|-------|-------|-------|-------|-------|
| <b>Model (a) - With only Ad-Valorem TC (<math>\widehat{\tau}^{ice} - 1</math>, in %)</b> |       |       |       |       |       |       |
| Mean   | 6.6   | 5.8   | 5.2   | 3.3   | 3.7   | 3.2   |
| Median   | 5.2   | 4.4   | 4.1   | 2.1   | 2.7   | 2.6   |
| <b>Model (b) - With Additive &amp; Ad-Valorem TC</b>                                     |       |       |       |       |       |       |
| <i>Ad-valorem term (<math>\widehat{\tau}^{adv} - 1</math>, in %)</i>                     |       |       |       |       |       |       |
| Mean   | 3.5   | 2.6   | 3.1   | 1.5   | 2.1   | 1.6   |
| Median   | 2.5   | 1.7   | 1.9   | 1.0   | 1.7   | 1.4   |
| <i>Additive term (<math>\widehat{t}/\widehat{p}</math>), in %</i>                        |       |       |       |       |       |       |
| Mean   | 2.6   | 2.1   | 1.7   | 1.2   | 1.2   | 1.0   |
| Median   | 1.2   | 0.6   | 0.6   | 0.5   | 0.4   | 0.4   |
| # observations   | 14944 | 16844 | 25307 | 35005 | 38475 | 39460 |

Notes: TC = Transport Costs. Statistics are obtained weighting each observation by its share in trade (mode-dependent). Additive term expressed in fraction of FAS price.

Table C.2: Vessel: Transport costs estimates, Selected years, 4-digit

| Year   | 1974  | 1981  | 1989  | 2001  | 2009  | 2013  |
|--|-------|-------|-------|-------|-------|-------|
| <b>Model (a) - With only Ad-Valorem TC (<math>\widehat{\tau}^{ice} - 1</math>, in %)</b> |       |       |       |       |       |       |
| Mean   | 9.8   | 6.1   | 5.8   | 5.1   | 4.2   | 3.6   |
| Median   | 9.4   | 5.1   | 4.8   | 4.5   | 3.8   | 3.1   |
| <b>Model (b) - With Additive &amp; Ad-Valorem TC</b>                                     |       |       |       |       |       |       |
| <i>Ad-valorem term (<math>\widehat{\tau}^{adv} - 1</math>, in %)</i>                     |       |       |       |       |       |       |
| Mean   | 5.4   | 3.4   | 2.8   | 2.8   | 2.4   | 2.1   |
| Median   | 4.9   | 3.0   | 2.4   | 2.5   | 2.6   | 1.8   |
| <i>Additive term (<math>\widehat{t}^{add}/\widehat{p}</math>), in %</i>                  |       |       |       |       |       |       |
| Mean   | 4.6   | 2.6   | 3.1   | 2.4   | 2.1   | 1.5   |
| Median   | 2.9   | 1.3   | 1.9   | 1.5   | 1.3   | 0.8   |
| # observations   | 19196 | 17916 | 29387 | 36677 | 37643 | 38820 |

Notes: TC = Transport Costs. Statistics are obtained weighting each observation by its share in trade (mode-dependent). Additive term expressed in fraction of FAS price.

### C.2 Goodness-of-fit tests at the 4-digit level

We now report the goodness-of-fit exercise (conducted by transport mode) at the 4-digit product classification level (for the selected years). The results are reported in Tables C.3 (for air) and C.4 (for vessel).

Table C.3: Air: Measures of Goodness-of-fit, 4-digits

|                       | Year     |          |          |          |           |          |
|-----------------------|----------|----------|----------|----------|-----------|----------|
|                       | 1974     | 1981     | 1989     | 2001     | 2009      | 2013     |
| <b>R<sup>2</sup></b>  |          |          |          |          |           |          |
| Model (a)             | 0.48     | 0.49     | 0.50     | 0.50     | 0.45      | 0.35     |
| Model (b)             | 0.63     | 0.66     | 0.65     | 0.66     | 0.54      | 0.45     |
| <b>SER</b>            |          |          |          |          |           |          |
| Model (a)             | 0.8      | 0.9      | 0.83     | 0.87     | 0.88      | 0.93     |
| Model (b)             | 0.67     | 0.74     | 0.69     | 0.80     | 0.80      | 0.86     |
| <b>Log-likelihood</b> |          |          |          |          |           |          |
| Model (a)             | -17505.6 | -21813.5 | -30960.6 | -44067.6 | -49375.6  | -53197.9 |
| Model (b)             | -14895.8 | -18589.9 | -26553.5 | -37297.9 | -45747.6  | -49899.1 |
| <b>AIC criteria</b>   |          |          |          |          |           |          |
| Model (a)             | 36243.1  | 44966.9  | 63417.1  | 89747.2  | 100317.13 | 107963.7 |
| Model (b)             | 31873.6  | 39495.8  | 55777.1  | 77439.9  | 94059.1   | 102224.3 |
| <b>Test LL</b>        |          |          |          |          |           |          |
| 2×(ll(UR) -ll(R))     | 5219.5   | 6447.1   | 8814.1   | 13539.4  | 7256.0    | 6597.5   |
| # restrictions        | 640      | 698      | 778      | 833      | 824       | 818      |
| p-value               | 0.00     | 0.000    | 0.00     | 0.00     | 0.00      | 0.000    |

Notes: Model (a) = with only ad-valorem transport costs. Model (b) = with additive & ad-valorem transport costs.  $R^2$  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 C.4: Vessel: Measures of Goodness-of-fit, 4-digits

|                       | Year      |          |          |          |          |          |
|-----------------------|-----------|----------|----------|----------|----------|----------|
|                       | 1974      | 1981     | 1989     | 2001     | 2009     | 2013     |
| <b>R<sup>2</sup></b>  |           |          |          |          |          |          |
| Term I only           | 0.50      | 0.45     | 0.47     | 0.41     | 0.37     | 0.35     |
| Terms A & I           | 0.66      | 0.62     | 0.62     | 0.58     | 0.51     | 0.46     |
| <b>SER</b>            |           |          |          |          |          |          |
| Model (a)             | 0.58      | 0.64     | 0.61     | 0.0.72   | 0.79     | 0.82     |
| Model (b)             | 0.48      | 0.53     | 0.51     | 0.61     | 0.69     | 0.75     |
| <b>Log-likelihood</b> |           |          |          |          |          |          |
| Model (a)             | -16460.1  | -16951.6 | -26771.4 | -39008.3 | -43888.9 | -47161.6 |
| Model (b)             | -12743.65 | -13546.9 | -21752.8 | -33281.0 | -39078.9 | -43399.2 |
| <b>AIC criteria</b>   |           |          |          |          |          |          |
| Model (a)             | 34464.2   | 35491.2  | 55272.9  | 79800.7  | 89459.8  | 95987.2  |
| Model (b)             | 28271.3   | 29877.8  | 46595.6  | 69743.9  | 81155.7  | 89692.4  |
| <b>Test LL</b>        |           |          |          |          |          |          |
| 2×(ll(UR) -ll(R))     | 12385.80  | 11226.8  | 17354.7  | 20113.5  | 16608.2  | 12589.6  |
| # restrictions        | 797       | 814      | 881      | 910      | 886      | 874      |
| p-value               | 0.000     | 0.000    | 0.000    | 0.000    | 0.000    | 0.000    |

Notes: Model (a) = with only ad-valorem transport costs. Model (b) = with additive & ad-valorem transport costs.  $R^2$  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.



## D Eliminating the composition effects: More details

In this section, we explain in more details the method employed to eliminate the country- and product- dimensions of the estimated transport cost.

### D.1 More on our methodology

#### D.1.1 Additive and ad-valorem transport costs: Excluding the composition effects

In this section, we detail our methodology to extract the time trend in the pure (or fitted) transport cost component, for each the ad-valorem and the additive component.

**For the ad-valorem component** Consider first the multiplicative transport cost component. Rewriting Equation (9) by taking the exponential, we get:

$$\hat{\tau}_{ikt} = \exp \left( \delta + \sum_{i \neq \text{AFG}} \alpha_i \cdot \mathbb{1}_i + \sum_{k \neq 011} \beta_k \cdot \mathbb{1}_k \right) \cdot \exp \left( \sum_{t \neq 1974} \gamma_t \cdot \mathbb{1}_t \right) \cdot \exp(\epsilon_{ikt})$$

Based on this equation, we deduce after estimation that:

$$\begin{cases} \text{For the year 1974:} & \hat{\tau}_{is74} = \exp(\delta + \alpha_i + \beta_s), \\ \text{For any year } t > 1974: & \hat{\tau}_{ist} = \exp(\delta + \alpha_i + \beta_s) \times \exp(\gamma_t) \end{cases}$$

From this, we obtain the following recursive link:  $\hat{\tau}_{ist} = \hat{\tau}_{is74} \exp(\gamma_t)$ . Given that  $\tau > 1$ , we can rewrite to get the percentage change between year 1974 and any year  $t > 1974$ :

$$\Gamma_{ist} = 100 \cdot \frac{\hat{\tau}_{ist} - 1}{\hat{\tau}_{is74} - 1} = 100 \cdot \frac{\hat{\tau}_{is74} \exp(\gamma_t) - 1}{\hat{\tau}_{is74} - 1}$$

As such, the index of transport costs in year  $t$  (relative to the reference year 1974)  $\Gamma_{ist}$  only depends on the cost observed in 1974 and the time trend. At this stage though, it remains specific to a product-origin country pair. Next step is to build the index  $\Gamma_t^{adv}$  such that:

$$\Gamma_t^{adv} = 100 \frac{\bar{\tau}_{1974} \cdot \exp(\gamma_t) - 1}{\bar{\tau}_{1974} - 1} \quad (13)$$

that is, Equation (11) with  $\bar{\tau}_{1974} = \exp(\delta + \sum_i \alpha_i + \sum_k \beta_k)$  the mean (ad-valorem) transport cost in 1974.

**For the additive component** After estimating Equation (10), we can re-build the additive component according to:

$$\begin{cases} \text{For the year 1974:} & \hat{t}_{is74} = \delta + \alpha_i + \beta_s, \\ \text{For any year } t > 1974: & \hat{t}_{ist} = (\delta + \alpha_i + \beta_s) \cdot \exp(\gamma_t) \end{cases}$$

From this, we deduce the recursive link:  $\hat{t}_{ist} = \hat{t}_{is74} \times \exp(\gamma_t)$ . Given the constraint  $t > 0$ , we then obtain the percentage change from 1974 from:

$$\Gamma_{ist}^{add} = 100 \frac{\hat{t}_{ist}}{\hat{t}_{ik74}} = 100 \exp(\gamma_t) \quad (14)$$

Note that it is independent of the product-origin country pair, we can thus rewrite the time-trend series for the additive transport cost component as:

$$\Gamma_t^{add} = 100 \exp(\gamma_t) \quad (15)$$

## D.2 Comparing with ?

In this section, we investigate the difference of results with ? regarding the importance of the composition effects in characterizing the time trends of international transport costs. In order to do so, we start from the empirical specification implemented by ? (also making use of Hummels' Stata codes provided on his webpage<sup>30</sup>), to better identify the precise points of difference between our estimation strategies.

Let us start from Hummels' (2007) quotation (p. 146) according to which (for air), the “*unadjusted measure of ad valorem air shipping costs [is] the aggregate expenditures on air shipping divided by the value of airborne imports*”. From this, we get the “raw” ad-valorem measure of transport costs as the ratio between the imported total value and the exported total value ( $(p_{ikt} - \tilde{p}_{ikt})q_{ikt}/\tilde{p}_{ikt}q_{ikt}$ ), the mean yearly value being obtained by weighting each flow by its value in total trade flows (ie, yielding the observed aggregate value of  $\tau_t$  for air).

Consider now the “fitted ad-valorem rate”. ? uses “*a regression in which the dependent variable is the ad-valorem air freight cost in logs for commodity  $k$  shipped from exporter  $i$ <sup>31</sup> at time  $t$ . The independent variables include a separate intercept for each exporter-commodity shipped, the weight/value ratio in logs for each shipment, and year dummy variables*”. With the value of the shipment equal to  $\tilde{p}_{ikt}q_{ikt}$ , and denoting the air freight cost as  $TC_{ikt}$ , we can formulate the above sentence according to the following equation (suppressing the transport mode index to alleviate notation):

$$\begin{aligned} \ln TC_{ikt} &= \delta + \beta \ln \frac{q_{ikt}}{\tilde{p}_{ikt}q_{ikt}} + \sum_{i,k} \alpha_{ik} \cdot \mathbb{1}_{ik} + \sum_t \gamma_t \cdot \mathbb{1}_t + \epsilon_{ikt} \\ \Leftrightarrow \ln TC_{ikt} &= \delta + \beta \ln \frac{1}{\tilde{p}_{ikt}} + \sum_{i,k} \alpha_{ik} \cdot \mathbb{1}_{ik} + \sum_t \gamma_t \cdot \mathbb{1}_t + \epsilon_{ikt} \end{aligned} \quad (16)$$

where  $TC_{ikt}$  is measured by the ratio between (air) charges, i.e.  $(p_{ikt} - \tilde{p}_{ikt})q_{ikt}^{fas}$  and the value of the shipment, i.e.  $\tilde{p}_{ikt}q_{ikt}^{fas}$ , or equivalently :  $TC_{ikt} = \frac{p_{ikt} - \tilde{p}_{ikt}}{\tilde{p}_{ikt}}$ .

After estimating Equation (16), ? uses the predicted value of the regression, denoted

<sup>30</sup><http://www.krannert.purdue.edu/faculty/hummelsd/research/jep-transport-cost-data.php>

<sup>31</sup>To be consistent with our notations, we change the country subscript from  $j$  (Hummels' (2007) terminology) to  $i$  (our terminology).

$\widehat{TC}_{ikt}$ , to obtain the unweighed average in the product/origin country dimension  $i, k$ , that is  $\widehat{TC}_t$ , to be compared to the unfitted ad-valorem rate  $TC_t$  (by transport mode).

What are the main differences with our estimation strategy? Five points may be underlined. First, ? obtains the fitted ad-valorem rate in one step (considering the observed export-import price gap on the left-hand side of Equation (16)), while we use our two-step approach to extract the fitted transport cost (ie, composition effect excluded) from the (already estimated) unfitted rate. However, this is a slight difference, as in both cases it ultimately amounts taking the predicted value of Equation (16) (which eliminates the changes specific to the origin country/product/year triplet). Second, in contrast to us, ? does not purge its measure of the fitted ad-valorem rate by the country-sector fixed effects. This is not likely to make a substantial difference though, as these fixed effects are by nature constant over time. Accordingly, they only play as a scale effect as the estimate is in log.

A third difference is that we separate the country-product fixed effects ( $\sum_{i,k} \alpha_{ik} \cdot \mathbb{1}_{ik}$  in Equation (16)) in two separate components (origin country and product dimensions). As discussed in Section 4, this is constrained by the number of fixed effects to estimate. However, we do not view this as the major cause of difference between our two methods, as it can be inferred from the robustness analysis on this point in Section 4.

Fourth, we differ in the weighting scheme retained to obtain the yearly value of the fitted transport cost (in the terms of ?' (?) method, the switch from  $\widehat{TC}_{ikt}$  to  $\widehat{TC}_t$ , Equations (11) and (12) in our case). ? takes the unweighed average value over the  $i, k$  dimension, which implicitly attributes a weight equal to 1 to each flow. We proceed differently, as we weight each flow by its relative value on total trade flows observed in 1974. This choice is made for two reasons. First, this weighting scheme does not overweight the small flows in value. Second, considering the 1974 weighting scheme makes sense as starting point of our time trend analysis. However, we ensure that our results are not sensitive to an alternative weighting scheme, by building the average fitted transport costs rate as in ?, as reported in Section 4.

One last difference remains to be commented, which relies on the way the additive component of international transport costs is treated. As we show below, it turns out to be key in accounting for the difference of results with ?. Coming back to Equation (16), it can be shown that it encompasses the two extreme cases of only ad-valorem costs ( $\beta = 0$ ) and only additive costs ( $\beta = 1$ ), as also pointed out in ?. To see it clearly, rewrite Equation (16) in a simpler form as:

$$\ln TC_{ikt} \equiv \ln \left( \frac{p_{ikt} - \tilde{p}_{ikt}}{\tilde{p}_{ikt}} \right) = \Delta_{ikt} - \beta \ln \tilde{p}_{ikt}$$

with  $\Delta_{ikt} \equiv \sum_{i,k} \alpha_{ik} \cdot \mathbb{1}_{ik} + \sum_t \gamma_t \cdot \mathbb{1}_t$  agglomerates the various fixed effects for reading clarity.

Under the first polar case with  $\beta = 1$ , it becomes:

$$\begin{aligned} \ln \left( \frac{p_{ikt} - \tilde{p}_{ikt}}{\tilde{p}_{ikt}} \right) &= \Delta_{ikt} - \ln \tilde{p}_{ikt} \\ \Leftrightarrow \ln(p_{ikt} - \tilde{p}_{ikt}) &= \Delta_{ikt} \end{aligned}$$

Equivalently, we can write:

$$\begin{aligned} p_{ikt} &= \tilde{p}_{ikt} + \exp(\Delta_{ikt}) \\ \Leftrightarrow p_{ikt} &= \tilde{p}_{ikt} + t_{ikt} \end{aligned}$$

with  $t_{ikt}$  appropriately defined. This exactly corresponds to the case of additive costs only.

In the second polar case with  $\beta = 0$ , Equation (16) becomes:

$$\begin{aligned} \ln \left( \frac{p_{ikt} - \tilde{p}_{ikt}}{\tilde{p}_{ikt}} \right) &= \Delta_{ikt} \\ \Leftrightarrow \ln \left( \frac{p_{ikt}}{\tilde{p}_{ikt}} - 1 \right) &= \Delta_{ikt} \\ \Leftrightarrow \frac{p_{ikt}}{\tilde{p}_{ikt}} &= \exp(\Delta_{ikt}) + 1 \end{aligned}$$

that is,

$$p_{ikt} = \tau_{ikt} \tilde{p}_{ikt}$$

with  $\tau_{ikt}$  appropriately defined. This exactly corresponds to the case of ad-valorem transport costs only.

As noted by ?,  $0 < \beta < 1$  represents the elasticity of freight costs to the export prices, increasing with the relative weight of the additive component in international transport costs. By estimating Equation (16) with a constant  $\beta$ , ? assumes that the share of additive costs does not vary over time, country partner or product. This is an important difference with our method, as we measure transport costs by explicitly allowing for an additive component that varies in the three dimensions (ie, somehow allowing for a varying  $\beta$  across time, country partner and product). To establish this point clearly, we replicate the method adopted by ? exposed above on our database (which is the same as his until 2004). The results are reported in Figure ??.