Trade Costs: An empirical assessment based on Law-of-One Price deviations and the Direction of Trade*

Ozlem Inanc
Department of Management
Isik University
Kumbaba Mevkii 34980, Sile-Istanbul, Turkey
oinanc@isikun.edu.tr

Marios Zachariadis†
Department of Economics
University of Cyprus
PO Box 20537, 1678 Nicosia, Cyprus
zachariadis@ucy.ac.cy

March 2008

Abstract

The importance of trade costs in segmenting product markets cannot be captured by considering aggregate prices or in the absence of information on the direction of trade. We address this problem by utilizing product-specific prices along with cross-sectional productivity measures and bilateral trade flows that allow us to identify the probable source of any one product. Our empirical approach is in line with the framework proposed in Anderson and van Wincoop (2004). The data are shown to be consistent with this framework. In particular, trade costs in the form of transportation and distribution costs, are important in determining international price differences and segmenting international markets.

Keywords: law of one price, trade costs, market segmentation.

JEL Classification: F4

*We thank Tibor Besedes, Mario Crucini, and participants at the Summer 2007 Microeconomic Pricing in an Open Economy Workshop at CEU-Budapest, the Spring 2006 Midwest Macroecomomics Conference at Washington University St Louis, the Universite Cergy-Pontoise THEMA Spring 2006 Seminar Series, the Spring 2005 Midwest International Economics Meetings at Vanderbilt University, and the ECB-CFS-Bundesbank Fall 2005 Seminar Series, for useful comments and suggestions.

†Corresponding author.

1. Introduction

Crucini, Telmer, and Zachariadis (2005) (CTZ) make the case that the Law-of-One-Price (LOP) and Purchasing Power Parity (PPP) are essentially about the cross-sectional distribution of international relative prices rather than the time-series behavior of changes in these, and that "economic theory places much starker restrictions on LOP deviations than on their changes"; the implication being that the gap between theory and empirics can be bridged through the use of microeconomic price levels enabling exact comparisons across space. In line with this, Anderson and van Wincoop (2004) propose the use of price levels comparable across locations at a point in time as a promising route for inferring trade costs, arguing that "*it* is hard to see how information can be extracted about the level of trade costs from evidence on changes in relative prices." They suggest that to extract information about trade costs from price levels "[*a*] natural strategy would be to identify the source country for each product," noting that "[*u*]nfortunately survey data often do not tell us which country produced the good."

Trade costs can be mismeasured if the distance between two locations does not capture distance between exporter and importer. For example, if trade between two countries does not occur for a product, then that price difference will lie between the no-arbitrage bounds and will be less than the trade cost. On the other hand, if both countries export the product to each other, the overall impact of trade costs on that product's price difference between the two locations can be zero even if these costs are large. In fact, a bilateral price difference reflects the size of trade costs only when one of the locations is the source of the product to the other.

In this paper, we use *microeconomic price levels* along with cross-sectional productivity indices and bilateral trade flows between countries to identify the likely source of each product. We then examine an empirical model proposed by Anderson and van Wincoop (2004), where international price dispersion is determined by *transport costs* and local trade (distribution) costs, as well as by taxes, good-specific characteristics and differences in markups.

Transport costs and broader trade costs are of central importance in many macroeconomic models, as in the recent examples of Bergin and Glick (2003) and Atkeson and Burstein (2004). However, assessing these at the macroeconomic level has proved problematic. Anderson and van Wincoop (2004) argue persuasively that "average price dispersion measures are not very informative about trade costs." In general, the impact of trade costs in segmenting individual product markets will be underestimated when considering aggregate prices or the average (over products) of price deviations. When aggregate prices or mean price deviations are considered, it is likely that countries both export and import to and from each other some of the goods that go into the construction of the composite price. As a result, the impact of trade costs on price differences could wash out on average even if trade costs were important in segmenting markets as determinants of international price deviations for individual products. This is the "*averaging-out property*" put forth by Crucini, Telmer, and Zachariadis (2004). Using actual retail prices from the Economist Intelligence Unit (EIU), they show that distance does not matter for the *mean* of deviations from the law of one price, once they control for the Balassa-Samuelson income effect.

Even when prices of individual products are available across international locations, trade costs can be mismeasured when the source of the product being compared across locations is not accounted for. This might be behind the small or non-existent estimated impact of physical distance on deviations from LOP often found in the literature. Using actual retail prices from the EIU, Anderson and Smith (2004) estimate price elasticities with respect to distance of 0.012 for the overall sample of countries and 0.006 for price comparisons within North America. Looking at particular countries, the price elasticities become negative, with estimates of -0.005 for the US and -0.017 for Canada.[1] Engel, Rogers and Wang (2003) estimate higher price elasticities with respect to distance using US-Canada actual retail price data. They consider *absolute* deviations from LOP which alleviates the "averaging-out" problem discussed in the previous paragraph. Their estimate for distance in the full panel of 100 items is 0.00321, and estimates range from 0.00584 for "Food items" down to -0.00004 for the category "Miscellaneous Products".[2]

Ceglowski (2003) takes a first step towards addressing the above concerns by considering relative distance from a "core location" (*assumed* to be Toronto) *in addition to* intercity distance. Using actual retail prices for 45 consumer goods across 25 Canadian cities, she explains *absolute* log deviations from LOP within Canada.[3] In adding distance from a "core location", she allows for the fact that "[*g*]eography could play an additional role in price differentials when prices include freight costs from a central or core location" since "[*i*]f shipping costs are a positive function of distance, this possibility implies a city's prices would be higher the further it is from the core location." The estimated coefficient is positive and significant for about half of the products (see her table 4,

[1] Their dependent variable is the log deviation from LOP, $\ln(e_{jkt}p_{ijt}/p_{ikt})$, considered for all goods i and all possible $j-k$ bilateral comparisons in every period t. We report their estimates from Table 2 for the case of tradeable goods. The specific country results are from their Table 3.

[2] These estimates are reported in Tables 2A, 3A and 4A respectively of Engel, Rogers and Wang (2003).

[3] In specifications that use Toronto as the core location, she also considers log deviations from LOP in place of their mean *absolute* values. This allows one "to include the information contained in the sign of the price differentials" (p.395), which can be useful since the *sign of the price differentials* contains information about the *direction of trade* when comparisons are made relative to the exporting location. Results in her Table B1 show that relative distance from the core location has a significant positive relation with these average price differentials for over half of the products, as one would expect if distance from the core captures transport costs. However, the estimated coefficient of relative distance is negative for several other products.

specification 3). She concludes that her "analysis uncovers a positive link between relative distances from a core location and price disparities, suggesting that the price effects of distance may not be fully captured by simple intercity measures."

We attempt to resolve the concerns raised in the previous paragraphs by utilizing product-specific price differences[4] and identifying core locations that differ across types of goods. To identify *the* most likely source for each product, we use information on the productivity of each country in each industry. As an alternative identification strategy, we use realized trade flows to determine the price of the product in the probable source as a weighted average price of an importing country's actual trading partners. Utilizing the unique (in terms of breadth of the goods covered and exact comparability across locations) microeconomic dataset of price levels across the European Union (EU) from CTZ, along with information on the *direction of trade*, we identify economically meaningful measures of trade costs in general and transport costs in particular through their estimated impact on product-specific retail price differences between importing and source countries.

To measure transport costs, we use geographic *distance*[5] and an industry-specific tradeability measure. To account for local trade costs, we use real income per capita differences and industry-specific features of local costs as captured by real wage rates and a non-traded factor input content measure. We find that country-specific aspects of transport costs measured by geographic distance are important in explaining deviations from the law of one price, with estimated price elasticities with respect to distance as high as ten percent.[6] Distribution costs and industry-specific measures of trade costs are also shown to be important.[7] Overall, the data are consistent with a model where transport costs as well as local trade costs play important roles in the determination of international retail price differences.

The remainder of the paper is organized as follows. Section 2 offers the theoretical motivation behind our empirical application. Section 3 describes the price data and the construction of cross-sectional TFP indices and trade-weighted relative prices. Section 4 reports the main results of our estimation exercise, and section 5 briefly concludes.

---

[4] Earlier work utilizing microeconomic price levels includes Parsley and Wei (1996, 2001). However, the focus of these papers in on the variability of LOP log deviations over time (as in the seminal work of Engel and Rogers, 1996, who however consider more aggregate data), so that their estimates of distance coefficients are not directly comparable to estimates in the literature focusing on price dispersion across space.

[5] Trade across these European countries is less likely to be characterized by high policy-related and other unidentified trade barriers, allowing geographic distance to act as a better measure of transport costs.

[6] To the extent that transport costs are relatively less important across these European countries, our estimates are a lower bound for average transport costs characterizing world trade.

[7] Countries with higher cost for the local inputs component (higher real income per capita or wage rates) have higher prices, an industry-specific measure of local input content has positive impact on absolute price dispersion, and industry-specific tradeability has negative impact on absolute price dispersion.

2. Motivation

Anderson and van Wincoop (2004) propose the use of actual price data comparable across locations at a point in time as a promising route of extracting information about trade cost levels. They consider a framework where the price of a final good is determined by production costs, trade costs, markups, and taxes. Abstracting from markups and taxes they are able to impose arbitrage constraints and derive an inequality that constrains international relative prices. The assumption here is that if country j buys from country κ, then p_j=$\mathrm{c}_k\tau_{j\kappa}$, where c_κ is the cost of production in κ and $\tau_{j\kappa}$ is the trade cost of transporting the good from κ to j. Moreover, country j will buy from κ if $\mathrm{c}_\kappa\tau_{j\kappa}$ is the lowest among all potential sources. The inequality for any two importing countries j and k is $\frac{c_{z_j}\tau_{jz_j}}{c_{z_j}\tau_{kz_j}} \leq \frac{p_j}{p_k} \leq \frac{c_{z_k}\tau_{jz_k}}{c_{z_k}\tau_{kz_k}}$ or $\frac{\tau_{jz_j}}{\tau_{kz_j}} \leq \frac{p_j}{p_k} \leq \frac{\tau_{jz_k}}{\tau_{kz_k}}$, where p_j and p_k are retail prices in countries j and k, and z_j and z_k are the optimal sources for countries j and k respectively. When countries j and k purchase the good from the same source, κ, then the above inequality is reduced to $\frac{p_j}{p_k} = \frac{\tau_{j\kappa}}{\tau_{k\kappa}}$, with the relative price now tied down by trade barriers. Finally, they conclude that "in the specific case where κ is one of the two countries, the relative price captures exactly what we intend to measure", as $\frac{p_j}{p_\kappa} = \frac{\tau_{j\kappa}}{\tau_{\kappa\kappa}} = \tau_{j\kappa}$.

That is, if one could identify the probable source country then one could capture the exact level of trade costs. This is in line with what we do in this paper. Specifically, we use independent information on the productivity of each country in each industry to identify the most likely source for each product. Utilizing productivity to identify the source is consistent with the above framework where a country buys from the cheapest source, and with the models of Eaton and Kortum (2002) and Bernard, Eaton, Jensen, and Kortum (2003), where the most productive country for any one product is the sole source of that product to the rest of the world. Alternatively, we consider actual trade flows to construct the price in the source, κ, as a weighted average of each country's within-sample trading partners.

Under the maintained assumptions above, the relative price thus obtained could be attributed to trade costs. However, controlling for a number of additional potentially important determinants of international price differences is necessary in practice if we are to best isolate the impact of trade costs. Our point of departure is the framework outlined in Anderson and van Wincoop (2004), where final goods prices might differ internationally to the extent that transport costs, local trade costs, taxes, and markups exhibit variation across countries and goods.

Given the absence of direct measures of transportation costs for broad cross sections of goods and countries[8], we follow the usual practice of using physical distance between the capital cities of

[8] See Hummels (1999), for an approach that incorporates direct measures of transport costs utilizing freight data

the countries in our sample to capture transportation costs.[9] That is, once we identify the probable source for each product, we identify the size of transport costs by the estimated coefficient of distance from the source country. In addition, as suggested in Anderson and van Wincoop (2004), we allow for industry-specific differences in transport costs, first through a measure of tradeability[10] as in CTZ and following that, through the use of industry-specific distance interaction effects.

We also account for the presence of local distribution costs through income per capita differences and by considering industry-specific features of these local costs as captured by the non-traded factor input content of each good.[11] Industry-specific features of local costs are also captured by domestic real wage rates. Differences in taxes across goods are captured by group-specific dummies for classes of goods that are likely to face higher taxes and where broadly available, by VAT differences across goods and countries.

Finally, consistent with the empirical findings of Campbell and Hopenhayn (2005)[12] and the theoretical model of Melitz and Ottaviano (2005)[13], we assume that larger markets are more competitive so that demand elasticities are higher and markups lower there. We utilize population size to capture market size. Larger markets are likely to have a greater number of exporters serving them -in the presence of some fixed cost component in trade costs- and are also more likely to have domestic production of close substitutes for imports -in the presence of some fixed cost component to production inducing economies of scale- both factors leading to a more elastic perceived demand for imports and lower prices in large markets. It might also be that potentially price-discriminating exporters value large foreign markets more than small ones, exhibiting greater risk aversion for losing large markets, and are thus less likely to have higher markups there in the presence of demand uncertainty. However, given the difficulty of capturing variation in markups across countries, an alternative starting assumption would be that markups are of similar level across countries so that they do not affect international price differences.[14] In that case, coefficients for population size

for a small group of importing countries.

[9] Disdier and Head's (*forthcoming*) summary of findings in the literature suggests an inversely proportional relationship between trade and distance. Although other factors might also be important, this relation should at least in part be attributed to trade costs increasing with distance.

[10] Since this industry-specific measure is based on realized trade flows, it might partly capture industry-specific trade costs other than transport costs. We note that *industry-specific* measures can only be considered as determinants of *absolute* price deviations, since *actual* price deviations can be negative or positive depending on the *direction* of trade across countries and can only be explained by factors that do vary across countries.

[11] We follow Anderson and van Wincoop (2004) in classifying transport costs and distribution costs as two categories of trade costs, the second of which is related to the *local cost component* of final prices.

[12] Using data across 225 U.S. cities, they find that larger markets are more competitive for most of the industries examined, and that doubling the number of competitors in a market decreases markups by at least that much.

[13] They show (p.24) that "larger markets exhibit tougher competition resulting in lower average markups."

[14] This assumption is discussed in Anderson and van Wincoop (2004) and imposed in Crucini, Telmer, and Zachariadis (2004).

differences could be measuring scale economies common across industries and specific to countries. That is, population size might be capturing scale economies that lower average domestic production costs leading to lower domestic prices in larger countries.[15]

3. Data Construction and Methodology

We now describe the diverse data we compile from a number of different sources. This labor-intensive task involves a concordance that allocates individual consumption goods prices into available industry-level information for our explanatory variables. We also describe the construction of variables to be used in the empirical analysis. This involves the creation of novel empirical concepts, like observed probability-weights based on import shares as explained in detail later on in this section.

Denoting p_{ij} as the local currency price of good i in country j, p_{ik} as the local price of the same good in country k, and e_{jk} as the nominal exchange rate of country j in terms of currency units of country k, we define law-of-one-price deviations as

$$\ln q_{ijk} = \ln(e_{jk}p_{ij}/p_{ik})$$

The retail price data utilized here are the same as that used in CTZ.[16] A detailed description of the data is provided in the latter paper. A comprehensive list of the goods is available at http://bertha.tepper.cmu.edu/eurostat. These data originates from Eurostat surveys conducted in different European cities sampled at five year intervals between 1975 and 1990. The level of detail often goes down to the level of the same brand sampled across locations and enables exact comparisons across space at a given point in time. The price data for each cross-section is collected in a sequence of surveys where the same group of goods are collected within the same period for all countries.[17] The Eurostat survey covers 9 countries for 658 goods in 1975, 12 countries for 1090 goods in 1980, and 13 countries for 1805 and 1896 goods respectively for 1985 and 1990. The nine EU countries in the 1975 survey are Belgium, Denmark, France, Germany, Ireland, Italy,

[15] The scale of domestic production depends on exports so that population size is less likely to capture scale economies from the production side and more likely to capture scale economies in domestic distribution and retail.

[16] We use the common currency prices with the outliers having being removed as in CTZ. CTZ remove the price entry for a good in a certain country when the price in that country differs by a factor of five from the average common currency price for that good across countries.

[17] In what CTZ call '1985,' for instance, the prices of most services were collected in September-October 1985, while prices of most clothing items were collected in December of 1984. The nominal exchange rate data with which prices were converted into a common currency takes explicit account of this timing, taking the form of averages of daily data over the relevant time intervals.

Luxembourg, the Netherlands, and the UK. Greece, Portugal and Spain are added in 1980 and Austria in 1985.

Goods were allocated into a three-digit industry in order to be matched with the industry-specific measures of the non traded input share, tradeability, and the real wage rate, as well as to TFP and bilateral import flows the construction of which is discussed in the next few paragraphs. The non-traded input share of the good is the ratio of non-traded input costs to total cost for each industry. Non-traded inputs are assumed to include: utilities, construction, distribution, hotels, catering, railways, road transport, sea transport, air transport, transport services, telecommunications, banking, finance, insurance, business services, education, health and other services. This measure is taken directly from CTZ who compute it using the 1988 input-output tables of the UK. The tradeability for each industry is measured as the ratio of total industry trade between countries in the sample divided by total output of that industry across the same countries, as in CTZ.[18] In order to limit measurement error issues, we use three-year averages of tradeability using the two preceding years along with each cross-section's sampling year.

The distance measure we use is the great-circle distance between the airports of the capital cities and is measured in kilometers. The capital city of each country is the sampling location of the price data for all countries but Germany for which reported prices are an average from a number of cities within that country. Thus, for Germany we use distance relative to Frankfurt which is a transportation center for that country. Population and real GDP per capita are obtained from PWT 6.1 for each of the cross-sections. The latter measure is the constant price chain series GDP per capita with code name rgdpch.

We also use data on VAT rates for 23 different categories of goods and services for all countries in our sample in 1990. For 1975, 1980, and 1985 VAT is not observed for Greece which entered the European Community (EC) in 1980, and for Portugal, and Spain which entered the EC in 1985. This is the same VAT data as in CTZ, assembled from the European Commission publication "VAT rates applied in the member states of the European Community" (2002) and other sources.

Data required for TFP calculation come from two World Bank sources: the Trade and Production Database, and the Database on Investment and Capital for Agriculture and Manufacturing. The Trade and Production Database collects production and trade information for 67 developing and developed countries from different sources and merges them into a common classification. The main source for its production data is the UNIDO and OECD joint collection program. We use value added in current dollars, fixed capital formation, wages and salaries, and the number of em-

[18] Both shares are listed in detail in tables A1 and A2 in the data appendix in CTZ.

ployees for 28 three digit manufacturing industries. Depending on the country, the coverage of data is from the late '70s to late '90s. Value added in current dollars is deflated to obtain value added in constant dollars using price deflators from the OECD STAN database. Wages in current dollars were deflated using the same price deflators from the OECD STAN database to obtain wages in constant dollars. The real wage rate utilized in the regressions was constructed as wages and salaries in constant dollars over the number of employees.

The Database on Investment and Capital for Agriculture and Manufacturing reports the total capital stock for the manufacturing sector. We calculate total manufacturing sector investment, using capital formation data for 28 manufacturing industries from the Trade and Production Database. We then construct each industry's investment share in total manufacturing for each country. Finally, we assume that the share of investment for the industry in total manufacturing for a specific year is equal to its share of the capital stock, and use this observed industry share and total manufacturing capital stock to calculate capital stock for each manufacturing industry.[19]

With the data at hand and under the assumption of a Cobb-Douglas production function, total factor productivity (TFP) between countries j and k for industry h is constructed following Harrigan (1997) as

$$TFP_{hjk} = (Y_{hj}/Y_{hk})(L_{hk}/L_{hj})^s(K_{hk}/K_{hj})^{1-s}$$

where Y denotes real value added, L is the number of employees, K is the capital stock for each industry and s is the average share of labor in total cost between j and k. In calculating TFP, we use three-year averages of the variables using the two preceding years along with each cross-section's sampling year. The data for constructing TFP is not available to us for 1975 and is only available for five of the above countries in 1980 limiting our ability to identify the source country. This is the reason why for regressions that utilize TFP information, we consider only 1985 and 1990 for which this is available for an identical sample of eight countries: Austria, Denmark, Germany, Greece, Ireland, Italy, Portugal, and the UK. Moreover, throughout the paper, we consider manufacturing goods prices since we do not have disaggregate data for constructing TFP across industries of the service sector. Excluding services industries is not necessarily a handicap of the analysis as we are primarily interested in transport costs faced by traded goods.[20] The availability of the TFP measure across industries is reported in table A1 in the Data Appendix.

[19] A data appendix made available upon request provides additional details on the construction of the capital stock.

[20] Arbitrage models as in Lee (2004) show that price differences across countries will equal the trade costs for products that are traded while endowment or productivity differences will determine the exact degree of deviations from the law of one price for products that are not traded in equilibrium.

To identify the probability-weighted source for each good sold in each country of the Eurostat price dataset, we utilize bilateral trade flows from the OECD *International Trade by Commodity Statistics* (ITCS) database. Here, we use the full sample of countries and years allowed by the CTZ price data, with the exception of Luxembourg, as the data requirements of TFP construction no longer constrain us. Utilizing this broader sample of countries is desirable as it enhances our ability to assess probable source for each product among a broader group of possible origin countries.

The ITCS database includes annual bilateral flows in current $US between 269 international locations for 2581 goods categories for the period 1960-2000. We inspected this list of traded goods categories and came up with a list of 68 product categories chosen to best relate to the products from the Eurostat price data. These 68 categories which are described in the first column of table A2 in the Data Appendix, were then aggregated by ISIC code into 42 separate 4-digit categories of the manufacturing sector, shown in the second column of Table A2, that are finally mapped onto the disaggregated product prices from the Eurostat data.[21] We end up with imports for each of 42 industries of each country in our sample from each other.[22] That is, we consider imports of country j from each of the other countries in our Eurostat price data for each industry h. For each importer j and industry h, the probability-weighted source price for a specific product is defined as the weighted average of the prices of exporters of that product with weights calculated using bilateral trade flows for each cross-section.

Denoting im_{jkt}^{h} as imports of country j from country k for industry h in cross-section t, the weight of exporter country k for importer j in industry h is defined as $w_{jkt}^{h} = im_{jkt}^{h} / \sum_{k=1}^{n-1} im_{jkt}^{h}$, where n is the number of countries in sample. However, some exporting countries have missing prices for some goods so that the sum of the above weights would not add up to one in these cases. To cope with this, we re-scale the weights.[23] The price in the probability-weighted origin is then simply given by the weighted sum of exporting countries' observed prices:

$$p_{j\kappa t}^{h} = \sum_{k=1}^{n-1} w\_new_{jkt}^{h} * p_{kt}^{h}$$

where we have one probability-weighted source, κ, for each importer j in each industry h. We can then compare the price of each product sold in the importing location relative to this probability-

[21] There is a many to one mapping from goods for which we have prices to the 4-digit categories in the trade data. Future work should focus on disaggregated trade data that can be closely matched to products in the price surveys.

[22] As we are constrained by the number of countries for which we have price data, we actually use eight countries for 1975, eleven for 1980, and twelve for 1985 and 1990. We note that while in the price data, Belgium and Luxembourg prices are given seperately, the bilateral flows dataset includes the aggregate of Belgium and Luxembourg reducing the number of countries we can consider by one for each cross-section.

[23] For each good, we consider only imports from countries for which the price is observed so that the new weights $w\_new_{jkt}^{h}$ are given by multiplying w_{jkt}^{h} by $\sum_{k=1}^{n-1} im_{jkt}^{h}$ over the new imports sum.

weighted source. The same weights are used in order to construct real GDP per capita, the real wage rate, population and distance variables of the probability-weighted origin relative to which we compare the respective variables of the importing country.

Finally, we add the effect of domestic production of the importer country into the analysis. Domestically consumed production of country i for industry h is defined as the difference between total output and exports of country i for that industry. Total output and exports data were obtained from the Trade and Production Database at the 3-digit level of the manufacturing sector. We treat domestically consumed production of country j for industry h in cross-section t as an import from itself and re-define the weight of exporter country k for importer j in industry h as $w^h_{jkt} = im^h_{jkt}/\sum_{k=1}^{n} im^h_{jkt}$, where n is the number of countries in sample including the importer country j itself. We then re-scale the weights as explained in the previous paragraph. The price in the probability-weighted origin is again given by the weighted sum of exporting countries' prices: $p^h_{j\kappa t} = \sum_{k=1}^{n} w\_new^h_{jkt} * p^h_{kt}$, where the price of the importing country itself is now included in this calculation Again, real GDP per capita, the real wage rate, population, and distance for the probabilistic exporter are calculated by using these same weights. These weighted variables are then used to construct log differences relative to the importing country. To facilitate the construction of relative distance, distance from the importing country is defined as $dist_{jj} = (A_j/\pi)^{0.5}$ where A_j is the surface area of importer country j in squared kilometers.

4. Estimation and Results

Based on the framework discussed in section 2, price differences between importing and source locations for each product are determined by transport costs and international differences in local distribution costs, taxes, and markups. We consider the following regression equation

$$q_{ijk} = a_0 + a_1 Dist_{jk} + a_2 y_{jk} + a_3 Pop_{jk} + \varepsilon_{ijk} \tag{4.1}$$

where q_{ijk} is the log deviation from the Law of One Price (LOP) for good i between countries j and k, a_0 is a constant term, and ε_{ijk} is a random error[24]. $Dist_{jk}$ is the (log) distance separating the capital cities of the two countries and is meant as a proxy for transportation costs impeding trade and maintaining price differentials across j and k. The variable y_{jk} is the log difference in real

[24]As shown in Crucini, Telmer, and Zachariadis (2000), it is necessary to correct the standard errors for heteroskedasticity in this specific context, where we use aggregative values of the explanatory variable to explain a highly disaggregated dependent variable. This creates a heteroskedastic pattern in the variance of the regression term as shown in the earlier paper. This type of aggregation also makes goodness of fit measures difficult to interpret, so that the low R^2s reported here should be taken with caution.

GDP per capita between j and k and captures the local (distribution) cost component suggested by the theoretical framework from CTZ and Anderson and van Wincoop (2004). That is, GDP per capita captures a "wage effect" whereby richer countries will have higher non-traded sector labor costs. Alternatively, we directly measure the wage component of local distribution costs by using the real wage rate w_{hjk} in place of y_{jk} where h denotes a three-digit industry classification with a one-to-many mapping into individual goods i. Finally, Pop_{jk} is the log difference in population size in 000's between countries j and k and is meant to capture the effect of domestic market size. The inclusion of population size is consistent with standard gravity models. A description of how all the variables were constructed is undertaken in the next section.

Initially, equation (4.1) is estimated for *all* possible *unique* bilateral price comparisons j-k (similar to what is done in some of the literature) with each bilateral comparison made according to alphabetical order rather than relative to countries more likely to be the source of a product. In this case, the *sign of the price differentials* will not be informative about the *direction of trade* as price comparisons are not made relative to the exporting location. This renders meaningless the coefficient of geographic distance as a proxy for transport costs. The estimate for the price elasticity with respect to distance is statistically indistinguishable from zero equal to -0.009 for 1985, and 0.055 and significant beyond the one percent level in 1990. These estimates are devoid of meaning as distance between two arbitrary countries does not necessarily capture distance between exporter and importer. Again, if trade between two countries does not occur for a product, then that price difference lies between the no-arbitrage bounds and is less than the trade cost. It is also possible that both countries being compared export the same product to each other. To the extent that this is the case, the final price for these products incorporates a similar transportation cost in both countries so that there might be little or no impact of transportation costs on price differences for these products between the two countries. Moreover, considering all possible bilateral comparisons tends to average out around zero the impact of transportation costs on prices producing unreliable estimates.[25] Thus, we view estimation using all bilateral comparisons as a mere reference point with which to compare the distance coefficient estimates we obtain once we use information on the probable source of the products.

---

[25] It is possible that k is the main exporter to j for some product i and does not import this from j, and that j is the main exporter to k for some product i' and does not import this from k. The overall result is a possible washing out of the effect of transport costs across goods for an aggregate or average measure of prices, related to the "averaging-out" property discussed in Crucini, Telmer, and Zachariadis (2004), and can be addressed by using absolute price differences or an appropriate variance measure for each product across countries.

Utilizing information on relative productivity

In general, in the absence of some information regarding the *direction of trade*, the distance coefficient cannot capture transport costs well in the context of "directional regressions" such us those estimated for equation (4.1) above. Considering LOP deviations as in equation (4.1), rather than their absolute values, allows us "to include the information contained in the sign of the price differentials" (Ceglowski, 2003, p.395), in accord with the premise that distance from the source captures transport costs. However, the *sign of the price differentials* contains information about the *direction of trade* only when price comparisons are made relative to the exporting location. This is why Anderson and van Wincoop (2004) argue that to estimate the precise role of transportation costs in determining differences in price levels for a good between locations, we would need to know the potential source of that good.

A way to address the problem of identifying the exporting country, is to assume the *most* productive country in the sample to export the good to others, in line with the theoretical models of Eaton and Kortum (2002) and Anderson and van Wincoop (2004).[26] Thus, we first rank countries according to their productivity in each industry and then denote the *most* productive country to be the *source* or reference country for that specific industry. Under the assumption that the most productive country for a certain industry will be the main exporter of goods of that industry, we construct good-specific log relative prices between each country j relative to the main exporter country κ for each industry h. As we show next, this methodology goes some distance into identifying the source country, providing improved estimates for transport costs.[27]

In what we refer to as Model 1, we consider regression equation (4.1) to explain $q_{ij\kappa}$, the log deviation from the Law of One Price (LOP) for good i between country j and κ, where κ is now the most productive country in industry h assumed to be the main source for product i in country j. To estimate Model 1, we construct the dependent variable of prices relative to the most productive

[26] An other way to address the problem discussed above is to assume that the *more* productive among any two countries being compared will export the good to the other. A problem with this is that, given measurement error associated with TFP construction, comparing countries with similar productivity is likely to give the wrong ordering because of measurement error more often than when considering price comparisons relative to the *most* productive country in the data. The latter is preferable as it avoids an ordering based on comparisons among countries closer together in tems of productivity. Yet an other approach would be to assume one of the locations to be the main exporter using a-priori information, similar to the assumption of Ceglowski (2003) that Toronto is the core location. However, estimation results using Germany or the U.K. as the core location, suggest that both the sign and significance of distance coefficients are not robust across periods or reference countries.

[27] Before turning to estimation using price differences relative to the most productive country, we evaluate the hypothesis that productivity is inversely related to prices, consistent with productivity being a determinant of the direction of trade. Estimating (4.1) with an added term for total factor productivity differences, TFP_{hjk}, across countries for each three-digit industry h, the elasticity of price differences with respect to TFP_{hjk} is $-.035$ for 1985 and $-.042$ for 1990, both statistically significant beyond the one percent level.

Table 4.1: Price differentials relative to Most Productive Country for each industry

	1990			1985		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Population	-.050*	-.010***	-.0003	-.056*	-.001	.015
	(-11.14)	(-1.90)	(-0.02)	(-13.91)	(-0.28)	(1.27)
GDP-per-capita	.422*	-.048	-.009	.280*	.027	-.104
	(12.33)	(-1.18)	(-0.08)	(11.84)	(1.05)	(-1.20)
Distance	.090*	.064*	.093**	.041*	.100*	.130*
	(6.61)	(6.13)	(2.08)	(2.98)	(10.00)	(3.94)
Tradeability		-.061*	-.034		-.046*	-.063**
		(-3.86)	(-0.93)		(-3.04)	(-2.13)
Non-tradable input share		.008*	.006		.010*	.005
		(3.49)	(1.47)		(4.74)	(1.14)
Large cars		.084*			.147*	
		(2.97)			(2.70)	
Vices		.187*	.151*		.162*	.089**
		(5.96)	(2.88)		(5.99)	(2.19)
Constant	.047*	.255*	.290*	.016**	.228*	.311*
	(6.54)	(10.85)	(4.92)	(2.25)	(9.83)	(5.87)
R^2(in percentage)	4.8	5.0	15.7	5.4	5.7	17.6
Observations	3186	3186	123	3373	3373	132
Countries	8	8	8	8	8	8

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. In Model 2 we take absolute values for all variables. In Model 3 we use the mean of absolute LOP deviations. The eight EU countries considered here are: Austria, Denmark, Germany, Greece, Ireland, Italy, Portugal, and the UK.

location, using the industry-specific country ranking implied by cross-sectional TFP levels. $\text{Pop}_{j\kappa}$ and $y_{j\kappa}$ are population and real GDP per capita or real wage rate log differences between country j and origin country κ. $\text{Dist}_{j\kappa}$ denotes log distance between source κ and destination j. All explanatory variables are demeaned so that the constant a_0 can be interpreted as the price deviation relative to source κ at average levels of distance, real GDP per capita, and population size in the sample. Our specification incorporates information regarding the direction of trade and can thus help assess the overall level of trade costs and the transport cost component of these implied by the estimated coefficient for physical distance. Results from this estimation framework are summarized in columns 1 and 4 of Table 4.1 for 1990 and 1985 respectively.

Distance from the origin has a positive and significant impact on international price differences, suggesting a role for transportation costs as a determinant of these. Based on 1990 estimates, doubling distance increases prices by 9 percent, substantially greater than the 5.5 percent estimate for the specification with all unique bilateral price comparisons (not shown in the Tables). The improvement in terms of the estimated price elasticity with respect to distance, is even more striking for 1985. This changes sign (from -0.9 percent to 4.1 percent) becoming positive and strongly significant once we account for the probable source of products. When the most productive country for each industry is chosen as the reference location, distance consistently has a positive and significant effect on relative price levels. As the distance between source and destination country increases, transportation costs go up and so does the price of the good in the destination country. We conclude that our approach goes some distance in capturing the likely source country for each industry, even if the existence of multiple products within any industry creates aggregation bias that might still wash out the impact of transport costs or distance to a considerable degree.[28]

Local costs as captured by real GDP per capita appear to have strong effects on price differences, with elasticities equal to 28 percent for 1985 and 42 percent for 1990. Moreover, according to our estimate of the constant term, the importing country typically had prices which were 4.7 percent higher than the source at mean levels of the explanatory variables in 1990. Finally, population size has a negative effect on price differences with an estimated price elasticity of minus 5.6 percent in 1985 and minus 5 percent in 1990. The finding that higher population is associated with lower prices in a country suggests a potential role for markup differences across countries due to differences in demand elasticities that are positively related to the size of the market. Scale economies in distribution related to domestic market size might also be behind this finding.

Absolute Law of One Price Deviations

Next, we utilize absolute law of one price deviations relative to the most productive country. This formulation allows us to consider the impact of good-specific variables that are common across countries and which help explain overall price dispersion. Specifically, we consider tradeability and the non-traded factor component of goods as in CTZ. This formulation also allows aggregation into mean absolute price differences which gives more meaningful measures of the goodness of fit. We plot the bivariate relation between mean absolute price differences and distance for 1985 and 1990 in Figures 4.1 and 4.2 respectively. The visual evidence supports a positive relation between these two variables.

Model 2 describes the relation of absolute price differences with absolute values of variables included in Model 1 and industry-specific variables like tradeability and the local factor input content of goods in each industry. Using absolute price differences serves three purposes. First, it

[28] This approach does not fully resolve the problem of identifying the source country for each good since the measure of productivity is at the three-digit level. Moreover, for each destination country there might be more than one main exporter of goods in a certain industry and this exporter might or might not be among the countries in our sample. We begin to address these problems in the section *Utilizing Trade Flows*, where we use bilateral imports among the countries in our sample to obtain the probability that a good sold in a certain location was imported from any of the countries in the sample, and by making use of the share of imports from non-EU countries to restrict the sample to goods that are more likely to be imported from countries in our sample.

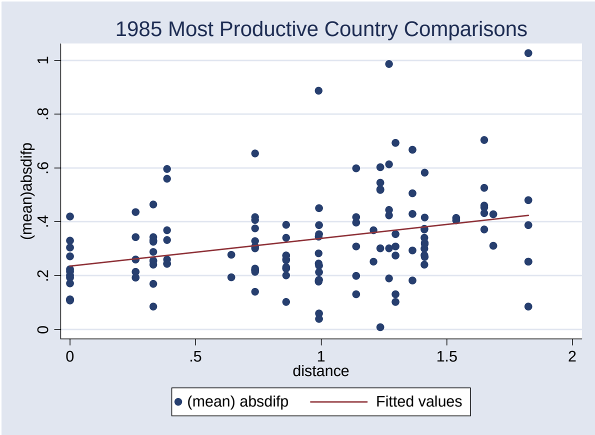

Figure 4.1:

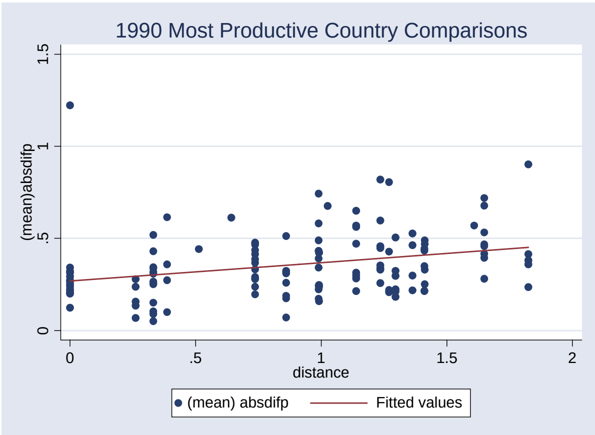

Figure 4.2:

allows the use of distance as a meaningful determinant of (absolute) price dispersion even in the absence of accurate source country information. This is the case since it resolves the "averaging-out" problem as long as prices are available at the product level to enable appropriate aggregation. Second, it makes it possible for us to consider industry-specific variables that cannot be considered in the context of Model 1. So, we now include tradeability and the non-traded factor input content variables from CTZ. We expect goods characterized by a higher degree of tradeability to have smaller absolute price dispersion and goods with higher local input content to have a higher degree of absolute price dispersion. In our empirical specification, these industry-specific variables enter along with the country-specific measures of transport costs and local distribution costs, where separate impact of industry and country-specific factors would suggest that these trade costs exhibit heterogeneity across both industries and countries. Third, we can now introduce dummy variables related to characteristics of categories of goods. These are intended to control for the degree of tax differences present for certain products where we have some a priori evidence (but no good-specific data) regarding particularly high differences across countries. We expect such goods to be characterized by a higher degree of absolute price dispersion.

We estimate the following regression equation for Model 2:

$$|q_{ij\kappa}| = a_0 + a_1 Dist_{j\kappa} + a_2|y_{j\kappa}| + a_3|Pop_{j\kappa}| + a_4 X_h + \varepsilon_{ij\kappa} \tag{4.2}$$

where X_h is a vector of industry-specific and category-specific variables capturing product characteristics as described above. Remaining variables are defined as in regression equation (4.1). The constant a_0 captures price dispersion at mean distance, real GDP per capita, and population size in the sample.

Results for Model 2 indicate that as distance between countries increases so does absolute price dispersion. For example, based on the 1985 estimates, a doubling in distance increases absolute price dispersion by ten percent. We also see that price differences are lower for goods that belong to more highly tradeable industries. To the extent that more tradeable goods face lower effective transportation costs this result suggests a role for transport costs in determining absolute price dispersion. Thus, both bilateral distance and industry-specific aspects of transport costs matter for absolute price dispersion.

Higher local input share implies higher absolute price dispersion as the model discussed earlier would predict. However, GDP per capita is statistically indistinguishable from zero, suggesting that higher income differences are not associated with higher price dispersion; while richer countries tend to have substantially higher prices as shown in Model 1, it is not the case that absolute price dispersion increases as the income gap across two countries becomes wider.[29]

Dummies for large cars and vices have positive and significant effects on absolute price deviations. If a good belongs to the group classified by one of these dummies, its price difference between countries will be larger, suggesting a role for tax differences in determining international price dispersion. Finally, the constant a_0 captures price dispersion relative to the source at sample average levels of the explanatory variables. This is equal to 22.8 percent in 1985 and 25.5 percent in 1990.

For Models 1 and 2, goodness-of fit measures are very low. Price data are more disaggregated than explanatory variables, thus the R^2 is not meaningful for these models. As explained in CTZ,

[29] The small sample of relatively similar income countries considered, resulting in small variation in income in these data, might be behind this finding.

this type of aggregation makes goodness of fit measures difficult to interpret so that the low R^2 reported here should be taken with caution. In order to alleviate the problem we aggregate the data. We use mean absolute price differences for each bilateral pair of countries in each three-digit industry and then run Model 2 again on the same explanatory variables as before. This is Model 3 for which results are reported in Table 4.1. The goodness-of-fit increases substantially.

The estimated coefficient for distance is positive and significant in both cross-sections for Model 3 as was the case in Model 2 but is now higher, at 13 percent for 1985 and 9 percent for 1990. Remaining coefficient estimates are insignificant except for tradeability in 1985 estimated at −6.3 percent at the one percent level of statistical significance, the estimated coefficient for vices[30] which is positive and significant in both cross-sections, and the constant a_0 which is equal to 31 percent in 1985 and 29 percent in 1990.

The Real Wage Rate as a proxy for local distribution costs

So far we have used GDP per capita to proxy for local distribution costs. This is estimated to be positive as expected in Model 1 but statistically indistinguishable from zero in Models 2 and 3. Here, we use an other proxy of the wage component of local distribution costs. We re-estimate Models 1, 2 and 3 utilizing information on industry-level real wage rates across countries. Since countries with higher GDP per capita will typically have higher wage rates, we do not include both measures to avoid the inherent collinearity problem for these two variables. Industry-level real wage rates capture the local cost component attributed to labor specific to each industry. That our wage measure captures variation across both industries and countries is an advantage relative to country-specific measures of real GDP per capita. This exercise also offers a robustness check for our coefficient estimates for distance, tradeability, and industry-specific local input costs.

We report results utilizing the wage rate in place of GDP per capita in Table 4.2. The real wage rate has positive impact on price differences in Model 1 with price elasticities of 15 percent for 1985 and 20.9 percent for 1990, in both cases about half those with respect to GDP per capita for Model 1 in Table 4.1. The population coefficient is estimated to be negative as expected if size is associated with smaller prices, with point estimates −3.1 percent for 1985 and −2.7 percent in 1990 both statistically significant beyond the one percent level. Moreover, according to the estimate of the constant term in Model 1, the importing country typically had prices which were 4.8 percent higher than the source at mean levels of the explanatory variables in 1990.

The estimated coefficient for distance is now bigger than the coefficients estimated when GDP

[30] Since we aggregate up to the 3-digit ISIC category level, the dummy for "large cars" cannot be included in the regression, as "large cars" is just one of several product types within 3-digit category 384.

Table 4.2: Price differentials relative to Most Productive Country for each industry

	1990			1985		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Population	-.027*	-.012**	-.002	-.031*	.001	.009
	(-7.46)	(-2.55)	(-0.15)	(-9.05)	(0.27)	(0.86)
Real Wage Rate	.209*	.034***	.096***	.150*	.054*	.046
	(12.06)	(1.66)	(1.96)	(8.31)	(3.71)	(1.15)
Distance	.101*	.046*	.049	.070*	.081*	.097*
	(7.21)	(3.98)	(1.24)	(4.60)	(7.63)	(2.71)
Tradeability		-.063*	-.030		-.051*	-.073**
		(-3.92)	(-0.82)		(-3.36)	(-2.46)
Non-tradable input share		.008*	.007***		.010*	.005
		(3.55)	(1.70)		(4.87)	(1.21)
Large cars		.088*			.152*	
		(3.17)			(2.81)	
Vices		.186*	.151*		.157*	.088**
		(6.00)	(3.06)		(5.83)	(2.24)
Constant	.048*	.254*	.276*	.014***	.229*	.317*
	(6.68)	(10.78)	(4.78)	(1.87)	(9.86)	(5.95)
R^2(in percentage)	5.3	5.0	17.8	4.3	6.2	17.4
Observations	3186	3186	123	3373	3373	132
Countries	8	8	8	8	8	8

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. In Model 2 we take absolute values for all variables. In Model 3 we use the mean of absolute LOP deviations. The eight EU countries considered here are: Austria, Denmark, Germany, Greece, Ireland, Italy, Portugal, and the UK.

per capita was included instead of wage rates. The distance coefficient is now estimated to be ten percent for 1990 and seven percent for 1985, compared to nine and four percent respectively in the estimations presented in Table 4.1 utilizing GDP per capita. Comparing these estimates of distance with the ones obtained using all unique bilateral comparisons (not shown in the Tables), these are now considerably larger. For 1990, the distance coefficient point estimate using all bilateral comparisons had been equal to 3.5 percent, and for 1985 this had been negative at -0.5 percent and statistically indistinguishable from zero. The improvement in terms of estimating the distance coefficient using most productive country comparisons relative to all bilateral price comparisons, is thus even more pronounced when we consider wage rates instead of GDP per capita.

Estimates of the variables in Models 2 and 3 in Table 4.2 are qualitatively similar to those in Table 4.1. The coefficient estimates for the industry-specific measures of tradeability and the local factor input content are virtually unchanged relative to those reported in Table 4.1, estimated to be negatively related to prices in the first instance and positively for the second. The coefficient estimates for distance become smaller relative to the specification with real GDP per capita. Finally, the population size coefficient is again estimated to have the wrong negative sign in Model 2 for

Table 4.3: 1990 with VAT

	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Population	-.034*	- .008	- .003	-.016*	-.010**	-.007
	(-6.74)	(-1.52)	(-0.21)	(-4.10)	(-2.20)	(-0.54)
Real GDP per capita	.336*	- .062	.002			
	(9.47)	(-1.56)	(0.01)			
Real Wage Rate				.178*	.055**	.113**
				(10.09)	(2.54)	(2.14)
Distance	.064*	.062*	.091**	.077*	.035*	.042
	(4.74)	(5.99)	(2.01)	(5.39)	(2.85)	(1.02)
VAT	.735*	.395*	.214	.820*	.454*	.503
	(6.44)	(3.77)	(0.58)	(7.32)	(4.07)	(1.34)
Tradeability		-.052*	-.034		-.054*	-.029
		(-3.33)	(-0.92)		(-3.43)	(-0.78)
Non-tradable input share		.008*	.006		.008*	.007***
		(3.48)	(1.45)		(3.60)	(1.72)
Large cars		.075*			.080*	
		(2.61)			(2.86)	
Vices		.200*	.153*		.200*	.155*
		(6.34)	(2.88)		(6.41)	(3.09)
Constant	.047*	.251*	.293*	.048*	.248*	.281*
	(6.57)	(10.65)	(4.96)	(6.70)	(10.50)	(4.88)
R^2(in percentage)	6.0	5.4	15.8	6.9	5.6	18.5
Observations	3186	3186	123	3186	3186	123
Countries	8	8	8	8	8	8

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. In Model 2 we take absolute values for all variables. In Model 3 we use the mean of absolute LOP deviations. The eight EU countries considered here are: Austria, Denmark, Germany, Greece, Ireland, Italy, Portugal, and the UK.

1990. Price dispersion relative to the source at average levels of distance, real wage rates, and population size in the sample is equal to 31.7 percent in 1985 and 27.6 percent in 1990 for Model 3, almost identical to the estimates of the constant term in Table 4.1.

To test robustness, we reconsider Models 1 to 3 adding VAT log differences as an explanatory variable. VAT is not observed for Greece, Portugal, and Spain except for 1990. Thus, we do not consider earlier periods as this would reduce our small sample to five countries and limit our ability to "guess" the probable source for each industry.[31] We report results in Table 4.3. For Model 1, the estimated coefficient for VAT differences is positive, very high, and strongly significant. The remaining estimates are for the most part similar to those in Tables 4.1 and 4.2. The estimated distance coefficient in Model 1 for the specification using price comparisons relative to the most productive country now falls to 6.4 percent and to 7.7 percent in the first and fourth columns of

[31] We want the maximum possible number of countries so that the most productive country in our sample is more likely to be the actual source.

Table 4.3, relative to 9.0 and 10.1 percent in Tables 4.1 and 4.2. Although lower than prior to the inclusion of VAT differences, these estimates are still higher than those obtained using all bilateral comparisons. For Models 2 and 3, the distance coefficients before and after inclusion of VAT are nearly unchanged and so are coefficient estimates for tradeability and local input content.

Utilizing trade flows

Assuming the most productive country in an industry to be the sole exporter of goods of that industry to countries in our sample does not completely resolve the problem of identifying the source. It is possible that a product is exported by more than one country. To cope with this, we use information about industry-specific bilateral trade flows across the countries in our sample so as to take into consideration that the same type of good can be exported by more than one country.

However, the goods could also be imports from countries outside the EU sample we have price data for. To the extent that this is the case, our within-sample import weights will not reflect the true probability that a good sold in one location is imported from an other location in the sample. Indeed, for some countries and industries, important exporters are outside the EU sample we have price data for.[32] To address this problem, the ratio of imports from the EU over total imports was constructed for each importer and industry, and if this was lower than 50 percent the good belonging to that industry was dropped from the dataset. This approach raises the likelihood that any good considered is actually imported from an EU country, so that we can better identify the source and more precisely estimate transport costs relevant to our sample of countries.

We proceed to utilize realized trade flows among the countries in our sample so as to determine the direction of trade and construct price comparisons for each product consumed in the importing country relative to countries likely to be *a source* for that product. The probability that a country in our sample is the exporter to a given destination for a good belonging to a given industry is constructed for each industry and destination as the ratio of imports from that country to the given destination over total imports to that destination. For each destination country and industry, we construct a weighted price as the sum of weighted exporting country prices where the weights are simply the ratios from above and as described in detail in the data section. Finally, the prices in the destination country are compared to this weighted sum to obtain log price differences for each good. This is the dependent variable we use now to estimate equation (4.1). Source κ is now a weighted sum of probable exporters with probabilities obtained as above.

[32] The share of EU imports for countries in our sample in 1990 was 84 percent for "furniture except metal industries" but only 51 percent for "tobacco and tobacco product industries." The import share from the EU also varies across countries for the same industry, with the share of EU imports for France, Italy and Greece in "tobacco and tobacco product industries" in 1990 higher than 90 percent and those for Denmark and Spain 11 and 8 percent respectively.

Table 4.4: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980	1975
	Model 1	Model 1	Model 1	Model 1
Population	-.029*	-.041*	-.011**	-.010**
	(-6.87)	(-8.71)	(-1.99)	(-2.31)
GDP-per-capita	.186*	.180*	.270*	.253*
	(8.12)	(5.17)	(6.89)	(8.42)
Distance	.031*	.058*	.045*	.075*
	(3.49)	(5.39)	(3.51)	(6.21)
VAT	.606*	.748*	.804*	1.12*
	(6.53)	(6.29)	(5.51)	(10.18)
Constant	.028*	.043*	.025*	.012**
	(5.11)	(8.82)	(3.86)	(2.23)
Pseudo-R^2(in percentage)	2.2	4.5	2.8	5.6
Observations	6848	5840	2775	2759
Number of countries	12	9	8	8

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. The eight countries in the 1975 sample are: Belgium, Denmark, France, Germany, Ireland, Italy, the Netherlands, and the UK. Austria is added in 1985. Greece, Portugal and Spain are added in 1990 since VAT is not observed for these countries except in the 1990 sample.

The price data have been cleansed of outliers as in CTZ. However, using trade quantities introduces an additional source of outliers given the well known measurement problems with trade flows measures. In order to handle outliers we minimize an absolute loss function and obtain the median estimator minimizing absolute deviations from the median rather than squared deviations from the mean. The median as an estimate of central tendency is not as greatly affected by outliers as the mean, alleviating the outliers problem.

In Table 4.4, we report estimates from this specification. The distance coefficients are estimated precisely and always positive for 1975, 1980, 1985, and 1990. The estimated price elasticity with respect to distance is as high as 7.5 percent in 1975 declining down to 3.1 percent by 1990.[33] These estimates taken in their totality suggest that transport costs are important for the determination of international price differences. Moreover, these estimates -using actual realizations of trade flows across countries- offer a clear improvement relative to those obtained using arbitrary comparisons, and are qualitatively similar to those obtained under the assumption that the most productive country in an industry is the sole exporter for products of that industry, providing support for productivity's role as predictor of the direction of trade.

An interesting feature that emerges, is the falling importance of transport costs in absolute

[33] As a robustness check, we utilize industry-specific real wage rates. As expected, wages impact positively on prices. The effect of distance decreases monotonically from 6.3 percent in 1980 down to 3.1 percent in 1990.

terms as witnessed in the declining estimated coefficient for the impact of physical distance on prices from 1975 to 1990 for this group of European countries. This is consistent with economic intuition as transport technologies have likely been improving over time for these countries.

The estimates for the impact of the local cost component of trade costs reported in Table 4.4 are positive and precisely estimated for each year in our sample, with price elasticities ranging from 27 percent in 1980 to 18 percent in 1985. Comparing the price elasticity of local trade costs in 1975 to that in 1990, these fall from 25.3 to 18.6 percent suggesting falling importance for this component of trade costs during the period.

The size of the population is consistently estimated to have a negative impact on prices with an estimated negative price elasticity, ranging from one percent in 1975 to 4.1 in 1985. As long as demand elasticities are positively related to the size of the market, this latter finding is consistent with markups being higher in smaller less competitive markets. Finally, the estimate of the constant term suggests that the importing country typically had prices 2.8 percent higher than the source at median levels of the explanatory variables in 1990.[34]

In addition, VAT differences have a strong but declining positive impact on price differences ranging from 112 percent in 1975 down to 61 percent in 1990 as tax rates become more homogeneous over the period. That is, while VAT rate differences have been very strong determinants of intra-European price differences, they have been steadily falling in importance throughout the period from 1975 to 1990 as would be expected from the EU policy of tax harmonization.

So far we have not accounted for consumption of domestic production. We now address this shortcoming of our analysis by allowing for the possibility that a product consumed in the home country can be produced domestically. *Domestically consumed production* of country j for industry h is defined as the difference between total output of country j for industry h and exports of country j for that industry. Results are reported in Table 4.5.[35] As seen there, distance coefficients fall for all three cross-sections relative to what is reported in Table 4.4, perhaps reflecting lower within-country transport costs. For 1985 the estimated distance coefficient decreases from 5.8 to 1.4 percent.

The domestic distribution cost as proxied by real GDP per capita has a positive impact on prices. Once again, this exhibits a falling tendency during the period, ranging from a high of 33.5

[34] Similarly to demeaning explanatory variables previously in our OLS regressions, we remove the median for all explanatory variables so that the constant is the price deviation relative to the source at median levels of distance, real GDP per capita, VAT, and population size in the sample.

[35] We do not have output data for these countries and industries for 1975. We also note that since domestic production is calculated at 3-digit aggregation, the weights are generated at that level for the specifications considered here rather than at the 4-digit level considered previously. Estimation without domestic production utilizing 3-digit weights gave estimates close to the ones reported in Table 4.4.

Table 4.5: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980
	Model 1	Model 1	Model 1
Population	-.028*	-.037*	-.024*
	(-25.30)	(-52.71)	(-25.25)
GDP-per-capita	.193*	.309*	.335*
	(21.92)	(47.28)	(37.30)
Distance	.009*	.014*	.011*
	(4.79)	(11.09)	(6.87)
VAT	.085*	.058*	.068*
	(6.92)	(6.93)	(6.28)
Constant	.013*	.003*	.002*
	(13.95)	(4.44)	(2.59)
Pseudo-R^2(in percentage)	2.6	6.5	5.3
Observations	5555	4917	2810
Number of countries	11	8	7

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. The seven countries in the 1980 sample are: Denmark, France, Germany, Ireland, Italy, the Netherlands, and the UK. Austria is added in 1985, and Greece, Portugal and Spain are added in 1990 as VAT is not observed for these countries except in the 1990 sample.

percent in 1980 down to 19 percent in 1990. The price elasticity of population is estimated to be negative and significant in all cases. VAT differences have positive and rather stable effect on prices from 1980 to 1990.[36] Finally, the estimate of the constant term suggests that the importing country typically had prices 1.3 percent higher than the source at median levels of the explanatory variables in 1990.

Heterogeneity in transport costs across industries

Utilizing information on the source of products sold in any two locations, transportation costs measured by distance are estimated to be important in determining deviations from the law of one price for individual goods. Moreover, distance has been shown to have a robust positive significant impact on absolute price dispersion. Consistent with a number of recent theoretical developments, we consider a specification with industry-specific distance coefficients to explore the relative importance of transportation costs across different industries.[37] Considering industry heterogeneity in transport costs, we show that our estimates of the latter are consistent with

[36] Estimates for the specification with real wage rates are qualitatively similar to those in Table 4.5 with all variables having expected signs and statistically significant. The coefficient estimate for distance ranges from 1.0 percent in 1980 to 0.7 percent in 1990. The impact of VAT on price differences falls from a high of 23 percent in 1980 down to 9.4 percent in 1990.

[37] A number of recent theoretical models consider heterogeneity in transport costs. A recent example is Bergin and Glick (2003) who consider heterogenous trade costs to develop an open economy macroeconomic framework where non-tradedness is an endogenous decision in the face of such good-specific costs.

Table 4.6: Ranking industries according to relative transportation costs

Industry Description:	(1) Value[a]	(2) Tradeability[b]	(3) Distance Coeffs[c]
Transport equipment	**1**	8	5
Machinery except electrical	**2**	3	**1**
Machinery electric	3	10	10
Other manufactured products	4	**2**	6
Professional and scientific equipment	5	**1**	15
Leather products	6	4	8
Furniture except metal	7	18	14
Wearing apparel except footwear	8	9	3
Footwear except rubber or plastic	9	5	4
Rubber products	10	13	9
Miscellaneous petroleum and coal products	11	15	7
Fabricated metal products	12	16	16
Textiles	13	7	13
Printing and publishing	14	**23**	21
Other chemicals	15	11	17
Beverages	16	19	19
Glass and Products	17	20	20
Tobacco	18	21	12
Paper and products	19	12	**2**
Other nonmetallic mineral products	20	**22**	**23**
Food products	21	17	11
Non-ferrous metals	**22**	6	18
Iron and steel	**23**	14	**22**
rank correlation with column (3)	0.59	0.57	1.0

Notes: [a]: Ranking from more expensive to cheaper goods. [b]: Ranking from highly tradeable industries to low tradeability industries. [c]: Ranking of industry-specific distance coefficients from low to high estimated price impact. These coefficient estimates were based on price comparisons relative to the most productive country in each industry for 1985.

common measures of tradeability.

We use a slightly modified version of Model 2 adding industry-specific distance coefficients and excluding the industry-specific tradeability measure from CTZ. We implement this using price comparisons relative to the most productive country in each industry.[38] Once we obtain industry-specific distance coefficients, we rank industries according to the estimated distance coefficient values, with the industry having the lowest distance coefficient ranked first. To see how this ranking relates to other measures of transportation costs, we report in Table 4.6 the ranking of industries according to (1) the average value of goods within that industry classification, with industries ranked from more expensive to cheaper ones, and (2) the degree of tradeability characterizing a certain industry, with industries ranked from highly tradeable to less tradeable ones. To obtain the "value" of the typical good in each industry, we average common currency prices of each good across countries and aggregate this average price across goods that fall in the same ISIC classification. Assuming a fixed component to transportation costs, per unit transportation costs attributed to this fixed component should decline with the value of the good as expensive goods have lower per unit costs. Tradeability is constructed as described in the data section and has a direct interpretation as an inverse measure of effective trade costs.

As long as industry-specific distance coefficients estimates capture relative importance of transport costs across industries, they should be closely related to measures of value and tradeability. Indeed, the correlation between the value ranking in column (1) and the distance coefficient ranking in column (3) is of the right sign at 59 percent and statistically significant beyond the one percent level, and that between the latter and the tradeability ranking is 57 percent. This suggests that the distance coefficient estimates utilized in this paper are indeed closely related to transport costs.

5. Conclusion

Trade costs are important in a number of international macroeconomic models with implications for price deviations across countries. Transport costs are a component of trade costs that has received particular attention in the literature. While technological progress in the transport sector can be expected to reduce their absolute size over time, the relative importance of transport costs can be increasing as policy-related costs of trade decline over time. Moreover, progress in transport technologies might allow previously non traded goods with higher per unit transport costs to enter

[38] As in the case of tradeability and local factor input content measures in Model 2 earlier, industry-specific factors are informative about the absolute level of price dispersion but not about whether a price is higher or lower in a certain geographic location. Thus, we consider absolute price deviations as in Model 2 rather than Model 1-type regressions.

international trade. Thus, the relative importance of transport costs in determining price wedges and international quantity flows might remain even as these are declining for any one good.

To enable us to estimate the costs of trading a good internationally, we rank countries based on their productivity in individual industries and compute product-specific international price differences relative to the most productive location for each industry. We have also used information on realized trade flows to determine the probable source of each product as a weighted average of countries from which a destination country actually imports from. Identifying the source has made it possible to consider price comparisons relevant to the direction of trade and trade costs.

In our application, physical distance relative to the origin has a precisely estimated positive impact on international deviations from LOP that is larger than estimates obtained when arbitrarily assigning an equal probability of being the source to each country. Based on the benchmark specifications for Model 1 in Tables 4.1 and 4.2, the estimated price elasticity of distance was as high as ten percent in 1990. These estimates assign a much bigger importance to distance and trade costs than what is implied by the estimated price elasticities with respect to distance in the literature on international price dispersion.

To conclude, the data are consistent with a model where transport costs and distribution costs are important determinants of international price differences. Our results suggest that utilizing relative productivity and bilateral trade flows along with relative prices from survey data, can help identify trade costs and their role in segmenting product markets. Future work should aspire to utilize microeconomic information on trade flows along with microeconomic relative prices in order to further improve our understanding of trade costs and the relative importance of determinants of international price differences.

Data appendix

Table A1: Industry availability of the TFP level data

Industry Description:	ISIC
Food products	311
Beverages	313
Tobacco	314
Textiles	321
Wearing apparel except footwear	322
Leather products	323
Footwear except rubber or plastic	324
Furniture except metal	332
Paper and products	341
Printing and publishing	342
Other chemicals	352
Miscellaneous petroleum and coal products	354
Rubber products	355
Glass and Products	362
Other nonmetallic mineral products	369
Iron and steel	371
Non-ferrous metals	372
Fabricated metal products	381
Machinery except electrical	382
Machinery electric	383
Transport equipment	384
Professional and scientific equipment	385
Other manufactured products	390

Table A2: Availability of the import flows data

Industry Description:	ISIC
Meat and meat preparations	3111
Dairy products and bird's eggs	3112
Vegetables and fruits	3113
Fish, crustaceans, mollucs, preparations thereof	3114
Margarine, imitat. lard & other prepared edible fats	3115
Fixed vegetable oils and fats	3115
Cereal and cereal preparations	3116
Macaroni, spaghetti and similar products	3117
Bakery products	3117
Sugar and honey	3118
Sugar confectionery and other sugar preparations	3119
Cocoa	3119
Chocolate & other food preptions containing cocoa	3119
Coffee and coffee substitutes	3121
Tea	3121
Spices	3121
Edible products and preparations n.e.s	3121
Alcoholic beverages	3133
Non alcoholic beverages n.e.s	3134
Tobacco and tobacco manufactures	3140
Textile fibres (except wool tops) and their wastes	3210
Textile yarn, fabrics, made up articles, related products	3210
Articles of apparel and clothing accessories	3220
Leather, leather manufactures, n.e.s	3230
Footwear	3240
Furniture and parts thereof	3320
Pulp and waste paper	3410
Paper, paperboard, articles of paper, paper, pulp/board	3410
Registers, exercise books, notebooks, etc	3420
Printed matter	3420
Artificial resins, plastic materials, cellulose esters and ethers	3513
Dyeing, tanning and colouring materials	3521
Essential oils & perfume materials; toilet polishing and cleaning preparations	3523
Chemical materials and products n.e.s	3529
Coal coke and briquettes	3540
Petroleum, petroleum products and realted materials	3540
Rubber manufactures, n.e.s	3550
Other artificial plastic materials, n.e.s	3560
Combs,hair slides and the like	3560

cont.

Glassware	3620
Clay construct. materials & refractory construct. materials	3691
Portland cement, cement fondu, slag cement etc.	3692
Nails,screws, nuts, bolts, etc. iron and steel	3710
Aluminium foil, of a thickness not exceeding 0.20 mm.	3720
Other tools for use in hand	3811
Cutlery	3811
Office machines and automatic data processing equipment	3825
Sewing machines, furniture for sewing mach. & parts	3829
Household type refrigerators & freezers	3829
Telecommunications & sound recording apparatus	3832
Gramophone records, recorded tapes etc..	3832
Household type, elect. & non electrical equipment	3833
Elect. apparel such as switches, relays, fuses, plugs etc.	3839
Batteries and accumulators and parts	3839
Filament lamps, no infra red ultra violet lamps	3839
Int combustion piston engines for outboard prop.	3841
Passenger motorcars,for transport of pass.&goods	3843
Motorcvcles, motorscooters, invalid carriages	3844
Photographic apparatus, optical goods, watches	3850
Medical instruments and appliances	3850
Orthopedic appliances, surgical belts	3850
Pins & needles, fittings, base metal beeds etc.	3900
Children's toys, indoor games	3900
Other sporting goods and fairground amusements	3900
Pens, pencils and fountain pens	3900
Jewelry, goldsmiths and other art. of precious metals	3900
Musical instruments, parts and accessories	3900
Meahanical lighters and parts	3900

References

Anderson, E. James, and Eric van Wincoop (2004) "Trade Costs," *Journal of Economic Literature*, 42(3), 691 – 751.

Anderson, Michael and Stephen Smith, "Borders and Price Dispersion: New Evidence on Persistent Arbitrage Failures," manuscript, Gordon College, April 2004.

Atkeson, Andrew, and Ariel Burstein, "International Relative Prices in New Ricardian Models of International Trade," manuscript, UCLA, November 2004.

Bergin, R. Paul and Reuven Glick, Endogenous Nontradability and Macroeconomic Implications, NBER Working Paper No. W9739, June 2003.

Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum (2003) "Plants and Productivity in International Trade," *The American Economic Review*, 93(4), 1268-1290.

Campbell, Jeffrey and Hugo Hopenhayn (2005) "Market Size Matters," *The Journal of Industrial Economics*, LIII(1), 1-25.

Ceglowski, Janet (2003) "The law of one price: intranational evidence for Canada," *Canadian Journal of Economics*, 36(2), 373–400.

Crucini, Mario, Chris Telmer, and Marios Zachariadis, (2005) "Understanding European Real Exchange Rates," *The American Economic Review*, 95:3, 724-738.

______ "Dispersion in Real Exchange Rates," Vanderbilt University Working Paper 00-w13, 2000.

______ "Price Dispersion: The role of Borders, Distance, and Location," Carnegie Mellon University GSIA Working Paper No. 2004-E25, 2004.

Disdier, Anne-Celia and Keith Head (forthcoming) "The Puzzling Persistence of the Distance Effect on Bilateral Trade", *Review of Economics and Statistics.*

Eaton, Jonathan and Samuel Kortum (2002) "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741-79.

Engel, Charles, and John H. Rogers (1996) "How Wide Is the Border?" *The American Economic Review*, 86(5), 1112-25.

Engel, Charles, John H. Rogers and Shing-Yi Wang (2003) "Revisiting the Border: An Assessment of the Law of One Price Using Very Disaggregated Consumer Price Data" International Finance Discussion Papers 2003-777.

European Commission (2002) "VAT rates applied in the member states of the European Community," Brussels.

Harrigan, James (1997) "Estimation of Cross-Country Differences in Industry Production Functions," *Journal of International Economics*, 47(2), 267-93.

Hummels, David, "Toward a Geography of Trade Costs," unpublished manuscript, University of Chicago, January 1999.

Inkoo, Lee, "Law-of-One-Price in the presence of Commodity Arbitrage: Theory and Evidence," manuscript, Vanderbilt University, 2004.

Melitz, Marc and Gianmarco Ottaviano (2005) "Market Size, Trade, and Productivity," NBER Working Paper Series No. 11393.

Organisation for Economic Co-Operation and Development (OECD), (1998), *International Sectoral Database* (ISDB), Paris, France.

Parsley, David C., and Shang-Jin Wei (1996) "Convergence to the Law of One Price Without Trade Barriers or Currency Fluctuation," *The Quarterly Journal of Economics*, 111(4), 1211-1236.

Parsley, David C., and Shang-Jin Wei (2001) "Explaining the border effect: the role of exchange rate variability, shipping costs, and geography," *Journal of International Economics*, 55 87–105.