DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

International Price Dispersion and the Direction of Trade

Ozlem Inanc and Marios Zachariadis

Discussion Paper 2006-02

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: ++357-2-892430, Fax: ++357-2-892432
Web site: http://www.econ.ucy.ac.cy

International Price Dispersion and the Direction of Trade*

Ozlem Inanc
Department of Economics
Louisiana State University
Baton Rouge, LA 70803
oinanc1@lsu.edu

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

March 2006

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 theoretical framework of Eaton and Kortum (2002) and the variation of this 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: segmented markets, trade costs, transport costs, distribution costs, market size, international price dispersion.

JEL Classification: E0, F4

*We thank Tibor Besedes, Mario Crucini, and participants at the Spring 2005 Midwest International Economics Meetings at Vanderbilt University, the European Cenral Bank-University of Frankfurt Fall 2005 Seminar Series, the University of Cyprus Economics Department Fall 2005 Seminar Series, and the LSU Economics Department Fall 2004 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 about the time-series behavior of changes in these relative prices, 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 that enable exact comparisons across locations. Anderson and van Wincoop (2004) propose the use of price level data that are comparable across locations at a point in time as a promising route for inferring trade cost levels, arguing that "*i*t is hard to see how information can be extracted about the level of trade costs from evidence on changes in relative prices." They go on to suggest that in order 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."

In this paper, we consider microeconomic price levels along with information on the productivity of each country in each industry which we use in order to identify *the* most likely source for each product. This is consistent with the models of Eaton and Kortum (2001, 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. 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.

We consider a variation of the Eaton and Kortum (2002) model proposed in Anderson and van Wincoop (2004). In this framework, international price dispersion is determined by transport costs, local trade (distribution) costs, taxes, good-specific characteristics, and differences in markups. As a measure of transport costs, we use geographic distance and an industry-specific tradeability measure. We account for local trade costs through income per capita differences as in Crucini, Telmer, and Zachariadis (2004) and also consider industry-specific features of local costs as captured by the non-traded factor input content measure used in CTZ and by 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 considering VAT levels for different goods and countries. Finally, we utilize population size to capture market size as an inverse measure of the markup, consistent with the assumption that larger markets tend to be more competitive with demand elasticities higher and markups lower there.

Transport costs and broader trade costs are of central importance in many macroeconomic models.[1] However, assessing these costs 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).

Even when prices of individual products are available across international locations, trade costs will be mismeasured in the absence of information regarding the source of the product being compared across locations.[2] Transport costs would be mismeasured since the distance between the two countries does not necessarily capture distance between exporter and importer. If trade between two countries does not occur for a certain product, then that price difference will lie between the no-arbitrage bounds and will be less than the trade cost.[3] 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 countries can be zero even if these costs are positive and large for each country. A bilateral price difference truly reflects the size of trade costs when only one of any two countries being compared is the source of that product to the other.

In this paper, we aim to resolve the abovementioned problems by utilizing product-specific international price differences along with cross-sectional productivity indices and bilateral trade flows between countries to identify the likely source of any one product. Utilizing the unique -in terms of

[1] For instance, Atkeson and Burstein (2004) consider a theoretical model where trade costs are essential in explaining the time series relation between international relative prices of tradeable goods and the real exchange rate.

[2] This might be behind the finding in Anderson and Smith (2004) and elsewhere of a small or non-existent average impact of transport costs, captured by physical distance, on deviations from LOP.

[3] Since the average trade cost between countries that do not trade with each other is likely to be greater than between those that do, the price gap is likely to be greater between locations that do not trade even though this falls within the bounds determined by trade costs.

breadth of the goods covered and their exact comparability across locations- microeconomic dataset of absolute prices across the European Union 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.[4]

We find that country-specific aspects of transport costs measured by geographic distance, and distribution costs measured by real income per capita, are important in explaining deviations from the law of one price and absolute price dispersion. Industry-specific transport costs as measured by the extent to which a good is traded and industry-specific distribution costs as measured by the local cost content of final products, are shown to be important in determining absolute price dispersion across countries. In addition, market size appears to be an important explanation for international price dispersion. 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, VAT rate differences have been very strong determinants of price differences across the European countries in the sample. However, the impact of these tax differences has been declining throughout the period from 1975 to 1990 as would be expected from the EU policy of tax harmonization. Overall, the data are consistent with models where transport costs and differences in distribution costs, market size, and retail taxes play important roles in the determination of international retail price differences.

Heterogeneity in trade costs and productivity and the interaction between these are central to the price implications of a number of recent papers. For example, in the model of Bergin, Glick, and Taylor (forthcoming), heterogeneity in trade costs and productivity across goods may reverse the usual Balassa-Samuelson effect if the productivity advantage relates to goods with high trade costs.[5] We explore the issue of industry heterogeneity in transport costs and show that our estimates of the latter are consistent with common measures of tradeability. We also allow for productivity heterogeneity across industries. Our findings regarding TFP and the resulting

[4]Trade across these European countries is less likely to be characterized by high policy-related and other unidentified trade barriers, enabling us to better capture transport costs via a geographic distance measure. However, to the extent that transport costs across these countries are relatively less important, our estimates of these are a lower bound for average transport costs characterizing world trade.

[5]A similar implication arises in Benigno and Thoenissen (2003) that allow for home bias and market segmentation so that productivity advantage is consistent with lower domestic prices. This is so since domestically produced goods comprise a larger fraction of domestic consumption than foreign consumption.

identification strategy we implement to identify the source, is consistent with lower product prices in countries that have higher productivity in the industry to which that product belongs.

The remainder of the paper is organized as follows. Section 2 describes the unique dataset of microeconomic prices from CTZ and the construction of cross-sectional TFP indices and trade-weighted relative prices. Section 3 offers the theoretical motivation behind our empirical application pursued there. Section 4 briefly concludes.

2. Data description

Let's denote 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. Then, we can define law-of-one-price deviations as

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

We use the same retail price data as CTZ.[6] A detailed description of the data is provided in the latter paper.[7] 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.[8] 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, Luxembourg, the Netherlands, and the UK. Greece, Portugal and Spain are added in 1980 and Austria in 1985.

Each good was assigned to a three-digit industry to be mapped into 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

[6] We take from CTZ the common currency prices with the outliers having being removed. 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.

[7] A comprehensive list of the goods is available at http://bertha.tepper.cmu.edu/eurostat.

[8] 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.

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.[9] We use three-year averages of tradeability using two preceding years along with the cross-sections sampling years in order to limit measurement error issues.

The distance measure utilized here is the greatest 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 the reported prices are an average from a number of cities within that country. Thus, for Germany, we use distance relative to Frankfurt, a geographic and economic center. 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), the OECD publication "Taxing Consumption", and the Ernst and Young publication "Vat and Sales Taxes Worldwide: A Guide to Practice and Procedures in 61 Countries" (1996).

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 sources for production data are the UNIDO and OECD joint collection program. We obtained from this database value added in current dollars and fixed capital formation, as well as wages and salaries and the number of employees for 28 three digit manufacturing industries. Depending on

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

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.[10] 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. Using capital formation data for twenty-eight manufacturing industries from the Trade and Production Database, we also obtained total manufacturing sector investment. We then obtained each industry's share of 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 then use the observed industry share and total manufacturing capital stock to calculate capital stock for each manufacturing industry. The data appendix provides additional details on the construction of the capital stock.

With the data at hand, and following Harrigan (1997), under the assumption of a Cobb-Douglas production function, TFP between countries j and k for industry h can be described 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 total factor productivity (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 we initially utilize price data for 1985 and 1990 for which TFP 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 could not obtain the data for constructing TFP for services at a disaggregate level and because we are primarily interested in trade costs faced by traded goods.[11] The availability of

[10] We obtain volumes expressed in US dollars as vol_{us} =(VALUK*VALU$_{95}$)/100, where VALUK is the volume index for value added and VALU$_{95}$ is the base year figure for the current price variable. We then obtain the valued added deflator as VALU/vol$_{us}$. Since 1990 is the base year for the capital stock of the manufacturing sector, we use the value added deflator for 1990.

[11] 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

the TFP measure across industries is reported in table A1 in the Data Appendix.

We utilize bilateral trade flows from the OECD *International Trade by Commodity Statistics* (ITCS) database, in order to identify the probability-weighted source for each good sold in each country of the Eurostat price dataset. We are now able to 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 since it also enhances our ability to assess the probable source for each product among a broader group of possible source 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.[12] We end up with imports for each of 42 industries of each country in our sample from each other.[13] 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^h_{jkt} 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 $w^h_{jkt} = im^h_{jkt} / \sum_{k=1}^{n-1} im^h_{jkt}$, 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.

from the law of one price for products that are not traded in equilibrium.

[12] There is a many to one mapping from goods for which we have prices to the 4-digit categories in the trade data. Ideally, future work should focus on more disaggregated trade data that can be closely matched to the products in the price surveys. However, this labor intensive task would face two inherent problems. First, for disaggregated products the problem of "empty cells" is a greater concern. Second, the measurement error is greater for highly disaggregated product categories relative to aggregates.

[13] 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.

To cope with this, we re-scale the weights.[14] 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-weighted source. The same weights are used in order to construct the 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 $w_{jkt}^{h} = im_{jkt}^{h} / \sum_{k=1}^{n} im_{jkt}^{h}$, 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: $p_{j\kappa t}^{h} = \sum_{k=1}^{n} w\_new_{jkt}^{h} * p_{kt}^{h}$, 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.

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

3. Motivation and Estimation

Theoretical 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 i buys from country κ, then $p_i = c_k \tau_{i\kappa}$, where c_κ is the cost of production in κ and $\tau_{i\kappa}$ is the trade cost of transporting the good from κ to i. Moreover, country i will buy from κ if $c_\kappa \tau_{i\kappa}$ is the lowest among all potential sources. The inequality thus derived is $\frac{c_{z_i}\tau_{iz_i}}{c_{z_i}\tau_{jz_i}} \leq \frac{p_i}{p_j} \leq \frac{c_{z_j}\tau_{iz_j}}{c_{z_j}\tau_{jz_j}}$ or $\frac{\tau_{iz_i}}{\tau_{jz_i}} \leq \frac{p_i}{p_j} \leq \frac{\tau_{iz_j}}{\tau_{jz_j}}$, where p_i and p_j are retail prices in country i and j, and z_i and z_j are the optimal sources for country i and j respectively. When countries i and j purchase the good from the same source, κ, then the above inequality is reduced to $\frac{p_i}{p_j} = \frac{\tau_{i\kappa}}{\tau_{j\kappa}}$, 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."

That is, once we identify the probable source country then we can capture the exact level of trade costs.[15] 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 (2001), 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. 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),

[15] Given the absence of product-specific source information our aim is necessarily less ambitious. We estimate an improved measure of the relative importance of two components of broadly defined trade costs: transport costs and distribution costs, while controlling for other potential determinants of international relative prices.

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 and the problems associated with cif/fob ratios in levels as discussed in Hummels and Lugovskyy (2003), we follow the usual practice of using physical distance between the capital cities of the countries in our sample to capture transportation costs. 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[16] 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.[17] 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, we assume that larger markets are more competitive so that demand elasticities are higher and markups lower there, and use 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 smaller ones thus exhibiting greater risk aversion for losing large markets, and are less likely to charge higher prices there in the presence of demand uncertainty. On the other hand, population size might capture scale economies that simply lower the average domestic cost of production leading to lower domestic prices. However, the scale of domestic producion

[16] Since this industry-specific measure is based on realized trade flows, it might partly capture industry-specific trade costs other than tranport costs. Moreover *industry-specific* measures can only be considered as determinants of *absolute* price deviations, since *actual* price deviations are related to the *direction* of trade across countries and can only be explained by factors that have variation across countries.

[17] 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.

also 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 the domestic distribution or retail sector. In any case, given the difficulty of capturing variation in markups across countries, an alternative starting assumption would be that markups exist but are similar across countries so that they do not impact on international price differences. This assumption is imposed in Crucini, Telmer, and Zachariadis (2004) and discussed in Anderson and van Wincoop (2004). In that case, the coefficients of population size differences would be interpreted instead as measures of scale economies common across industries and specific to countries.

Estimation and results

Based on the above, we expect that the price difference between the importing location and the source country for a particular final product would be largely determined by transport costs and international differences in local distribution costs, taxes, and markups. Thus, we attempt to infer estimates of the impact of each potential determinant of international price differences by utilizing physical distance as a measure of the importance of transport costs, income per capita or domestic industry-specific real wage rates as measures of the local cost component comprising the price of final goods, and population size as capturing differences in markups, also allowing where possible for VAT differences across industries and countries. In addition, for the absolute price differences specifications we are able to consider product category-specific differences in taxes and industry-specific measures of tradeability and the non-traded factor input content to capture the importance of industry-specific transport costs and local distribution costs respectively.

As a first step, we consider the following regression equation for *all* possible *unique* bilateral price comparisons j-k

$$q_{ijk} = a_0 + a_1 Dist_{jk} + a_2 y_{jk} + a_3 Pop_{jk} + \varepsilon_{ijk} \tag{3.1}$$

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[18], and ε_{ijk} is a random error[19]. $Dist_{jk}$ is the (log) distance separating the

[18] All explanatory variables are demeaned so that the constant can be interpreted as the price deviation relative to source k at average levels of distance, real GDP per capita, and population size in the sample.

[19] 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^2's reported here should be taken with caution.

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 GDP per capita between j and k and captures the local 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.[20] In this sense, GDP per capita is a measure of the local distribution costs discussed above. 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 also consistent with gravity models used to assess international quantity flows.

In considering *all* possible *unique* bilateral price comparisons j-k, we compare each pair of countries once with each bilateral comparison made based on alphabetical order rather than relative to countries more likely to be a source for the product. This is then an arbitrary comparison using no information regarding the source of each product and renders the coefficient of geographic distance proxying for transport costs meaningless. This case will be a reference point with which to compare trade cost estimates obtained utilizing information on the probable source of each product.

Estimates and t-statistics from estimating the above specification (Model 1) with OLS and correcting standard errors for the inherent heteroskedasticity are presented in Table 3.1. The distance coefficient is estimated to be negative and statistically indistinguishable from zero for 1985 and equal to 5.5 percent and significant in 1990. Considering all possible bilateral comparisons tends to average out around zero the impact of transportation costs on prices producing unreliable estimates. The estimated coefficient of distance is perhaps devoid of meaning here as distance between two arbitrary countries does not necessarily capture distance between exporter and importer. If trade between two countries does not occur for that product, then that price difference will lie between the no-arbitrage bounds and will be less than the trade cost. Moreover, when comparing two countries it is possible that both export some of the same products to each other. To the extent that this is the case, the final price for these products will incorporate a similar transportation cost in both countries so that there might be little or no impact of transportation costs on the price difference for these products between the two countries.[21] In general, in the absence of some

[20] Crucini, Telmer, and Zachariadis (2004) explore the relation between distribution costs and GDP per capita.

[21] It is also possible that k is the main exporter to j for some product i and does not import this product from j, and that j is the main exporter to k for some product i' and does not import this product from k. In the first case, this would induce the distance coefficient to be positive as transport costs increase the price in country j relative to k

Table 3.1: All unique bilateral Comparisons

	1990			1985		
Model:	1	2	3	1	2	3
Population	-.049*	.007**	.011	-.066*	.006**	.003
	(-17.14)	(2.51)	(1.34)	(-22.54)	(1.96)	(0.29)
GDP-per-capita	.283*	.047*	.132**	.295*	.036**	.036
	(17.20)	(2.91)	(2.60)	(19.02)	(2.46)	(0.76)
Distance	.055*	.075*	.095*	-.009	.100*	.116*
	(6.40)	(11.97)	(4.58)	(-0.98)	(15.79)	(5.69)
Tradeability		-.087*	-.076*		-.057*	-.089*
		(-11.46)	(-4.22)		(-7.24)	(-5.36)
Non traded input share		.003*	.007*		.010*	.011*
		(2.63)	(2.66)		(9.12)	(3.94)
Large cars		.143*			.255*	
		(7.67)			(6.95)	
Vices		.218*	.194*		.227*	.172*
		(12.83)	(5.69)		(13.51)	(4.99)
Constant	.097*	.359*	.314*	.076*	.274*	.311*
	(24.00)	(29.29)	(8.54)	(18.65)	(21.78)	(8.93)
R^2(in percentage)	3.5	6.3	23.2	4.1	5.8	16.9
Observations	12315	12315	473	13995	13995	530
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.

information regarding the source of each product and the direction of trade, the distance coefficient will not capture transport costs well in the context of "directional regressions" such us the one in Model 1.

GDP per capita and population enter in expected ways in Model 1. Per capita GDP shows a strong positive relation with price differences between countries. The price elasticity of real GDP per capita is 29.5 percent for 1985 and 28.3 percent for 1990, exhibiting remarkable stability over this five year period. Moreover, higher population is associated with lower prices in a country suggesting a potential role for markup differences across countries due to differences in demand elasticities that are positively related to the size of the market. In this case, the markup would be lower in larger markets as evident in the negative estimated coefficients for population size. Alternatively, scale economies in distribution related to the domestic size of the market might be behind this finding.

while in the second case the distance coefficient would be negative. The overall result is a possible washing out of the average effect of transport costs across goods. This is related to the "averaging-out" property discussed in Crucini, Telmer, and Zachariadis (2004) and can be addressed by considering absolute price differences for each product across countries or an appropriate variance measure. We pursue this in Model 2 below.

Model 2 describes the relation between absolute price differences and the absolute values of the variables that are included in Model 1 as well as additional industry-specific variables like tradeability and the local factor input content of goods in each industry. Taking absolute values of the price differences serves three purposes. First, it allows us to use distance as a meaningful determinant of (absolute) price dispersion even in the absence of source country information. This is the case since it resolves the "averaging-out" problem, as pointed out by Crucini, Telmer, and Zachariadis (2004). Second, it allows us to consider the two industry-specific variables from CTZ which are closely related to a theoretical model where final goods are produced by combining local inputs with traded inputs. We would 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. Finally, we can now introduce two 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 would expect such goods to be characterized by a higher degree of absolute price dispersion.

Thus, we estimate the following regression equation for Model 2:

$$|q_{ijk}| \ = a_0 + a_1 Dist_{jk} + a_2|y_{jk}| + a_3|Pop_{jk}| + a_4 X_h + \varepsilon_{ijk} \tag{3.2}$$

where X_h is a vector of industry-specific and category-specific variables capturing product characteristics as described above. The remaining variables are defined as in regression equation 3.1. The constant a_0 now captures price dispersion at mean distance, real GDP per capita, and population size in the sample. The 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 (tradeability) matter -about equally- for absolute price dispersion.

Furthermore, higher local input share implies higher absolute price dispersion as the model discussed earlier would predict. Moreover, income per capita differences enter as a positive determinant of price dispersion, suggesting that both country-specific and industry-specific aspects of distribution costs matter for absolute price dispersion. However, the estimated impact of income on *absolute* relative prices across countries is many times smaller than its impact on the actual *level* of relative prices. By considering absolute price differences we might be underestimating the importance of the local cost component in determining price levels. In this case, the gains made in terms of estimating the transport cost component of trade costs using absolute price dispersion in Model 2 would appear to be a loss in terms of our ability to estimate the distribution costs component of trade costs.

Finally, population coefficient estimates suggest absolute price dispersion increases with differences in population size, indicating a possible role for markup differences as determinants of international price dispersion. The dummies for large cars and vices also 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.

For Models 1 and 2, goodness-of fit measures are very low. Price data are more disaggregated than explanatory variables, therefore the R^2 is not meaningful for these models. As explained in CTZ, 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 follow CTZ and aggregate the data. Specifically, 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 3.1. The goodness-of-fit increases substantially for both cross sections. The coefficient estimates for most of the variables are similar qualitatively to those reported for Model 2. The estimated coefficient for distance is positive and significant in both cross-sections for Model 3 as was the case in Model 2, while the estimated coefficients for local costs are generally higher than in Model 2. The estimated coefficient for category-specific taxes is about the same as in Model 2 in the case of vices. However, since in Model 3 we aggregate according to 3-digit ISIC category, the dummy for "large cars" has not been included in this regression since this product category is one of several in category 384 which includes all transport equipment.

Table 3.2: All unique bilateral Comparisons

	1990			1985		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Population	-.041*	.009*	.015***	-.043*	.007**	.005
	(-15.58)	(3.39)	(1.91)	(-16.39)	(2.50)	(0.58)
Real Wage Rate	.148*	.061*	.106*	.098*	.055*	.082*
	(20.05)	(7.41)	(4.50)	(13.14)	(7.49)	(3.78)
Distance	.035*	.056*	.068*	-.005	.075*	.075*
	(4.04)	(8.76)	(3.28)	(-0.54)	(11.10)	(3.52)
Tradeability		-.088*	-.077*		-.058*	-.092*
		(-11.62)	(-4.28)		(-7.39)	(-5.75)
Non-tradable input share		.003*	.007*		.010*	.010*
		(2.73)	(2.72)		(9.14)	(4.07)
Large cars		.150*			.259*	
		(8.07)			(7.00)	
Vices		.219*	.196*		.224*	.167*
		(12.97)	(5.86)		(13.39)	(4.94)
Constant	.098*	.358*	.314*	.076*	.274*	.318*
	(24.25)	(29.32)	(8.78)	(18.53)	(21.91)	(9.38)
R^2(in percentage)	4.6	6.8	25.6	2.9	6.2	19.1
Observations	12315	12315	473	13995	13995	530
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.

As a robustness check and to account for broader local costs (including production 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 but specific to each industry. The fact 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 exercuse also offers a robustness check for our coefficient estimates for distance, tradeability, and industry-specific local input costs. We report results utilizing wage rates in Table 3.2. We can see that the real wage rate has positive impact on price differences in Model 1 and on absolute price differences in Models 2 and 3. The wage impact on prices is about half the GDP impact for Model 1 but larger than the GDP impact for Model 2, and more robust than the GDP impact for Model 3. We also see that the coefficient estimates for the industry-specific measures of tradeability and the local factor input content are virtually unchanged. Finally, the estimates for the distance coefficient are qualitatively similar but smaller across the board for all three Models and both years relative to the specifications that include GDP per capita in Table 3.1. This might suggest that real wage rates capture an aspect of local production costs that would otherwise be in part attributed to transport costs.

Utilizing information on relative productivity

Overall, the results for Models 2 and 3 summarized in Table 3.1, indicate that there is a positive and significant relation between distance and absolute price dispersion. However, the interpretation of the coefficients related to transport costs can be problematic for the reasons outlined in the previous section and in the introduction. Moreover, as shown in Table 3.1 for Model 1, the effect of distance on price differences is estimated to be statistically indistinguishable from zero for 1985 for instance, perhaps pointing to the argument put forth by Anderson and van Wincoop (2004). That is, without knowing the potential source for a good we cannot estimate the precise role of transportation costs in determining differences in the price levels for that good between countries.[22] One way to address the problem is to assume that the *more* productive among any two countries being compared will export the good to the other country.[23] A problem with this approach would be that given the measurement error associated with TFP construction, comparing countries with similar productivity is likely to often give the wrong ordering simply because of measurement error. A related and preferable method, is to consider price comparisons only relative to the *most* productive country in the data, to avoid an ordering based on comparisons among countries that are closer together in tems of productivity.

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 can then construct the good-specific log relative prices between each country j relative to the main exporter country κ for each industry h.

Admittedly, this approach does not fully resolve the problem of identifying the source country

[22] One approach would be to just assume one of the countries to be the main exporter using a-priori information. This is unsatisfactory conceptually for obvious reasons and, as one would expect, this approach does not give reliable results. Table A3 in the appendix reports the estimation results for Germany or the U.K. used as reference countries in each case. The sign and significance of the distance coefficients are not robust across periods or reference countries.

[23] Thus, one could consider adding to Model 1 an interaction term between the inverse of the productivity difference and distance across any two countries. This would capture the idea that for each bilateral comparison, the less productive country will be importing the product from the more productive country and thus have higher prices than the latter country according to the extent of transportation costs present. Implementing this, we obtain consistently positive but small estimated coefficients for this measure, with relatively large standard errors.

for each good in our price sample since our measure of productivity is at the three-digit level and suffers from an obvious aggregation bias. 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 next section 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 the EU countries in our sample. However, as we show next, the current methodology goes some distance into identifying the source country and thus providing a meaningful measure of transport costs.

Before turning to estimation using price differences relative to the most productive country, we attempt to evaluate the hypothesis that productivity is inversely related to prices, consistent with productivity being a determinant of the direction of trade. We consider a specification similar to (3.1) adding now a term for productivity differences across countries:

$$\text{1985:} \quad q_{ijk} = \underset{\underset{(18.60)^*}{[.076]}}{a_0} + \underset{\underset{(-0.90)}{[-.008]}}{a_1} Dist_{jk} + \underset{\underset{(19.56)^*}{[.318]}}{a_2} y_{jk} + \underset{\underset{(-23.02)^*}{[-.068]}}{a_3} Pop_{jk} + \underset{\underset{(-3.48)^*}{[-.035]}}{a_4} TFP_{hjk} + \varepsilon_{ijk}$$

$$\text{1990:} \quad q_{ijk} = \underset{\underset{(6.49)^*}{[.098]}}{a_0} + \underset{\underset{(6.63)^*}{[.057]}}{a_1} Dist_{jk} + \underset{\underset{(17.78)^*}{[.325]}}{a_2} y_{jk} + \underset{\underset{(-17.84)^*}{[-.051]}}{a_3} Pop_{jk} + \underset{\underset{(-4.40)^*}{[-.042]}}{a_4} TFP_{hjk} + \varepsilon_{ijk}$$

TFP_{hjk} is the difference in total factor productivity across countries j and k for industry h, where industry h is a three-digit classification with a one-to-many mapping into individual goods i. Above, we report the estimates and t-statistics for the variables in this regression for 1985 and 1990. The estimates for TFP suggest a negative impact on prices. These estimates suggest the relevance of productivity in determining the *direction* of international trade and as a result international price differences. The negative impact of TFP is also consistent with the theoretical model of Benigno and Thoenissen (2003) and parameterizations of the model in Bergin, Glick, and Taylor (forthcoming).

Given that TFP is a determinant of the direction of price differences across countries, we now go ahead to consider the following regression equation:

$$q_{ij\kappa} = a_0 + a_1 Dist_{j\kappa} + a_2 y_{j\kappa} + a_3 Pop_{j\kappa} + \varepsilon_{ij\kappa} \tag{3.3}$$

where $q_{ij\kappa}$ is the log deviation from the Law of One Price (LOP) for good i between country j and κ,

Table 3.3: 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* (-11.14)	-.010*** (-1.90)	-.0003 (-0.02)	-.056* (-13.91)	-.001 (-0.28)	.015 (1.27)
GDP-per-capita	.422* (12.33)	-.048 (-1.18)	-.009 (-0.08)	.280* (11.84)	.027 (1.05)	-.104 (-1.20)
Distance	.090* (6.61)	.064* (6.13)	.093** (2.08)	.041* (2.98)	.100* (10.00)	.130* (3.94)
Tradeability		-.061* (-3.86)	-.034 (-0.93)		-.046* (-3.04)	-.063** (-2.13)
Non-tradable input share		.008* (3.49)	.006 (1.47)		.010* (4.74)	.005 (1.14)
Large cars		.084* (2.97)			.147* (2.70)	
Vices		.187* (5.96)	.151* (2.88)		.162* (5.99)	.089** (2.19)
Constant	.047* (6.54)	.255* (10.85)	.290* (4.92)	.016** (2.25)	.228* (9.83)	.311* (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.

the most productive country in industry h assumed to be the main source for product i in country j. To estimate equation (3.3) we utilize the industry-specific country ranking implied by cross-sectional TFP levels in constructing the dependent variable of prices relative to the most productive location. Again, $\mathrm{Pop}_{j\kappa}$ and $y_{j\kappa}$ are the population and real GDP per capita log differences between countries j and κ respectively, $\mathrm{Dist}_{j\kappa}$ denotes the log distance between source κ and destination j, and $\varepsilon_{ij\kappa}$ is a random error. As the explanatory variables are demeaned, the constant a_0 captures the price deviation relative to source κ at average levels of distance, real GDP per capita, and population size in the sample. Regression equation (3.3) incorporates information regarding the direction of trade and can thus assist in inferring the overall level of trade costs and the level of the transport costs component of trade costs as the estimated coefficient for physical distance. Results from this estimation framework are summarized in Table 3.3.

Model 1, the first specification of Table 3.3, indicates that distance has a positive and significant impact on international price differences, suggesting a role for transportation costs as a determinant of these. Based on the 1990 estimates, a doubling in distance would lead to an increase in prices of

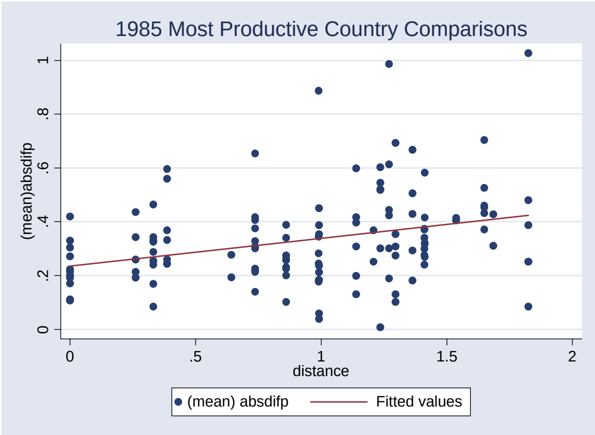

Figure 3.1:

9 percent, substantially greater than the 5.5 percent increase for the specification with all unique bilateral price comparisons in Table 3.1. The improvement in terms of the estimated distance coefficient is even more striking for 1985. Comparing Model 1 across Tables 3.1 and 3.3, we see that the estimated coefficient of distance changes sign becoming positive and strongly significant once we account for the probable source of the traded 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 distance and transport costs to a considerable degree.

In addition, local costs as captured by real GDP per capita appear to have a strong effect 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. This would suggest that markups are about 5 percent lower in larger countries.

Next, we utilize absolute law of one price deviations relative to the most productive country to estimate a specification similar to (3.2). This formulation allows us to consider the impact of good-

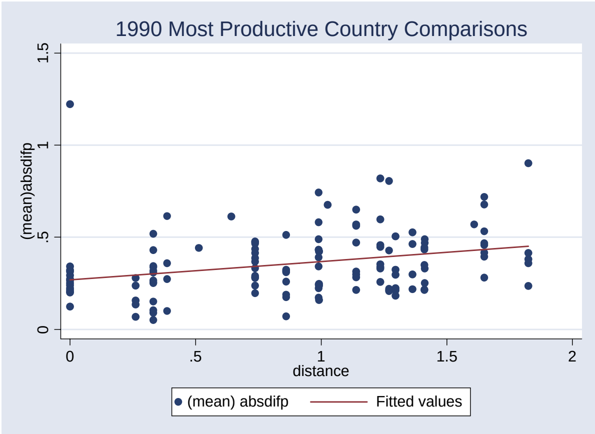

Figure 3.2:

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 Crucini, Telmer, and Zachariadis (2004). This also allows aggregation into mean absolute price differences (in Model 3) which allows us to obtain 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 3.1 and 3.2 respectively. The visual evidence supports a positive relation between these two variables.

The estimates for Models 2 and 3 are reported in Table 3.3. The distance coefficient always has a positive significant impact on absolute price differences. However, for Models 2 and 3, there appears to be little gain in terms of the effect of distance on absolute price differences relative to the estimates utilizing all unique bilateral price comparisons reported in Table 3.1. This is in contrast to the significant gains achieved when we utilize the productivity information to identify the source in Model 1. Accounting for industry-specific productivity resolves some of the problems associated with the lack of information on the source of each product, so that considering absolute price deviations in Models 2 and 3 does not have as much of an additional impact on the distance coefficient in addition to the gains achieved in Model 1. The remaining parameter estimates are for the most part similar to those for Models 2 and 3 in Table 3.1, with the exception of population which is now estimated to have a small negative impact on absolute price dispersion in Model 2 for 1990, and GDP per capita that is now statistically indistinguishable from zero for both Models and both cross-sections. The latter finding suggests that once we consider comparisons relative to

Table 3.4: 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.

the most productive country, higher income differences are no longer associated with higher price dispersion. That is, 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.[24] Finally, the constant a_0 captures price dispersion relative to the source with average levels of distance, real GDP per capita, and population size in the sample. This is equal to 31 percent in 1985 and 29 percent in 1990.

In Table 3.4, we re-estimate Models 1, 2, and 3 replacing GDP per capita by wage rates that vary both across industries and countries. In Model 1, wage differences are positively associated with price differences with price elasticities of 15 percent for 1985 and 20.9 percent for 1990. 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 per capita is included instead of wage rates. The distance coefficient is now estimated to be ten

[24] The small sample of relatively similar income countries considered here and the resulting small variation in income for these data might be the reason behind the latter finding.

percent for 1990 and seven percent for 1985, compared to nine and four percent respectively in the estimations presented in Table 3.3 utilizing GDP per capita. Comparing these estimates of distance with the ones obtained using all unique bilateral comparisons in Table 3.2, we see that these are now considerably larger. For 1990, the distance coefficient point estimate was equal to 3.5 percent while for 1985 this was negative and statistically indistinguishable from zero. The improvement in terms of estimating the distance coefficient using the most productive country comparisons, is thus even more pronounced when we include wage rates instead of GDP per capita.

Estimates of the variables in Models 2 and 3 in Table 3.4 are qualitatively similar to those in Table 3.3. Again, the population size coefficient is estimated to have the wrong negative sign in Model 2 for 1990. 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 3.3. However, for these absolute price comparisons the coefficient estimates for distance become smaller relative to the specification with GDP per capita. Finally, price dispersion relative to the source at average levels of distance, real GDP per capita, 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 3.3.

Finally, for 1990 for which we have VAT data for all countries in our sample, we reconsider Models 1 to 3 for the specification with all bilateral price differences and the one relative to the most productive country, adding now VAT log differences as an explanatory variable on the RHS. VAT is not observed for Greece, Portugal, and Spain except in the 1990 sample. For this reason, we do not consider VAT for 1985 since this would reduce our small sample to just five countries, and further limit our ability to "guess" the probable source and destination countries for each industry.[25] We report results in Tables 3.5 and 3.6 utilizing GDP per capita and wage rates respectively. For Model 1, the estimated coefficient for VAT differences is positive, very high, and strongly significant. The remaining estimates we obtain are for the most part similar to those in Tables 3.1 to 3.4. For the specification using all bilateral comparisons, the distance coefficients for Model 1 are virtually unchanged at 5.2 and 3.4 percent relative to 5.5 and 3.5 percent in the specifications without the VAT variable reported in Tables 3.1 and 3.2 for the specifications with GDP and wages respectively. However, the estimated distance coefficient in Model 1 for the specification using price comparisons

[25] Ideally, we would like the maximum possible number of countries so that the most productive country in our sample will be more likely to be the source in the actual trade data.

Table 3.5: 1990 with VAT

	All Unique Bilateral Comparisons			Most Productive Country		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Population	-.029*	.005***	.010	-.034*	- .008	- .003
	(-9.01)	(1.78)	(1.17)	(-6.74)	(-1.52)	(-0.21)
Real GDP per capita	.210*	.053*	.133*	.336*	- .062	.002
	(12.37)	(3.31)	(2.63)	(9.47)	(-1.56)	(0.01)
Distance	.052*	.074*	.095*	.064*	.062*	.091**
	(6.14)	(11.86)	(4.60)	(4.74)	(5.99)	(2.01)
VAT	1.02*	.440*	.173	.735*	.395*	.214
	(15.38)	(7.46)	(0.80)	(6.44)	(3.77)	(0.58)
Tradeability		-.080*	-.076*		-.052*	-.034
		(-10.50)	(-4.17)		(-3.33)	(-0.92)
Non-tradable input share		.003*	.007*		.008*	.006
		(2.71)	(2.64)		(3.48)	(1.45)
Large cars		.127*			.075*	
		(6.76)			(2.61)	
Vices		.228*	.195*		.200*	.153*
		(13.41)	(5.71)		(6.34)	(2.88)
Constant	.099*	.354*	.316*	.047*	.251*	.293*
	(24.57)	(28.82)	(8.57)	(6.57)	(10.65)	(4.96)
R^2(in percentage)	5.5	6.8	23.3	6.0	5.4	15.8
Observations	12315	12315	473	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.

relative to the most productive country now falls to 6.4 percent in Table 3.5 and to 7.7 percent in Table 3.6, relative to 9.0 and 10.1 percent in Tables 3.3. and 3.4. Although lower than prior to the inclusion of VAT differences, these estimates are still higher than those obtained using all bilateral comparisons. Finally, for Models 2 and 3, the distance coefficients before and after the inclusion of VAT differences are nearly unchanged and so are the coefficient estimates for tradeability and the local input content, while population size is again estimated to have the wrong negative sign for Model 2 in 1990.

Utilizing trade flows

Assuming the most productive country in an industry to be the sole exporter of goods of that industry to the countries in our sample does not completely resolve the problem of identifying the source. It is possible that a similar 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

Table 3.6: 1990 with VAT

	All Unique Bilateral Comparisons			Most Productive Country		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Population	-.024* (-8.56)	.007* (2.65)	.013 (1.61)	-.016* (-4.10)	-.010** (-2.20)	-.007 (-0.54)
Real Wage Rate	.124* (16.56)	.072* (8.63)	.109* (4.67)	.178* (10.09)	.055** (2.54)	.113** (2.14)
Distance	.034* (4.01)	.052* (7.99)	.067* (3.23)	.077* (5.39)	.035* (2.85)	.042 (1.02)
VAT	1 .01* (15.72)	.516* (8.64)	.310 (1.47)	.820* (7.32)	.454* (4.07)	.503 (1.34)
Tradeability		-.080* (10.53)	-.077* (-4.22)		-.054* (-3.43)	-.029 (-0.78)
Non-tradable input share		.003* (2.84)	.007* (2.68)		.008* (3.60)	.007*** (1.72)
Large cars		.134* (7.10)			.080* (2.86)	
Vices		.232* (13.68)	.198* (5.95)		.200* (6.41)	.155* (3.09)
Constant	.099* (24.83)	.352* (28.80)	.318* (8.83)	.048* (6.70)	.248* (10.50)	.281* (4.88)
R^2(in percentage)	6.7	7.4	25.9	6.9	5.6	18.5
Observations	12315	12315	473	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.

sample so as to take into consideration that the same type of good can be exported by more than one country in the sample.

However, the goods could also be imports from countries other than 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. For instance, in 1990, the share of imports from European Union (EU) countries for our sample is 84 percent for "furniture except metal industries" but only 51 percent for "tobacco and tobacco product industries." Moreover, the import share from the EU varies between countries for the same industry. For example, in 1990 the share of EU imports for France, Italy and Greece in "tobacco and tobacco product industries" is higher than 90 percent, whereas the share for Denmark is 11 percent and for Spain only 8 percent. This tells us that, for some countries and industries, important exporters are outside the EU sample we have price data for. In order to alleviate this problem, we consider 50 percent as a cutoff point for the fraction of imports from the EU by each

Table 3.7: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980	1975
	Model 1	Model 1	Model 1	Model 1
Population	-.044*	-.056*	-.025*	-.037*
	(-12.28)	(-13.66)	(-5.10)	(-8.36)
GDP-per-capita	.251*	.279*	.292*	.268*
	(12.48)	(11.35)	(9.24)	(8.09)
Distance	.044*	.047*	.044*	.080*
	(5.28)	(4.97)	(4.01)	(6.03)
Constant	.042*	.056*	.037*	.021*
	(8.45)	(10.25)	(5.42)	(3.63)
Pseudo-R^2(in percentage)	1.9	2.4	1.3	3.4
Observations	6848	7322	3392	2759
Number of countries	12	12	11	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. Greece, Portugal and Spain are added in 1980, and Austria is added in 1985.

country in each industry. That is, for each importer and industry, the ratio of imports from the EU over total imports is constructed and if this is lower than the 50 percent cutoff point, the good belonging to that industry is dropped from the dataset. This approach increases the likelihood that a certain good we consider in the price comparisons is actually imported from an EU country. The advantage of this approach, is that for these goods we can better identify the source and thus estimate more precisely transport costs relevant to our sample of countries.

We proceed to utilize realized trade flows among the countries in our sample in order to determine the direction of trade and construct price comparisons for each product consumed in the importing country relative to countries that are 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 the 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.

Once again, we estimate an equation similar to equation (3.3) where source κ is now a weighted sum of probable exporters and these probabilities are obtained as described above. In Table 3.7, we report estimates from this specification. The price data have already been cleansed of outliers

Table 3.8: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980
	Model 1	Model 1	Model 1
Population	-.049*	-.060*	-.010***
	(-10.79)	(-14.69)	(-1.85)
Real Wage Rate	.141*	.147*	.134*
	(10.58)	(13.08)	(9.25)
Distance	.038**	.040*	.052*
	(2.54)	(2.56)	(2.89)
Constant	.030*	-.001	.001
	(5.06)	(-0.16)	(0.10)
Pseudo-R^2(in percentage)	2.7	4.1	2.0
Observations	5910	5423	2766
Number of countries	11	10	10

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. The ten countries in the 1980 sample are: Denmark, France, Germany, Ireland, Italy, the Netherlands, the UK, Greece, Portugal and Spain. Austria is added starting in 1985. The Netherlands is not included in the 1985 cross-section because it's missing the wages data for many industries.

following CTZ. However, the trade quantities used here introduce an additional source of outliers given the well known measurement problems with measures of trade flows. Thus, in order to handle outliers, we minimize an absolute loss function and obtain the median estimator so that coefficients are estimated by minimizing absolute deviations from the median rather than squared deviations from the mean. Since as an estimate of central tendency the median is not as greatly affected by outliers as the mean, this alleviates the outliers problem.[26,27]

The estimated coefficients for distance reported in Table 3.7 are estimated precisely and are always positive for 1975, 1980, 1985, and 1990. The estimated price elasticity of distance is as high as 8.0 percent in 1975 but declines down to 4.4 percent by 1990. 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 in Table 3.1, but are qualitatively similar to those obtained assuming the most productive country in an industry to be the exporter for products of that industry.

[26] We also tried the Cook's D criterion to identify outliers which are then assigned smaller weights relative to other observations using iteratively reweighted least squares robust regressions. This method assigns a weight to each observation, with well behaved less influential observations assigned higher weights, and only very extreme outliers completely removed from the sample. Results were very similar to those in Table 3.7.

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

Table 3.9: 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 estimates for the impact of the local cost component of trade costs reported in Table 3.7 are positive and precisely estimated for each year in our sample, with a price elasticity ranging from about 29 percent in 1980 down to about 25 percent in 1990. The size of the population is consistently estimated to have a negative impact on prices with an estimated negative price elasticity, ranging between 2.5 percent in 1980 and 5.6 in 1985. Finally, the estimate of the constant term tells us that the importing country typically had prices which were 4.2 percent higher than the source at median levels of the explanatory variables in 1990.

As a robustness check to the use of GDP per capita, we utilize industry-specific real wage rates and report corresponding results in Table 3.8. Here, we do not consider the 1975 cross-section since the wage measure is not available for that year. As expected, the real wage rate has a strong positive impact on prices while population enters negatively in all cross-sections. The estimated price elasticity of distance ranges from a high of 5.2 percent in 1980 down to 3.8 percent in 1990.

Finally, we consider VAT differences as an additional explanation of price differences across countries and report results for this specification in Table 3.9. 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. The estimated effect of distance ranges from a high of 7.5 percent in 1975 down to 3.1 percent in 1990. Similarly, the price elasticity

Table 3.10: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980
	Model 1	Model 1	Model 1
Population	-.035* (-7.03)	-.045* (-9.23)	.006 (1.45)
Real Wage Rate	.111* (7.99)	.080* (4.28)	.119* (7.90)
Distance	.031** (2.01)	.048** (2.49)	.063* (4.31)
VAT	.615* (5.74)	.869* (7.86)	.893* (7.77)
Constant	.023* (3.86)	.015** (2.22)	.009 (1.58)
Pseudo-R^2(in percentage)	3.1	7.6	3.5
Observations	5910	3975	2164
Number of countries	11	7	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 starting in 1985 while the Netherlands is excluded for that year. Greece, Portugal and Spain are added in 1990.

of the local component of distribution costs captured by GDP per capita is estimated positive and significant for all cross sections. The impact of population size on prices is again negative across the board.

When we use the real wage rate instead of real GDP per-capita, for Table 3.10, VAT again has strong but declining positive effect on prices for all years, ranging from 89 percent in 1980 down to 61.5 percent in 1990. Similarly, the real wage rate has strong positive effect on price differences for all years. On the other hand, population enters negatively and significantly for 1985 and 1990 but positively and statistically insignificant for 1980. As usual, the effect of distance decreases monotonically by more than half, from 6.3 percent in 1980 down to 3.1 percent in 1990.

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 at home can be an import or produced domestically. *Domestically consumed production* of country i for industry h is defined as the difference between total output of country i for industry h and exports of country i for that industry. As we did previously, in order to increase the likelihood that a certain good we consider in the price comparisons is actually imported from an EU country, we consider a within-sample import ratio of 50 percent as a cutoff point. Results are reported in Tables 3.11 to 3.14.[28]

[28] We cannot use the year 1975 since we do not have total output data for these countries. We also note that since

Table 3.11: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980
	Model 1	Model 1	Model 1
Population	-.031*	-.036*	-.025*
	(-31.97)	(-55.32)	(-27.85)
GDP-per-capita	.202*	.277*	.268*
	(25.18)	(67.03)	(37.88)
Distance	.011*	.014*	.016*
	(6.51)	(13.10)	(10.95)
Constant	.012*	.014*	.004*
	(14.24)	(29.38)	(5.99)
Pseudo-R^2(in percentage)	2.3	3.8	2.8
Observations	5555	6399	3630
Number of countries	11	11	10

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. The ten countries in the 1980 sample are: Denmark, France, Germany, Ireland, Italy, the Netherlands, the UK, Greece, Portugal and Spain. Austria is added to the sample in 1985.

As we can see in Table 3.11, when domestic production is considered distance coefficients fall for all three cross-sections relative to what is reported in Table 3.7, perhaps reflecting lower within-country transport cots. For example, for 1985 the estimated distance coefficient decreases from 4.7 to 1.4 percent. Moreover, we can see again a tendency for a monotonically declining impact of distance over time as this falls from 1.6 in 1980 down to 1.1 in 1990. The domestic distribution cost as proxied by real GDP per capita, is similar to the specification without domestically consumed production for all three cross-sections. Finally, the price elasticity of population is estimated to be negative and significant for all cross sections.

In Table 3.12, we report estimates obtained by replacing real GDP per capita with the real wage rate. Accounting for the effect of domestically consumed production, price elasticities for distance and the real wage rate are positive and significant in all three cross-sections while the price elasticity for population is always estimated to be negative and statistically significant. Again, we see a decline in the price elasticity with respect to distance from 1.4 percent in 1980 down to 0.9 percent in 1990.

Finally, we include VAT differences as an explanatory variable and report the results in Table 3.13. VAT differences have positive and significant effects for all years. Similarly, the distance and

domestic production is calculated at 3-digit aggregation, the weights are generated at that level when we run the regressions with domestic production. The estimates without domestic production were generated by using weights in 4 digits. For the sake of comparability we also run the regressions without domestic production by using weights in 3 digits and estimates were very close to the ones reported in Tables 3.7 to 3.10.

Table 3.12: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980
	Model 1	Model 1	Model 1
Population	-.027*	-.029*	-.010*
	(-27.53)	(-57.85)	(-13.92)
Real Wage Rate	.065*	.094*	.080*
	(18.38)	(62.63)	(30.01)
Distance	.009*	.007*	.014*
	(5.44)	(7.65)	(11.29)
Constant	.008*	.002*	.002*
	(9.62)	(5.12)	(3.10)
Pseudo-R^2(in percentage)	2.1	4.5	1.7
Observations	5454	5423	3537
Number of countries	11	10	10

Notes: * p-value <0.01, ** p-value<0.05, *** p value<0.10. The ten countries in the 1980 sample are: Denmark, France, Germany, Ireland, Italy, the Netherlands, the UK, Greece, Portugal and Spain. The Netherlands is not included in the 1985 cross-section because it's missing the wages data for many industries. Austria is added to the sample in 1985.

Table 3.13: 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.

Table 3.14: Regressions using comparisons relative to trade-weighted probabilistic exporter.

	1990	1985	1980
	Model 1	Model 1	Model 1
Population	-.023*	-.027*	-.005*
	(-23.12)	(-58.17)	(-5.21)
Real Wage Rate	.062*	.232*	.149*
	(17.36)	(40.94)	(27.51)
Distance	.007*	.008*	.010*
	(4.26)	(8.09)	(6.94)
VAT	.094*	.057*	.229*
	(8.12)	(10.13)	(4.90)
Constant	.009*	.002*	.001***
	(10.62)	(3.59)	(1.88)
Pseudo-R^2(in percentage)	2.6	10.0	3.4
Observations	5454	3975	2743
Number of countries	11	7	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. The Netherlands is not included in the 1985 cross-section because it's missing the wages data for many industries. 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.

GDP per capita coefficient are estimated to be positive and population negative and significant for all cross sections. Estimates for the specification with real wage rates and VAT differences as explanatory variables reported in Table 3.14 are qualitatively similar to those in Table 3.13, 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 as these rates become more homogenized across countries over the period.

Heterogeneity in transport costs across industries

We have shown that once we utilize information regarding the source of products sold in any two locations, transportation costs as 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 positive significant and robust impact on absolute price dispersion in our sample of bilateral country comparisons. Here, we consider a specification with industry-specific distance coefficients that aims to explore the relative importance of transportation costs across different industries. This is again in line with Anderson and van Wincoop (2004) who consider heterogeneity in transport costs in their extension of Eaton and Kortum (2001) who assumed identical trade costs. As was the case with the measures of tradeability and local factor input content used in Model 2 previously, 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 the model with absolute price deviations as in Model 2, rather than the directional regressions from Model 1. Specifically, we consider a slightly modified version of Model 2 adding now industry-specific distance coefficients and excluding the industry-specific tradeability measure from CTZ. We implement this by utilizing information on the source of individual products to consider price comparisons relative to the most productive country in each industry.

Table 3.15: Ranking industries according to relative transportation costs

Industry Description:	(1) Ranking according to Value[a]	(2) Ranking according to Tradeability[b]	(3) Ranking according to 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.

Once we obtain industry-specific distance coefficients, we then rank the industries according to how high the distance coefficient is estimated to be, with the industry with the lowest distance coefficient ranked first and the one with the highest coefficient ranked last. To see how this ranking relates to the other measures of the importance of transportation costs we also report the ranking of the industries according to (1) the average value of goods within that industry classification, and (2) the degree of tradeability characterizing a certain industry. To obtain the "value" of the typical good in each industry used for the ranking in Table 3.15, we average the common currency prices of each good across countries and then aggregate this average price across all goods that fall in the same ISIC classification. Assuming a fixed component to transportation costs, then the per unit transportation cost attributed to this fixed component should decline with the value of the good considered in column (1) of Table 3.15, with expensive goods having lower per unit costs. Tradeability is constructed as described in the data section. As we have argued there, tradeability has a direct interpretation as an inverse measure of effective trade costs.

If the above reasoning is valid, and as long as our industry-specific distance coefficients capture the relative importance of transportation costs across industries, these estimates should be closely related to the measures of value and tradeability considered here. 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. Moreover, the correlation between tradeability ranking in column (2) and the distance coefficient ranking in column (3) is similar in value and again statistically significant beyond the one percent level. As a robustness check we run the regressions using wage rates in place of GDP per capita. In this case, the correlation between the value ranking and the distance coefficients ranking is 35 percent and that between the tradeability ranking and the distance coefficients ranking is 45 percent, both statistically significant at the five percent level.

4. Conclusion

Trade costs are important in a number of international macroeconomic models with implications for price deviations across countries. Transport costs are one 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 international trade. Thus, the relevance of transport costs in determining price wedges and international quantity flows might remain important even as technological progress lowers the level of transport costs 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.

One commonly used measure for transport costs is physical distance from the origin. Here, distance relative to the most productive country has a precisely estimated positive impact on international deviations from LOP that is larger than the estimates obtained when arbitrarily assigning an equal probability of being the source to each country. Our estimates of the impact of transport costs using actual realizations of trade flows across countries 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. This confirms that productivity is a strong predictor of the direction of trade and that the assumption of the most productive country in an industry being the main exporter for all products of that industry is not a bad first approximation.

An interesting feature that emerges is the falling importance of transport costs in absolute terms as witnessed in the declining estimated coefficient for the impact of physical distance on prices during the period from 1975 to 1990. This is consistent with economic intuition as transport technologies have been improving over time. It is also in accord with much of the literature documenting the declining importance of physical distance over time. Distribution costs are also important in determining international deviations from LOP, confirming the well-known fact that countries with higher income per capita -and thus a higher cost for the local inputs component- have higher prices. Moreover, industry-specific measures of local input costs have a positive impact on absolute price dispersion. Overall, the data is consistent with models where transport costs, distribution costs, market size, and taxes play important roles in the determination of international price differences.

We conclude that utilizing relative productivity along with relative prices from survey data can help in identifying 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 .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.app. 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 m.	3900
Musical instruments, parts and accessories	3900
Meahanical lighters and parts	3900

Construction of the capital stock for each industry

The Database on Investment and Capital for Agriculture and Manufacturing reports the total capital stock for the manufacturing sector (TK). In order to calculate capital stock for each manufacturing industry, 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. We calculate total manufacturing sector investment by using capital formation data for twenty-eight manufacturing industries, and then obtain each industry's share of total manufacturing for each country. However, since some countries have missing observations for some industries the shares of the remaining industries are overestimated. In order to resolve this problem, we use the following approach for each cross-section: Let us denote $\mathrm{I}_{\max}$ as total investment in the manufacturing sector for countries that have no missing values. Then, the industries that have missing investment values for at least one country are excluded and the sum of capital formation for the remaining industries is denoted for each country as I_j. We now define $Fraction = \frac{1}{N}\sum_{j=1}^{N}(I_j/I_{\max_j})$ where N is the number of countries that are used to calculate $\mathrm{I}_{\max}$. We assume this fraction is the same for countries that have missing capital formation data for one or more industries. Then for each industry h and country j, we define $weight_{hj} = \frac{I_{hj}}{N_{hj}}$. If a country has missing data, then the share of the capital stock for each industry h is defined as $Share_{hj} = [TK_j] \times [Fraction]$ and its capital stock is now given by $K_{hj} = weight_{hj} \times share_{hj}$. If the country does not have missing data then we assume the share of investment for each industry is simply equal to its share in the capital stock given as $share_{\max_{hj}} = \frac{I_{hj}}{I_{\max_{hj}}}$ and $K_{hj} = [TK_j] \times [\mathrm{share}_{\max_{hj}}]$.

Appendix

Table A3: Arbitrary Reference country comparisons

Germany	1990			1985		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Population	-.068* (-10.72)	-.011** (-2.00)	.009 (0.53)	-.060* (-9.89)	-.001 (-0.19)	.008 (0.55)
GDP-per-capita	.835* (11.68)	.416* (6.13)	.260 (1.55)	.395* (8.05)	.046 (0.87)	-.017 (-0.12)
Distance	.267* (7.40)	-.093* (-3.38)	.040 (0.61)	.042 (1.48)	.093* (4.06)	.144** (2.18)
Tradeability		-.057* (-4.44)	-.043 (-1.36)		.011 (0.85)	-.015 (-0.69)
Local inputs share		.001 (0.67)	.003 (0.68)		.003* (2.60)	.005** (2.26)
Large cars		.169* (4.92)			.266* (4.02)	
Vices		.135* (5.76)	.095** (1.98)		.143* (6.11)	.089*** (1.83)
Constant	.047* (3.09)	.261* (11.73)	.251* (4.30)	.037** (2.33)	.281* (15.87)	.285* (6.32)
R^2(in percentage)	5.7	5.7	23.2	4.5	4.8	17.3
Observations	3567	3567	133	4244	4244	161
UK						
Population	-.047* (-6.81)	.004 (0.52)	.006 (-0.31)	-.064* (-10.18)	-.011*** (-1.79)	-.019 (-1.21)
GDP-per-capita	.379* (9.61)	-.055 (-0.99)	.215 (1.40)	.278* (9.26)	.137* (3.46)	.275** (2.30)
Distance	.034** (2.27)	.109* (9.37)	.107* (2.68)	-.004 (0.31)	.115* (12.51)	.158* (5.75)
Tradeability		-.030*** (-1.81)	-.007 (-0.21)		-.042** (-2.48)	-.054** (-1.96)
Local inputs share		.003 (1.51)	.008*** (1.84)		.005* (3.96)	.006*** (1.93)
Large cars		.035 (1.28)			.097*** (1.84)	
Vices		.128* (4.22)	.147* (2.98)		.131* (4.43)	.097** (2.16)
Constant	.108* (8.47)	.336* (13.17)	.199* (3.24)	.027** (2.32)	.297* (15.20)	.287* (6.58)
R^2(in percentage)	3.4	4.7	21.0	3.2	5.1	29.0
Observations	2834	2834	130	4184	4184	160

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," April 2004, mimeo Gordon College.

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

Balassa, Bela (1964) "The Purchasing Power Parity Doctrine, A Reappraisal," *Journal of Political Economy*, 584-96.

Benigno, Gianluca and Christoff Thoenissen (2003) "Equilibrium Exchange Rates and Supply-Side Performance," *The Economic Journal* 113, C103-C124.

Bergin, R. Paul, Reuven Glick , and Alan M. Taylor (forthcoming) "Productivity, Tradability, and the Long-Run Price Puzzle," Journal of Monetary Economics.

Bergstrand, Jeffrey (1991) "Structural Determinants of Real Exchange Rates and National Price Levels: Some Empirical Evidence," *The American Economic Review* 81(1), 324-34.

Berka, Martin, "Non-linear Adjustment in International Prices and Physical Characteristics of Goods," manuscript, University of British Columbia, November 2002.

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.

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.

Dornbusch, Rudiger, S. Fischer, and Paul A. Samuelson (1977) "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods" *The American Economic Review*, 67(5), 823-39.

Eaton, Jonathan and Samuel Kortum (2001), "Technology, trade, and growth: A unified framework," *European Economic Review Papers and Proceedings*, 45(5) 742-755.

______ (2002) "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741-79.

Ernst and Young (1996) "Vat and Sales Taxes Worldwide: A Guide to Practice and Procedures in 61 Countries," John Wiley & Sons Ltd.

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

Harrigan, James, "Estimation of Cross-Country Differences in Industry Production Functions," NBER Working Paper Series No. 6121, August 1997.

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

Hummels, David, and Volodymyr Lugovskyy, "Usable Data? Matched Partner Trade Statistics as a Measure of International Transportation Costs," manuscript, Purdue University, May 2003.

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

OECD (1988) "Taxing Consumption," Paris, France.

Samuelson, Paul (1964) "Theoretical Notes on Trade Problems," *Review of Economics and Statistics*, 46, 145-54.