ELSEVIER

Journal of International Economics 68 (2006) 238–265

Journal of INTERNATIONAL ECONOMICS

www.elsevier.com/locate/econbase

Product quality and the direction of trade

Juan Carlos Hallak *

Department of Economics, University of Michigan, Ann Arbor, MI, 48109, USA

Received 14 September 2004; received in revised form 7 April 2005; accepted 22 April 2005

Abstract

A substantial amount of theoretical work predicts that quality plays an important role as a determinant of the global patterns of bilateral trade. This paper develops an empirical framework to estimate the empirical relevance of this prediction. In particular, it identifies the effect of quality operating on the demand side through the relationship between per capita income and aggregate demand for quality. The model yields predictions for bilateral flows at the sectoral level and is estimated using cross-sectional data for bilateral trade among 60 countries in 1995. The empirical results confirm the theoretical prediction: rich countries tend to import relatively more from countries that produce high-quality goods.

Keywords: Product quality; Linder; Gravity; Bilateral trade; North–South trade

JEL classification: D12; F1; F12

1. Introduction

Increasing evidence indicates that there are large differences across countries in the quality of the products that they produce and export. While traditional theories of international trade neglect the existence of product quality differences across countries, a substantial amount of theoretical work predicts that quality systematically affects the direction of international trade. In spite of the theoretical predictions, there is yet no

* Tel.: +1 734 763 9619; fax: +1 734 764 2769.
E-mail address: hallak@umich.edu.

0022-1996/$ - see front matter
doi:10.1016/j.jinteco.2005.04.001

evidence evaluating the empirical relevance of quality as a determinant of bilateral trade volumes.

Linder (1961) first noted the role of quality as a determinant of the direction of trade. He argued that richer countries spend a larger proportion of their income on high-quality goods. He also argued that closeness to demand is a source of comparative advantage, providing richer countries with a comparative advantage in the production of high-quality goods—the goods that they demand. He then conjectured that the congruence of production and consumption patterns leads countries with similar income per capita to trade more with one another. This is the Linder hypothesis, the earliest theory explaining the effects of quality differences on the direction of trade.

More recent theoretical work has developed general equilibrium models to formalize the role of quality as a determinant of trade patterns.[1] These models share two key features with Linder's theory. First, rich countries have a comparative advantage in the production of high-quality goods (stemming from productivity or factor endowment differences). Second, rich countries consume high-quality goods in larger proportions than poorer countries. Even though the models yield theoretical results in the spirit of the Linder hypothesis, they do not obtain this conjecture as a general result.

Recent empirical work provides support for the supply-side relationship between per capita income and quality production postulated by Linder and subsequent theorists. Schott (2004) shows that export unit values increase systematically with exporter per capita income and relative endowments of physical and human capital, while Hummels and Klenow (2005) use quantities exported and proxies for the number of varieties to argue that quality differences are necessary to explain (at least part of) the observed differences in unit values. In contrast to the evidence on the supply-side, the demand-side relationship between per capita income and quality consumption and the impact of this relationship on bilateral trade flows have not been the subject of similar empirical scrutiny. In particular, there is yet no attempt at estimating the existence and magnitude of such a quality-driven demand effect on the global patterns of bilateral trade.[2]

A vast related literature has concentrated on testing the Linder hypothesis. Since this is a conjectured corollary to a theory that places quality at center stage, tests of this hypothesis could be interpreted as evidence on the role of quality. This literature typically uses the gravity equation as a benchmark, and adds a "Linder term," a measure of income dissimilarity between pairs of countries. Even though the Linder hypothesis predicts a negative sign for the estimated coefficient on the Linder term, the empirical

[1] See Falvey and Kierzkowski (1987), Flam and Helpman (1987), Stokey (1991), and Murphy and Shleifer (1997).

[2] Brooks (2003) shows that differences in export shares among Colombian firms (in industries with exports largely oriented to the US) depend negatively on the industry-level quality gap relative to G7 countries, as measured by differences in unit value of exports to the US. Verhoogen (2004) shows similar evidence for Mexico using ISO 9000 certification as an indicator of quality upgrading.

results on the sign of this coefficient are mixed.[3] There is nevertheless an even more fundamental problem than failure to confirm the Linder hypothesis: the empirical framework used by this literature cannot properly identify the role of quality as a determinant of the direction of trade. First, the prediction that the intensity of trade is higher between countries with similar income per capita can also result from inter-sectoral non-homotheticities in demand, not related to quality. This is the case if income elasticities differ across sectors, and richer countries have a comparative advantage in sectors with high income elasticities.[4] Second, quality effects coexist with other traditional (inter-sectoral) determinants of trade, such as differences in factor proportions. But the gravity equation framework does not nest these different forces. It is thus unable to isolate the role of quality from other inter-sectoral determinants of trade.

This paper provides a testable framework to estimate the impact of quality on the direction of trade. In particular, it identifies the effect of quality operating on the demand side through the relationship between income and aggregate demand for quality. The theoretical framework, described in Section 2, yields predictions for bilateral trade flows at the sectoral level. By focusing on sectoral trade, the empirical specification embeds and controls for inter-sectoral determinants of comparative advantage. A parameter in the demand system captures the extent to which income per capita affects quality choice. In the empirical specification, this parameter translates into the coefficient on an interaction term between exporter quality and importer per capita income. If income affects quality demand – and therefore trade patterns – this coefficient is predicted to be positive.

Section 3 tests the model using a cross-section of bilateral trade flows (at the 3-digit level) among 60 countries in 1995. Following the classification of goods in Rauch (1999), only “differentiated” sectors are used in the estimation because the assumptions of the theoretical framework correspond more closely to the characteristics of those sectors. A key variable in the analysis is the measure of quality for each country and sector. I construct export price indices based on cross-country differences in unit values of US imports at the 10-digit level. I then interpret these indices as quality indices. The results support the theoretical prediction: rich countries tend to import relatively more from countries that produce higher quality goods.

In addition to quality, cross-country variation in export prices might also reflect differences in prices for goods of the same quality. Section 4 drops the assumption that export prices only reflect quality differences and allows factors other than quality, in particular exporter per capita income, to influence export prices. The resulting empirical specification now includes an additional term interacting exporter and importer per capita income. Since exporter income is correlated with export prices, the inclusion of this term also addresses the concern that the results of the previous section might be driven by other factors correlated with income but unrelated to quality. The results indicate that this is not the case. The interaction term between exporter price and importer income retains substantial explanatory power after the additional term is included.

[3] See surveys in Deardorff (1984), Leamer and Levinsohn (1995), and McPherson et al. (2001).

[4] The role of this type of non-homotheticity has been addressed by Markusen (1986), Hunter and Markusen (1988), Bergstrand (1989), Hunter (1991), Deardorff (1998), and Matsuyama (2000).

Section 5 estimates the model using "reference-priced" and "homogeneous" sectors, the remaining two types of goods in Rauch's classification. For reference-priced sectors, where the theory may still reasonably apply, the results also support the theoretical prediction. Surprisingly, they appear to be even stronger than the results for differentiated sectors, even though this is not true when the estimation is performed using a censoring model, which accounts for zero-valued trade flows. For homogeneous sectors, where the theory should not apply, the results are as expected: there is no quality effect on the pattern of bilateral trade. These results provide complementary evidence that it is quality – as opposed to other factors correlated with income per capita but unrelated to quality – that drives the previous results. Section 6 concludes.

2. Theoretical framework

2.1. The demand system

Demand in each country k is generated by a representative consumer with a two-tier utility function. The upper tier utility is weakly separable into subutility indices defined for each differentiated good sector $z=1, \ldots, Z$, and for each homogeneous good sector $g=Z+1, \ldots, G$,

$$U^k = U\left[u_1^k, \ldots, u_z^k, \ldots, u_Z^k, u_{Z+1}^k, \ldots, u_g^k, \ldots, u_G^k\right]. \tag{1}$$

The subutility index u_g^k is a general function of the quantity consumed of good g. The subutility index u_z^k takes the specific form:

$$u_z^k = \left[\sum_{h \in H_z} \left(\theta_h^{\gamma_z^k} q_h\right)^{\alpha_z}\right]^{\frac{1}{\alpha_z}} \quad 0<\alpha_z,\gamma_z^k<1 \quad \forall z,k \tag{2}$$

where u_z^k is defined over all varieties $h \in H_z$ in sector z. In (2), q_h and θ_h are the quantity and quality of variety h, and the parameter γ_z^k is the intensity of preference for quality of country k. None of the parameters is restricted to be the same across sectors.

The subutility functions u_z^k are an augmented version of the Dixit–Stiglitz structure of preferences. Quality enters as a utility shifter, while there is still a horizontal dimension of product differentiation. This specification of utility is designed to capture differences across countries in quality demand, stemming from their differences in income. For a given shape of the income distribution, we expect countries with higher average income to consume a larger proportion of high-quality goods. In the demand system that this utility generates, the parameter γ_z^k captures – in a reduced form – the effect of income on quality demand at the aggregate level.

The representative consumer uses two-stage budgeting. In the first stage, for a given expenditure allocation across sectors E_1^k, …, E_Z^k, …, E_G^k, expenditure on variety h in sector z is:

$$p_h^k q_h^k = \frac{\left(\frac{p_h^k}{\theta_h^{\gamma_z^k}}\right)^{1-\sigma_z}}{\sum_{r\in H_z}\left(\frac{p_r^k}{\theta_r^{\gamma_z^k}}\right)^{1-\sigma_z}} E_z^k = s^k(h)E_z^k \tag{3}$$

where $\sigma_z = 1/(1-\alpha_z) > 1$ is the elasticity of substitution, and p_h^k is the price of h faced by consumers in country k. Eq. (3) shows expenditure on h as a share $s^k(h)$ of total expenditure in sector z. This share depends on the value of γ_z^k, and it changes with this parameter according to:

$$\frac{\partial s^k(h)}{\partial \gamma_z^k} = \lambda_{hz}^k\left[\ln\theta_h - \sum_{r\in H_z} s(r)\ln\theta_r\right], \quad \lambda_{hz}^k = \frac{\left(\frac{p_h^k}{\theta_h^{\gamma_z^k}}\right)^{1-\sigma_z}(\sigma_z - 1)}{\sum_{r\in H_z}\left(\frac{p_r^k}{\theta_r^{\gamma_z^k}}\right)^{1-\sigma_z}} > 0. \tag{4}$$

Eq. (4) highlights the main characteristic of the demand system. For a variety h of above-average quality – the term in brackets in (4) is positive – a higher γ_z^k induces a larger share spent on h. For a variety h of below-average quality, a higher γ_z^k induces a smaller share spent on h. Countries with higher γ_z^k thus spend a larger share of their income on high-quality goods. Allowing γ_z^k to vary across countries, this demand system has the convenient property of accommodating in a simple form cross-country differences in the pattern of expenditures for goods of different quality.[5] A special case arises when γ_z^k is the same for every country. In that case, the demand system is equivalent to the demand system generated by the Dixit–Stiglitz structure of preferences, where there are no differences across countries in quality choice.[6] Since Dixit–Stiglitz preferences are standard in international trade models and empirical frameworks with product differentiation, the proposed demand system has the additional advantage of embedding a meaningful benchmark against which to assess the impact of quality on aggregate demand and trade.[7]

[5] I do not address here the effects of differences in higher moments of the income distribution. See Francois and Kaplan (1996) and Dalgin et al. (2004) for an empirical treatment of inequality and trade based on inter-sectoral non-homotheticities.

[6] In this case, quantity substitutes for quality at the same rate in every country. Quantities and prices can then be renormalized to “common-quality” units to obtain the typical Dixit–Stiglitz specification.

[7] An earlier version of this paper (Hallak, 2003) shows how income effects on quality choice can also be derived from a nested logit model of demand.

2.2. Bilateral trade flows at the sectoral level

Country i produces N_{iz} different varieties in sector z. These varieties are symmetric; they share the same quality and sell at the same price.[8] We can multiply Eq. (3) by the number of varieties N_{iz} to obtain country k's total imports from country i in sector z:

$$\mathrm{imp}_{iz}^{k} = N_{iz}\frac{\left(\frac{p_{iz}\tau_{iz}^{k}}{\theta_{iz}^{\gamma_z^k}}\right)^{1-\sigma_z}}{\sum_{r\in H_z}\left(\frac{p_r\tau_r^k}{\theta_r^{\gamma_z^k}}\right)^{1-\sigma_z}}E_z^k \tag{5}$$

where we use $p_{iz}^{k}=p_{iz}\tau_{iz}^{k}$, the equality between import price and the product of export price and trade cost factor between i and k.

Countries differ in the quality of the goods they produce and in their pattern of sectoral specialization. While recent empirical work (Schott, 2004) studies supply-side determinants of quality production, this paper instead takes the distribution of quality production across countries as given. Conditional on this distribution, it identifies the effect of quality on the direction of trade operating on the demand side through the relationship between income and quality choice.

The proposed demand system allows quality choice to depend on a country-specific parameter. In the special case with equal γ_z^k across countries, the demand system is equivalent to the standard Dixit–Stiglitz structure of preferences, which does not allow the demand for quality to vary across countries. I show next that Dixit–Stiglitz preferences impose a strong restriction on the relationship between bilateral flows at the sectoral level. This restriction is independent of the distribution of quality production across countries. Thus, it provides a key for identifying quality effects on the direction of trade.

Using (5), consider US (country k) imports of Rubber Tires (sector z) from Germany (country i). Alternatively, consider US imports of tires from Turkey (country j) by replacing i with j in (5). The ratio between these two expressions indicates the ratio of US imports of German tires relative to Turkish tires, and is independent of the expenditure level and the price index for tires in the US:

$$\mathrm{ratio}_{ij}^{k}(z) = \frac{N_{iz}}{N_{jz}}\left(\frac{p_{iz}\tau_{iz}^{k}/\theta_{iz}^{\gamma_z^k}}{p_{jz}\tau_{jz}^{k}/\theta_{jz}^{\gamma_z^k}}\right)^{1-\sigma_z}. \tag{6}$$

[8] This will not be a strong restriction so long as the variation of quality levels within countries is small compared to the variation of quality levels between countries.

Replacing k with l in (6), consider the same ratio for a different importer, Argentina (country l). Then, compare these by taking the ratio of the two ratios:

$$r_{ij}^{kl}(z) = \frac{\text{ratio}_{ij}^{k}(z)}{\text{ratio}_{ij}^{l}(z)} = \left(\frac{\tau_{iz}^{k}/\tau_{jz}^{k}}{\tau_{iz}^{l}/\tau_{jz}^{l}}\right)^{1-\sigma_z} \left[\frac{\left(\theta_{iz}/\theta_{jz}\right)^{\gamma_z^k}}{\left(\theta_{iz}/\theta_{jz}\right)^{\gamma_z^l}}\right]^{\sigma_z-1}. \tag{7}$$

To abstract from the impact of trade costs, assume for now that $\frac{\tau_{iz}^{k}/\tau_{jz}^{k}}{\tau_{iz}^{l}/\tau_{jz}^{l}} = 1$. Then

$$r_{ij}^{kl}(z) = \left[\frac{\left(\theta_{iz}/\theta_{jz}\right)^{\gamma_z^k}}{\left(\theta_{iz}/\theta_{jz}\right)^{\gamma_z^l}}\right]^{\sigma_z-1}. \tag{8}$$

There are three possible cases in (8). The first one is trivial. If Germany and Turkey produce the same quality of tires ($\theta_i=\theta_j$), then relative imports from Germany and Turkey are the same for both the US and Argentina ($r_{ij}^{kl}=1$). A much more relevant case is the second one, which is our benchmark case of no income effects on quality choice. If the US and Argentina have the same intensity of preference for quality ($\gamma^k=\gamma^l$), then relative imports from Germany and Turkey will still be the same for both importers ($r_{ij}^{kl}=1$), even if qualities are different ($\theta_i\neq\theta_j$). As argued before, standard Dixit–Stiglitz preferences can accommodate this case after an appropriate normalization to "common-quality" units. Thus interpreted, these preferences impose the restriction that any two countries' relative imports from any other two countries are identical, i.e., $r_{ij}^{kl}=1$, regardless of the quality produced by the two exporting countries.[9] The restriction does not neglect inter-sectoral determinants of trade. Suppose that Germany has a comparative advantage in producing tires. Then, Germany will have many firms in this sector. N_{iz} will be large, and Germany will be a large exporter of tires.[10] But exports from Germany will be large to both the US and Argentina without affecting r_{ij}^{kl}. A similar exercise focusing on an importer provides the same answer. Suppose that Argentina has a comparative advantage in producing tires. We then expect a large number of tire-producing firms there, leading to a low price index for tires since domestic goods do not pay trade costs. Relative prices of imported varieties will then be high, discouraging Argentina's imports from both Germany and Turkey. But again, r_{ij}^{kl} will not be affected. Finally, suppose that E_z^k is large for the US because of a combination of size and inter-sectoral non-homotheticities in demand. This will affect US imports from both countries proportionally, but it will still not affect the relative ratio r_{ij}^{kl}.

It is only in the third case, where both quality and the intensity of preference for quality are different ($\theta_i\neq\theta_j$ and $\gamma^k\neq\gamma^l$), that $r_{ij}^{kl}\neq 1$. Only then will quality affect the relative intensity of sectoral trade between different country pairs. If Germany's quality is higher ($\theta_i>\theta_j$) and US's intensity of preference for quality is higher ($\gamma^k>\gamma^l$), then the US will import relatively more from Germany while Argentina will import relatively more from Turkey. More generally, countries with higher γ^k will import relatively more from countries that produce higher quality goods.

[9] In the more general case, this is only true after controlling for differences in bilateral trade costs.

[10] Romalis (2004) derives the effect of comparative advantage on the number of firms in a Heckscher–Ohlin model that accounts for product differentiation and trade costs.

Taking logarithms on both sides of Eq. (5), we obtain the following prediction for bilateral trade:

$$\ln \mathrm{imp}_{iz}^{k} = \ln N_{iz} - \tilde{\sigma}_{z} \ln p_{iz} - \ln \sum_{r \in H_{z}} \left(\frac{p_{r}^{k}}{\theta_{r}^{\gamma_{z}^{k}}} \right)^{-\tilde{\sigma}_{z}} + \ln E_{z}^{k} - \tilde{\sigma}_{z} \ln \tau_{iz}^{k} + \tilde{\sigma}_{z} \gamma_{z}^{k} \ln \theta_{iz} \tag{9}$$

where $\tilde{\sigma}_z = \sigma_z - 1$. In a cross-section of bilateral trade flows at the sectoral level, the first two terms on the RHS are specific to exporting country i. These terms take the same value when i is the exporter, independent of who importer k is. In the econometric specification, they will be captured by sector-specific *exporter* fixed effects. Similarly, the next two terms are specific to importer k and take the same value independent of exporter i. These terms will be captured by sector-specific *importer* fixed effects. Only the last two terms are specific to the bilateral pair. I assume that trade costs are determined by:

$$\ln \tau_{iz}^{k} = \eta_{z} \ln \mathrm{Dist}_{i}^{k} + \tilde{\beta}_{z} \mathbf{I}_{i}^{k} - v_{iz}^{k} \tag{10}$$

where Dist_i^k is the bilateral distance between each country pair, v_{iz}^k is a random disturbance, and $\tilde{\beta}_z$ is a vector of parameters associated with dummy variables $\mathbf{I}_i^k$ indicating whether the country pair shares, respectively, a common border, a common language, a preferential trade agreement, a colonial relationship, or a common colonizer.[11] I also postulate a relationship between the intensity of demand for quality – captured by γ_z^k – and income:[12]

$$\gamma_{z}^{k} = \gamma_{z} + \mu_{z} \ln y^{k}. \tag{11}$$

Under the null hypothesis that income does not affect demand for quality, $\mu_z = 0$, and $\gamma_z^k = \gamma_z$. This is the benchmark Dixit–Stiglitz case. Under the alternative hypothesis that it does affect demand for quality, $\mu_z > 0$, and γ_z^k increases with income.

Combining (10) and (11) with (9), we obtain:

$$\ln \mathrm{imp}_{iz}^{k} = \varphi_{iz} + \psi_{z}^{k} - \tilde{\sigma}_{z} \eta_{z} \ln \mathrm{Dist}_{i}^{k} + \beta_{z} \mathbf{I}_{i}^{k} + \tilde{\sigma}_{z} \mu_{z} \ln \theta_{iz} \ln y^{k} + \varepsilon_{iz}^{k} \tag{12}$$

where φ_{iz} and ψ_z^k are fixed effects for exporter and importer country, respectively, $\beta_z = -\tilde{\sigma}_z \tilde{\beta}_z$, and $\varepsilon_{iz}^k = \tilde{\sigma}_z v_{iz}^k$. The next section estimates (12) using price (export unit value) indices as measures of quality.

3. Export price (unit value) indices as quality indices

There is wide variation in export unit values across countries, even when measured at very disaggregated product categories. Quality differentiation is often considered the main determinant of this variation. Following this view, in this section, I construct price indices at the sectoral level (p_{iz}) from export unit values, and interpret them as quality indices. Thus, all cross-country variation in unit values is attributed to differences in quality. First, I describe the data set used to estimate (12) and the construction of the price indices. Then, I address estimation issues and present the results. The relationship between price indices and quality is relaxed in the next section to allow for other determinants of export price.

[11] The empirical specification would not change if trade costs depended on quality, so long as this is modeled using a function of quality as an additive term in (10). This term would be subsumed into the exporter fixed effect.

[12] Throughout the paper – unless explicitly noted – income will refer to income per capita.

Table 1
Geometric mean of sectoral price indices

Country	GDP per capita (PPP)	Differentiated goods	Reference priced goods	Homogeneous goods
Switzerland	25,475	1.64	1.79	1.00
Norway	24,693	1.46	1.47	1.23
Japan	23,211	1.48	1.92	1.31
Canada	23,085	1.00	1.00	1.00
Denmark	22,947	1.51	1.48	1.58
Belgium–Lux.	22,700	1.45	1.41	0.97
Singapore	22,270	1.09	1.26	1.17
Hong Kong	22,166	0.79	1.28	1.06
Austria	22,089	1.43	1.71	1.16
Germany	21,478	1.46	1.49	1.14
Australia	21,267	1.29	1.36	1.00
Netherlands	20,812	1.41	1.39	1.20
Italy	20,512	1.33	1.41	1.25
France	20,492	1.54	1.64	1.42
Sweden	20,030	1.46	1.56	1.49
United Kingdom	19,465	1.39	1.47	1.25
Finland	18,764	1.60	1.39	1.24
New Zealand	17,705	1.30	1.39	0.95
Israel	17,394	1.33	1.57	1.27
Ireland	17,264	1.46	1.76	1.13
Spain	15,163	1.29	1.21	0.93
Portugal	13,613	1.20	1.24	1.12
South Korea	13,502	0.91	1.25	1.28
Taiwan	13,335	0.81	1.38	1.20
Greece	13,147	1.10	1.20	1.13
Saudi Arabia	10,766	1.12	1.27	1.13
Argentina	10,736	1.02	1.10	1.00
South Africa	8581	1.11	1.19	1.02
Malaysia	8145	0.94	1.14	1.04
Uruguay	7831	1.04	0.79	1.02
Chile	7544	1.00	0.98	1.19
Mexico	7061	0.74	0.95	1.02
Poland	6605	0.86	1.11	1.09
Brazil	6572	0.89	1.07	1.07
Romania	6430	0.81	0.72	1.02
Thailand	6217	0.79	1.15	0.96
Colombia	6151	0.87	1.04	1.11
Venezuela	5979	0.73	0.89	0.82
Costa Rica	5940	0.97	1.18	0.92
Turkey	5803	0.93	1.33	0.98
Bulgaria	5608	0.91	0.86	0.79
Tunisia	4870	1.02	1.36	1.36
Paraguay	4598	0.83	0.48	0.83
Peru	4329	1.02	1.06	0.98
Dominican Rep.	3997	0.87	0.93	0.89
Lebanon	3964	0.94	1.13	1.13
Philippines	3518	0.89	1.06	0.92
Guatemala	3444	0.87	0.98	0.90
Syria	3211	1.06	1.40	0.93

Table 1 (*continued*)

Country	GDP per capita (PPP)	Differentiated goods	Reference priced goods	Homogeneous goods
Ecuador	3162	0.89	1.02	0.89
Morocco	3052	0.94	1.34	2.37[a]
Indonesia	2869	0.80	1.01	0.88
Egypt	2869	1.05	0.88	0.68
Sri Lanka	2741	0.70	1.06	1.11
China	2560	0.63	1.02	1.06
India	1877	0.82	0.97	0.87
Pakistan	1733	0.82	0.70	0.81
Vietnam	1478	0.70	1.02	0.65
Bangladesh	1253	0.76	1.34	0.58
Nigeria	832	0.81	0.97	0.73
Average correlation b/w sectoral index and GDP Per Capita		0.45	0.36	0.23

The bottom row gives the average across sectors of the correlation between PPP GDP per capita and the price index.

[a] Morocco has price indices for only two sectors.

3.1. Data and sample selection

The data consist of a cross-section of bilateral trade flows and country-level variables for 60 countries in 1995. Bilateral trade data, disaggregated at the sectoral level, come from Feenstra (2000). The data set is based on the World Trade Analyzer assembled by Statistics Canada. I define sectors at the 3-digit SITC (Rev.2) level.

I follow Rauch's (1999) classification of 4-digit SITC sectors into three categories. "Homogeneous sectors" include goods that are internationally traded in organized exchanges, with a well-defined price (e.g., wheat). "Reference-priced sectors" include goods that are not traded in organized exchanges but have reference prices available in specialized publications (e.g., polyethylene). "Differentiated sectors" are those sectors that do not satisfy either of the two previous criteria. Rauch uses two standards to make his classification, one "liberal" and one "conservative." I use the liberal standard because it is more stringent in the classification of goods as differentiated. When a 3-digit sector includes 4-digit subsectors that belong to different classifications, I break the 3-digit sector down accordingly, each part including only the relevant 4-digit sectors.

There is a large proportion of bilateral country pairs with zero trade. The proportion is larger for smaller countries. Since I want to prevent zero-trade observations from dominating the sample, I concentrate only on relatively large countries. I include countries with a population larger than 3 million, and with more than US$ 2 billion imports of differentiated goods. Hungary is additionally excluded because sectoral trade data are of poor quality. Algeria, Iran, Libya, and the US are excluded because they lack data on export unit values. In the case of the US, export unit values are not available because they are obtained from a database on US imports. The final sample consists of 60 countries, listed in the first column of Table 1. I also drop very small sectors, keeping only sectors with a volume of trade (among the 60 selected countries) above US$ 2 billion. The final sample includes 114 differentiated sectors, 51 reference-priced sectors, and 38 homogeneous sectors. Their codes are listed in Table A1.

The distance variable measures great circle distance between capital cities and was prepared by Howard Shatz (1997). Dummies for border, common language, colonizer–colony relationship, and common–colonizer relationship were constructed using the CIA Factbook. Only "official" languages are considered in the construction of the common language variable. An exception is made for Malaysia–Singapore, which is recorded as having a common language. Colonial links are only considered if the colonizer–colony relationship was still in force after 1922. The indicator variable for Preferential Trade Agreement includes PTAs in force and with substantial coverage in 1995: Andean Pact, ASEAN, CACM, EFTA, EEA, EU, MERCOSUR, NAFTA, Australia–New Zealand, EC-Turkey, EFTA–Turkey, EC–Israel, EFTA–Israel. PPP GDP comes from the World Bank WDI.

3.2. Export price indices

I construct export price indices p_{iz}, for country i and 3-digit sector z, based on cross-country differences in export unit values at the 10-digit level. For a given export category, unit values measure the ratio between the value and the quantity of exports. They are the average price for the category. Composition problems are pervasive in unit value comparisons. If a category includes different goods, then differences in unit values might not merely reflect differences in prices but also differences in the composition of exports within the category. To minimize the incidence of composition problems, I measure unit values at the finest possible level of aggregation for which data are available. The NBER Trade Database compiled by Feenstra et al. (2002) classifies US imports by country of origin and type of good at the 10-digit level of the Harmonized Tariff Schedule (HTS), the level at which import duties are defined. For each of these categories and source countries, the database provides information on the customs value and quantity of US imports, and the units in which quantities are measured.[13] Examples of 10-digit categories are:

HTS Code	Description
1006204060	Rice, short grain, husked (brown)
1902112000	Uncooked pasta, not stuffed or otherwise prepared, containing eggs, exclusively pasta
5208292020	Woven fabrics of cotton, containing 85% or more by weight of cotton, weighing not more than 200 g/m^2, bleached, sateens
6203492010	Men's trousers and breeches of artificial fibers
8413702022	Centrifugal pumps for liquids, single-stage, single-suction, frame-mounted, with discharge outlet under 7.6 cm in diameter
8418210010	Refrigerators of household type, compression type, having a refrigerated volume of under 184 l

Based on cross-country differences in unit values at the 10-digit level, I construct export price indices p_{iz} at the 3-digit level using a slightly modified version of the Eltetö,

[13] Customs values do not include freight and are used as the basis for duty assessment. They are intended to serve as arm's length transaction values for commodities.

Köves, and Szulc (EKS) multilateral price index, which in turn is based on bilateral Fisher indices. Appendix A describes the methodology in detail.[14]

Several shortcomings are associated with the construction of these indices. Countries often report exports to the US in only a few or even none of the 10-digit categories included in a particular 3-digit sector. When countries are active in only a few 10-digit categories, the price index is very sensitive to measurement error. When countries are not active in any 10-digit category, the price index has a missing value. To increase the availability and reliability of the indices, I merge 2 years of data, 1995 and 1996, instead of using only 1 year. I also calculate the indices at the 2-digit level and then use the 2-digit indices for the relevant 3-digit categories. This procedure has two advantages. First, the indices are more reliable, as they are based on a larger number of observations. Second, bilateral trade observations for which the exporter's 3-digit price index would be missing are kept in the sample if the corresponding 2-digit index can be calculated.

The indices are based on a source database with considerable measurement error.[15] To deal with this problem, I remove observations with extreme unit values (four times above or below the category mean) and observations with very low quantity (below the lower of 50 units or a quarter of the category mean quantity). Lastly, aggregation problems may still be present at the 10-digit level, but I expect these problems to be minimized at such level of aggregation.

Table 1 provides summary measures of the export price indices.[16] Ordering countries by income, and normalizing the indices so that Canada has a value of 1 in every sector, the table shows the geometric average of the sectoral indices for each goods category. Plate 1a to c provide the same information. The correlation between export price indices and income is positive for all differentiated sectors, and for most reference-priced and homogeneous sectors. The average correlation across sectors is 0.45 for differentiated sectors, 0.36 for reference-priced sectors, and 0.23 for homogeneous sectors. This is consistent with the supply-side assumptions of most theoretical trade models accounting for vertical differentiation, and it confirms the findings of Hummels and Klenow (2005) and Schott (2004).

3.3. OLS estimation

Using the constructed price indices p_{iz} as measures of quality, Eq. (12) can be estimated using a cross-section of bilateral trade flows at the sectoral level. Unfortunately, it is not possible to separately identify the magnitude of μ_z. But since $\sigma_z > 0$, a test of the hypothesis $\sigma_z \mu_z = 0$ implies a test of the hypothesis $\mu_z = 0$.

Eq. (12) is a prediction for bilateral trade at the sectoral level. Estimating it with aggregate data would only be appropriate if the parameters were constrained to be equal across sectors. This restriction is not plausible in general. In particular, it will be strongly violated in the case of the exporter and importer fixed effects, which must be sector-specific as they control for inter-sectoral determinants of comparative advantage. I thus estimate (12) sector by sector. Also, to obtain a single estimate of the parameter of interest, I estimate (12)

[14] See also Diewert (1993), Ch. 5, for properties of this and similar indices.

[15] See General Accounting Office (1995).

[16] The MATLAB code used to construct the indices and the detailed tables with the export price indices by sector are available online: http://www-personal.umich.edu/~hallak/papers.

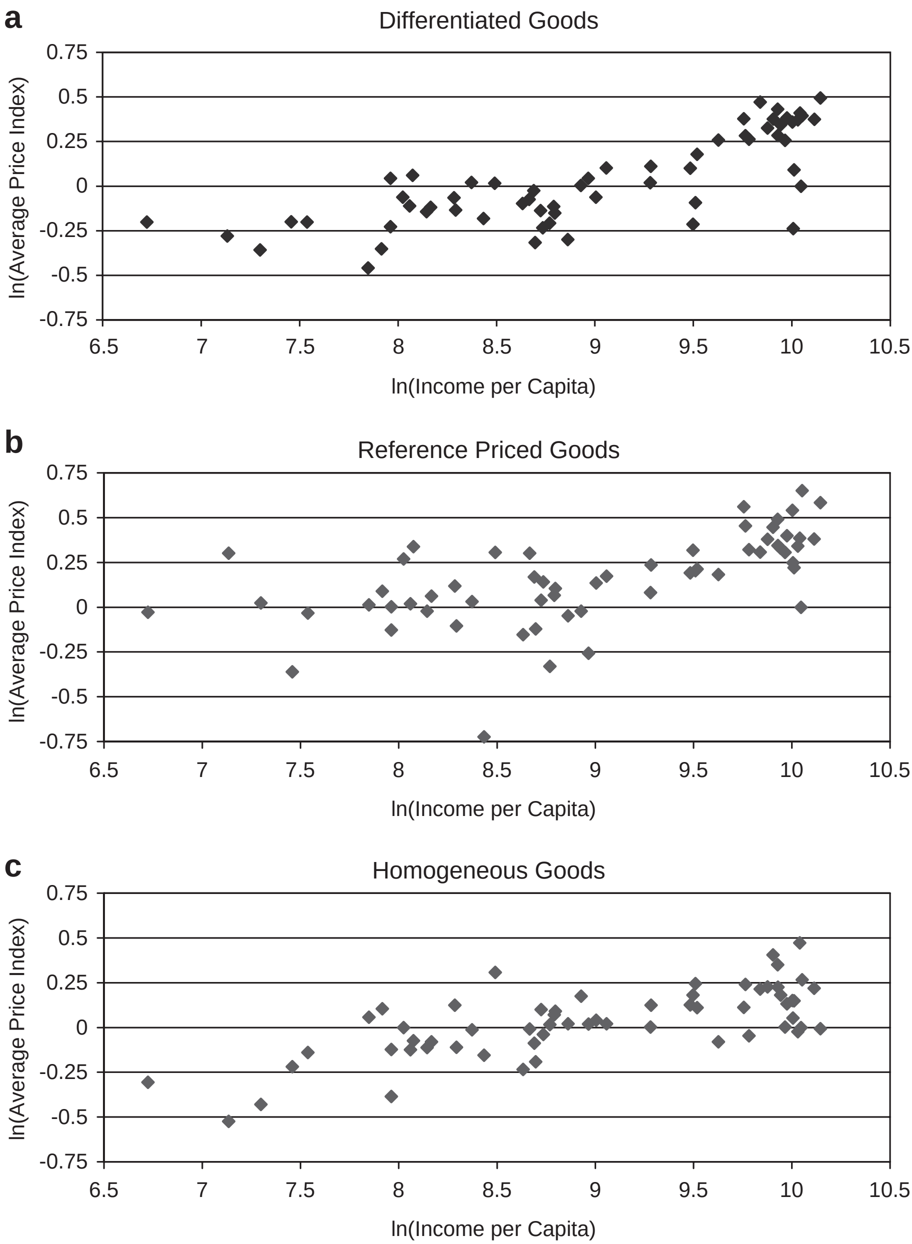

Plate 1. (a–c) Relationship between average price indices and income per capita.

pooling the observations across sectors. In this case, I allow all the parameters (including the exporter and importer dummies) to take sector-specific values, except for the cross-sector restriction on the interaction term: $\tilde{\sigma}_z \mu_z = \tilde{\sigma} \mu$.

Several sectors contain intermediate goods. If we interpret (1) and (2) as a production function of a final good based on the use of intermediate inputs, then (3) is the demand function for these inputs. Therefore, we can still derive (12) as long as transport costs are determined by (10) and the intensity of preference for quality is determined by (11). The last equation might not be appealing in the case of intermediate goods, for which it may be more reasonable to assume that γ_z^k is a function of the quality of the final good, θ_z^k, rather than of importer income. However, if quality supply and income are related by $E(\theta_z^k|y^k, \mathbf{x}) = \delta_{0z} + \delta_{1z}\ln y^k$, where $\mathbf{x}$ is the set of independent variables in (12), we can still derive this equation. The error term will contain an additional component, but it will be uncorrelated with the regressors.

Since (12) was derived from a theory based on product differentiation, I focus on "differentiated" sectors. In Section 5, I discuss the applicability of the theory for other sectors and estimate (12) using those sectors.

I first perform OLS estimation of (12) separately for each differentiated sector. Since I cannot report detailed regression results for all 114 sectors, I provide summary results for the distribution of the estimated coefficients according to sign and significance levels in Table 2 (top panel). In all cases, heteroskedasticity-robust standard errors are used. The first row of the table shows results for the coefficient of interest, the interaction term between exporter quality and importer income. Consistent with the theoretical prediction, this coefficient tends to be positive. It is greater than zero in more than 2/3 of the sectors (79), and lower than zero in less than 1/3 of the sectors (35). The coefficient is positive and significant in 44 sectors, and negative and significant in 15 sectors.

The median coefficient across sectors is 0.1839. The interpretation of this coefficient is related to the ratio of ratios in Eq. (8). Keeping trade costs constant, and denoting by $\hat{\mu}_z$ the estimate of $\tilde{\sigma}_z\mu_z$, we can use (12) to calculate the predicted ratio of ratios:

$$\ln\left(r_{ijz}^{kl}\right) = \hat{\mu}_z \ln\frac{\theta_i}{\theta_j}\ln\frac{y^k}{y^l}. \tag{13}$$

To understand this expression, take countries at the 75th percentile (Sweden and the UK) and the 25th percentile (Dominican Republic and Lebanon) of the income distribution in our sample. Consider Dominican Republic's ratio of imports from the UK relative to those from Lebanon. How much would this ratio increase if the Dominican Republic were to have the income of Sweden? We can find the answer by substituting in (13) for the median quality ratio between the UK and Lebanon across sectors and the income ratio between Sweden and the Dominican Republic: $r_{ijz}^{kl} = \exp(0.1839 * 0.5207 * 1.6115) = 1.154$. Dominican Republic's imports from the UK relative to Lebanon would increase by 15.4% (in the median sector). This is the relevant exercise for interpreting the estimated magnitude of the quality effect. Other typical measures of explanatory power such as the beta coefficient are very small because size (captured by the dummy variables) and distance dominate most of the variation in bilateral trade flows.

The last two columns of Table 2 show results for the pooled regression, where only the coefficient on the interaction term is constrained to be equal across sectors. The first of these columns shows the unweighted OLS results. The second column shows estimation results with observations weighted according to the precision of the price index. Denoting by G_{iz} the number of active 10-digit categories used in the construction of p_{iz}, I assume that the precision

Table 2
Price indices as quality indices differentiated goods (114 Sectors)—OLS and ML estimates

	Regressions by sector						Pooled regression	
	Sign		Significance (5%)			Median	Coefficient (unweighted)	Coefficient (weighted)
	Pos.	Neg.	Pos.	Not Sig.	Neg.			
A. OLS estimation								
$\ln(P_i)*\ln(y_k)$	79	35	44	55	15	0.1839	0.1120*** (0.0359)	0.1153*** (0.0396)
$\ln(\text{Distance}_{ik})$	0	114	0	0	114	−1.0365		
Border_{ik}	102	12	38	76	0	0.3328		
Common Lang_{ik}	112	2	95	19	0	0.5321		
PTA_{ik}	96	18	63	48	3	0.3846		
$\text{Colonial Link}_{ik}$	111	3	86	28	0	0.7912		
$\text{Common Colony}_{ik}$	84	30	17	95	2	0.2893		
B. ML estimation								
Imports equation								
$\ln(P_i)*\ln(y_k)$	73	41	45	56	13	0.1128		
$\ln(\text{Distance}_{ik})$	0	114	0	0	114	−1.1723		
Border_{ik}	102	12	30	84	0	0.3072		
Common Lang_{ik}	113	1	100	14	0	0.5718		
PTA_{ik}	83	31	44	63	7	0.2361		
$\text{Colonial Link}_{ik}$	111	3	95	19	0	0.9150		
$\text{Common Colony}_{ik}$	88	26	25	87	2	0.3460		
Fixed cost equation								
$\ln(\text{Distance}_{ik})$	90	24	38	73	3	0.1310		
Border_{ik}	106	8	47	65	2	0.8063		
Common Lang_{ik}	20	94	2	82	30	−0.2498		
PTA_{ik}	19	95	2	75	37	−0.4094		
Colonial Linkik	28	81	5	68	36	−0.6943		
$\text{Common Colony}_{ik}$	34	79	2	96	15	−0.2024		
Exporter GDP	25	89	5	63	46	−0.1064		
Importer GDP	79	35	32	75	7	0.0488		

Estimation of Eq. (12) in panel (A). Estimation of censoring model (Eqs. (12) and (14)) in panel (B). Columns 2 and 3 provide a breakdown of the total number of sectors by sign of the estimated coefficient. Columns 4 to 6 provide a breakdown by sign and significance. Heteroskedasticity-robust standard errors in all regressions. Clustering by country pair across sectors in pooled regression. Weighted pooled regression uses weights $w=\sqrt{\ln(G_{iz})}$, where G_{iz} is the number of "active" categories of country i in sector z.
***, **, *Significant at the 1%, 5%, and 10% level, respectively.

of p_{iz} is positively related to G_{iz}, and use weights $w=\sqrt{\ln(G_{iz})}$. In both cases, as in every other pooled regression, results are reported with heteroskedasticity-robust standard errors and clustering by country pair. The unweighted and weighted regressions show similar results. The pooled coefficient is substantially smaller than the median of the sectoral coefficients, but we can reject the null hypothesis that $\mu=0$ at the 1% level of significance. These results strongly confirm the theoretical prediction: rich countries tend to import relatively more from countries that produce high-quality goods.

The estimated coefficients for the rest of the variables have the expected signs in most sectors. Distance hinders trade, whereas Border, Common Language, PTA, Colonial relationships, and Common Colony relationships facilitate trade, presumably through a reduction in trade costs. A similar set of results for these variables is available from the pooled regression. Since the coefficients are allowed to take sector-specific values, there is not a unique value to report, as is the case with the (constrained) coefficient on the interaction term. To save space, I do not report these results, which are very similar to those from the sectoral regressions.

3.4. Fixed costs of exporting

Almost half of all sectoral bilateral pairs in the sample report no trade. Since the estimating equation uses the logarithm of bilateral trade, these observations must be discarded when using OLS, possibly inducing selection bias. To address the potential bias, this section develops a simple censoring model based on the assumption that fixed exporting costs explain the substantial fraction of zero values in bilateral trade. International trade only occurs when the profits it generates cover the fixed costs.

The use of a standard censoring model (Tobit) for estimation is not convincing for two reasons. First, we do not know the censoring value. Second, fixed exporting costs are likely to vary across bilateral pairs. I model the (unobserved) censoring value for country-pair ik in sector z as:

$$\log c_{iz}^{k} = \delta_{0z} + \delta_{dz}\log\text{Dist}_{i}^{k} + \delta_{z}\mathbf{I}_{i}^{k} + \delta_{xz}\log GDP_{i} + \delta_{mz}\log\text{GDP}^{k} + u_{iz}^{k} \tag{14}$$

where $\mathbf{I}_{i}^{k}$ is the same vector of dummy variables as in (12), GDP_i and GDP_k are total income of i and k, respectively, and u_{iz}^{k} is a normally distributed random disturbance.

Given the demand structure in (3), mark-ups are constant, and profits are a constant fraction of the value of exports: $\pi_{iz}^{k}=(1/\sigma_z)\ \text{imp}_{iz}^{k}$.[17] Countries trade if the profits that export sales generate are sufficient to cover the fixed costs. Hence, exports occur if $\pi_{iz}^{k}=(1)/(\sigma_z)\ \text{imp}_{iz}^{k}>F_{iz}^{k}$, which implies $\text{imp}_{iz}^{k}>\sigma_z F_{iz}^{k}=c_{iz}^{k}$. Therefore, up to a constant shift, (14) is in fact the equation determining bilateral fixed costs. The empirical specification is then a censoring model with two equations, the "imports equation" (12) and the "fixed cost equation" (14). In the first equation, the dependent variable is now a latent variable. The second equation determines the (unobserved) censoring point c_{iz}^{k}. Thus, (12) only takes non-zero values if $\text{imp}_{iz}^{*k}>c_{iz}^{k}$. Even though the censoring point is unobservable, the parameters of both equations can be estimated jointly by maximum likelihood. This specification is a more general version of the standard Tobit estimation, with unknown and random censoring value.[18] A shortcoming of this approach is the implicit assumption that the decision to export is made at the country-sector level, when fixed exporting costs are in fact borne at the firm level.[19]

[17] With a finite number of varieties, this is only true as a limiting property.

[18] A similar approach was taken by Cogan (1981) to model labor supply with fixed costs of entry into the labor market.

[19] In a recent paper, Helpman et al. (2004) develop a model of bilateral trade with heterogeneous firms, from which they derive two corrections to the OLS specification. The first correction takes account of observations with zero trade, but it does not affect the results significantly. The second correction accounts for firms that do not export, even when non-zero values are observed at the country level. Absent the last correction, the estimated coefficients on variables capturing trade costs confound the effect of these variables on fixed and variable exporting costs. This might not be a major problem here since our variable of interest is not a trade cost determinant.

The estimation is performed sector by sector, assuming a bivariate normal distribution for the random disturbances. The results are shown in panel (B) of Table 2. Despite some differences, the censoring model broadly confirms the OLS results. Even though the number of positive estimated coefficients is lower here, the coefficient on the interaction term is still positive in a majority of sectors (73 out of 114). More importantly, the results by significance level are almost identical; the estimated coefficient is positive and significant in 45 sectors (2/5), and it is negative and significant in 13 sectors (1/9). The median magnitude of the coefficient is considerably lower in the censoring model, and it is now closer to the coefficient obtained from the pooled specification under OLS. This indicates that discarding zero-valued observations does not induce important differences in sign and significance of the estimated coefficients, but it might overestimate the coefficient magnitude.[20]

The results for the rest of the variables in the *imports equation* are also very similar to those estimated under OLS. In the case of the *fixed cost equation*, most of the estimated coefficients tend to have the predicted sign, even though they are not significantly different from zero in most sectors.[21] Somewhat puzzling are the cases of Border and Importer GDP, which are estimated to have a positive impact on fixed exporting costs.[22]

3.5. Exporter income per capita

Both the OLS and MLE results are consistent with the theoretical prediction. However, export prices are strongly correlated with exporter (per capita) income, and thus with many factors other than quality supply that are also correlated with income. Export prices might then be spuriously capturing the effect of these other factors on bilateral trade. To address this concern, I estimate (12) substituting exporter income for exporter price in the interaction term. The results are displayed in Table 3, specification 2 (specification 1 reproduces the results of Table 2 to facilitate comparison). Using exporter income instead of exporter price only marginally changes the number of sectors with positive and negative coefficients. However, there is a substantial increase in the number of sectors with estimated coefficients significantly different from zero. The higher precision of estimates using exporter income suggests that the explanatory power of the interaction term using export prices might merely be the result of export prices acting as a noisy measure of exporter income. On the other hand, the reverse could also be true: exporter income might owe its explanatory power to its high correlation with exporter quality. To disentangle

[20] Computational constraints for performing a maximization routine with the very large number of dummy variables used in the pooled estimation prevent estimating the censoring model on the pooled data.

[21] It is not possible to identify the effect of Colonial Link and Common Colony on fixed costs in 5 and 1 sectors, respectively. When all bilateral observations that share a colonial relationship have non-zero trade, the MLE estimate is $-\infty$. This brings the predicted fixed cost down to zero for such observations, thus maximizing the likelihood of observing positive trade in all of them.

[22] Helpman et al. (2004) similarly find that sharing a Border decreases the probability of trading. They attribute this finding to the effect of territorial border conflicts among neighboring countries.

Table 3
Price indices as indicators of quality differentiated goods (114 sectors)—OLS and 2SLS estimates

		Regressions by sector						Pooled regression	
		Sign		Significance (5%)			Median	Coefficient (unweighted)	Coefficient (weighted)
		Pos.	Neg.	Pos.	Not Sig.	Neg.			
1. Price only									
OLS:	$\ln(P_i)*\ln(y_k)$	79	35	44	55	15	0.1839	0.1120*** (0.0359)	0.1153*** (0.0396)
2. Income only									
OLS:	$\ln(y_i)*\ln(y_k)$	78	36	56	35	23	0.1218	0.0782*** (0.0230)	–
3. Price and income									
OLS:	$\ln(P_i)*\ln(y_k)$	71	43	34	65	15	0.1591	0.0780*** (0.0386)	0.1142** (0.0446)
	$\ln(y_i)*\ln(y_k)$	69	45	51	36	27	0.0918		
2SLS:	$\ln(P_i)*\ln(y_k)$	75	39	33	66	15	0.1858	0.0648 (0.0541)	0.0981* (0.0579)
	$\ln(y_i)*\ln(y_k)$	66	48	35	51	28	0.1016		
4. Price, income, and distance to the US									
OLS:	$\ln(P_i)*\ln(y_k)$	69	45	34	68	12	0.1582	0.0958** (0.0393)	0.1405*** (0.0455)
	$\ln(y_i)*\ln(y_k)$	69	45	47	39	28	0.1113		
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	57	57	14	84	16	0.0006		
2SLS:	$\ln(P_i)*\ln(y_k)$	75	39	36	64	14	0.1923	0.0877 (0.0553)	0.1304** (0.0596)
	$\ln(y_i)*\ln(y_k)$	66	48	35	51	28	0.0823		
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	53	61	11	87	16	−0.0181		

Specification 1 replicates results shown in Table 1. Specification 2 substitutes income interaction for price interaction in Eq. (12). Specifications 3 and 4 estimate Eq. (17). Estimation by 2SLS uses lagged export price index (93–94) as instrument for current index (95–96). See Table 1 for description of columns, estimation of standard errors, and weights.
***, **, *Significant at the 1%, 5%, and 10% level, respectively.

whether it is quality or other factors correlated with income that drive the results, the next section relaxes the assumption that export prices are only determined by exporter quality.

4. Export price as “indicator” of quality

4.1. Empirical specification

Quality differences are presumably one of the main sources of cross-country variation in export prices. However, this variation might also reflect differences in prices for goods of the same quality, which might stem, for example, from differences in production costs. I postulate a reduced form specification for the determination of export prices that includes exporter income in addition to quality level. The inclusion of this variable attempts to capture cross-country variation in production costs systematically related to income. Distance of country i to the US is also included to control for selection bias in the quality composition of exports to the US. Export prices are thus assumed to be determined by:

$$\ln p_{iz} = \zeta_{0z} + \zeta_{1z}\ln\theta_{iz} + \zeta_{2z}\ln y_i + \zeta_{3z}\ln \mathrm{Dist}_i^{US} + \xi_{iz}. \tag{15}$$

Even though export prices are no longer exact measures of quality, they can still be used as "indicators" of quality (Wooldridge, 2002). The partial relationship between product quality and export price is given by ζ_{1z}. Since it is more costly to produce goods of higher than of lower quality, we assume that $\zeta_{1z}>0$. The sign of ζ_{2z} is instead ambiguous. Once we control for product quality, differences in comparative advantage are likely to drive any systematic relationship between income and production costs (and hence export prices). For example, high income countries often have a comparative advantage in capital-intensive sectors, as they tend to be capital abundant. Therefore, absent factor price equalization, we expect them to have a lower export price in those sectors ($\zeta_{2z}<0$). In contrast, we expect high income countries to have a higher export price in labor-intensive sectors ($\zeta_{2z}>0$), as they tend to have a comparative disadvantage in those sectors. There is, however, a strong caveat. Because of quality differences, sectors are likely to include both capital-intensive and labor-intensive goods. High-quality varieties might be capital intensive and low-quality varieties labor intensive. Thus, while the relationship between income and production costs is likely to vary within a sector according to quality, the parameter ζ_{2z} in (15) is only able to capture the sectoral average of this relationship.[23]

Hummels and Skiba (2004) provide evidence that the quality composition of exports increases with the distance between trading partners (the Alchian–Allen conjecture). In our sample of unit values of US imports, controlling for a country's average quality of exports to all countries (θ_{iz}), a larger distance to the US will imply a higher quality selection of exports to that market, and thus a higher observed price index p_{iz}. We thus expect $\zeta_{3z}>0$.

Solving for $\ln\theta_{iz}$ in (15), we obtain:

$$\ln\theta_{iz} = -\frac{\zeta_{0z}}{\zeta_{1z}} + \frac{1}{\zeta_{1z}}\ln p_{iz} - \frac{\zeta_{2z}}{\zeta_{1z}}\ln y_i - \frac{\zeta_{3z}}{\zeta_{1z}}\ln \mathrm{Dist}_i^{\mathrm{US}} - \frac{\xi_{iz}}{\zeta_{1z}}. \tag{16}$$

Substituting (16) into (12) yields the estimating equation:

$$\begin{aligned}\ln \mathrm{imp}_{iz}^k &= \varphi_{iz} + \psi_z^k - \tilde{\sigma}_z\eta_z\ln \mathrm{Dist}_i^k + \boldsymbol{\beta}_z\mathbf{I}_i^k + \zeta'_{1z}\ln p_{iz}\ln y^k \\ &\quad + \zeta'_{2z}\ln y_i\ln y^k + \zeta'_{3z}\ln \mathrm{Dist}_i^{\mathrm{US}}\ln y^k + \xi'_{iz}, \\ \zeta'_{1z} &= \frac{\tilde{\sigma}_z\mu_z}{\zeta_{1z}},\ \zeta'_{2z} = -\frac{\tilde{\sigma}_z\mu_z\zeta_{2z}}{\zeta_{1z}},\ \zeta'_{3z} = -\frac{\tilde{\sigma}_z\mu_z\zeta_{3z}}{\zeta_{1z}},\ \xi'_{iz} = -\frac{\tilde{\sigma}_z\mu_z}{\zeta_{1z}}\xi_{iz}\ln y^k + \tilde{\sigma}_z v_i^k.\end{aligned} \tag{17}$$

[23] Hallak (2003) discusses in more detail the idea of sectoral comparative advantage in the presence of quality differences.

There are three interaction terms in (17): the "price interaction" ($\ln p_{iz} \ln y^k$), the "income interaction" ($\ln y_i \ln y^k$), and the "distance (to the US) interaction" ($\ln \text{Dist}_i^{\text{US}} \ln y^k$). Since $\zeta_{1z}>0$ and $\tilde{\sigma}_z>0$, the sign of the coefficient on the "price interaction," ζ'_{1z}, corresponds to the sign of the coefficient on the parameter of interest, μ_z. Thus, we can still use the coefficient on this term to assess the effect of quality on the direction of trade. The parameter of interest μ_z is also a component of the coefficients associated with the remaining two interaction terms, ζ'_{2z} and ζ'_{3z}. However, in the first case, we cannot use the sign of ζ'_{2z} to test the sign of μ_z because the sign of ζ_{2z} is ambiguous. In the second case, since we assume that both $\tilde{\sigma}_z$ and ζ_{1z} are positive, $\zeta_{3z}>0$ and $\mu_z>0$ – as postulated – imply that $\zeta'_{3z}<0$. But we can only use the estimation of the last parameter to test joint hypotheses on the signs of μ_z and ζ_{3z}.

Since ξ_{iz} is a component of p_{iz}, the disturbance ξ'_{iz} and the regressor $\ln p_{iz} \ln y^k$ are correlated in (17). I therefore calculate $p_{iz,\,t-1}$ for the (merged) years 1993 and 1994, and use $\ln p_{iz,\,t-1} \ln y^k$ as an instrument for $\ln p_{iz} \ln y^k$. To the extent that ξ_{iz} captures classical measurement error in the price index, the instrument will be uncorrelated with the disturbance term. Since measurement error is substantial, I expect the use of this instrument to remove much of the correlation between regressor and error term. However, the disturbance ξ_{iz} might also capture omitted factors affecting export prices, not included in (15). For example, country i might have a technological advantage in sector z that allows it to produce at lower cost. In that case, ξ_{iz} will be persistent over time, and instrumenting with the lagged variable will not remove the correlation between regressor and error term. Since I have no alternative instrument, this is a concern to keep in mind when interpreting the results.

4.2. Estimation results

Specifications 3 and 4 in Table 3 show the results of estimating (17) by OLS and 2SLS. As before, the pooled regression constrains the coefficient on the price interaction to be the same across sectors. The coefficient on the income interaction (as those on all other variables) is allowed to vary across sectors to control for the (sector-specific) average relationship between income and production costs. Specification 3 restricts the coefficient on the distance (to the US) interaction to be zero. Specification 4 is the full model. We start by discussing the results of specification 3. Despite the inclusion of the income interaction in the regression, the price interaction retains considerable explanatory power. In the sectoral regressions using OLS, its sign is still positive in a majority of sectors, but the number of sectors with a significantly positive coefficient decreases substantially. In both the unweighted and the weighted pooled regressions, the coefficient on the price interaction is positive and significant.

When 2SLS is used to estimate, there is a slight increase in the number of sectors with a positive estimated coefficient on the price interaction, even though the results by significance level are almost unchanged. In contrast, the estimated coefficient on the income interaction is positive and significant in a substantially lower number of sectors (from 51 to 35). Comparing the results of specification 3 with those obtained under specifications 1 and 2, the estimates of the coefficient on the income interaction appear to be more sensitive to the inclusion of the price interaction than vice versa. This is consistent with our previous discussion. Once export prices are controlled for, the sign of the income interaction should depend on the relationship between income and production costs and should no longer be expected to be uniformly positive. The last two columns show the pooled regression results. In the unweighted regression, we cannot reject the null that $\mu=0$. This in part reflects the fact that, while the coefficient on the price interaction is positive in almost 2/3 of the sectors, it is still negative in a substantial number of them (39). But it might also be the case that the income interaction, because of its correlation with the price interaction, spuriously captures the quality effect in sectors where the price indices are not accurately measured. This interpretation is supported by the results of the last column, where both the magnitude and the significance of the coefficient increase as observations are weighted according to the precision of the price measure.

The last set of results corresponds to the full specification, where distance to the US is measured by the distance from capital cities to New York. The results do not change substantially. Focusing on the 2SLS estimates, the sectoral regressions show a slight increase in the number of positive and significant coefficients for the price interaction. The pooled regressions show that, while in the unweighted case the constrained coefficient on the price interaction is still not significantly different from zero (even though the p-value is 0.11), the coefficient on the price interaction is significant at the 5% level when observations are weighted according to the precision of the export price index. The estimates of the income interaction are almost identical. The estimated coefficient on the distance interaction is statistically insignificant in most sectors. Since the empirical specification captures the Alchian–Allen effect only indirectly – as it is not designed for that purpose – it is not surprising that it fails to identify this effect with precision.[24]

Table 4 shows the results of estimating Eq. (17) using the censoring model. As before, the median magnitude of the estimated coefficient on the price interaction is much closer to the estimates obtained in the pooled regressions (using either OLS or 2SLS) than to the results of the sectoral regressions. The distribution of the estimates by significance level is more similar to the OLS than to the 2SLS estimates, but the number of significant coefficients (both positive and negative) is larger when the censoring model is used to estimate.

The results of the full model confirm the findings of the previous section, even though it is much harder to identify the quality effect with precision when both the price and income interactions are included. In particular, the results indicate that export prices do not capture spurious effects unrelated to quality. On the contrary, since measurement error in the price variable is likely to be substantial, it might be income instead that captures spuriously the effect of quality when export prices are imprecisely measured. The next section looks for further evidence that it is the interaction of quality demand and quality

[24] Robustness tests using per capita GDP not adjusted for purchasing power parity as an alternative to PPP GDP, and using distance to the closest US coast (Los Angeles or New York) as an alternative to distance to New York yield very similar results. In the first case, the sectoral results by significance level are almost identical, but the coefficient on the price interaction term in the weighted pooled regression is only significant at the 10% level. In the second case, there is no significant difference to report.

Table 4
Price indices as indicators of quality differentiated goods (114 Sectors)—ML estimates

		Regressions by sector					
		Sign		Significance (5%)			Median
		Pos.	Neg.	Pos.	Not Sig.	Neg.	
1. Price only							
MLE:	$\ln(P_i)*\ln(y_k)$	73	40	45	55	13	0.1131
2. Income only							
MLE:	$\ln(y_i)*\ln(y_k)$	72	42	52	36	26	0.0797
3. Price and income							
MLE:	$\ln(P_i)*\ln(y_k)$	68	45	38	58	17	0.1068
	$\ln(y_i)*\ln(y_k)$	71	42	49	36	28	0.0842
4. Price, income, and distance to US							
MLE:	$\ln(P_i)*\ln(y_k)$	71	43	39	57	18	0.1117
	$\ln(y_i)*\ln(y_k)$	71	43	51	33	30	0.0883
	$\ln(\text{Dist}_i^{\text{US}})*\ln(y_k)$	64	50	22	77	15	0.0285

Estimation of censoring model (Eqs. (14) and (17)). Heteroskedasticity-robust standard errors in all regressions.

supply that drives the results, by estimating the model using reference-priced and homogeneous sectors.

5. Reference-priced and homogeneous goods

The intuitive idea that richer countries import relatively more from countries that produce higher quality goods is based on two premises. On the supply-side, it requires that there are indeed cross-country differences in product quality. On the demand side, it requires that richer countries consume higher quality goods. The extent to which the first premise is valid varies between types of goods. Assuming that quality differentiation can be approximated by the cross-country dispersion of export price indices, quality differentiation appears to be stronger for differentiated goods than for referenced-priced goods, and stronger for the latter than for homogeneous goods, as the average dispersion across sectors is, respectively, 0.43, 0.40, and 0.34 for those sectors.[25] The extent to which the second premise is valid for different types of goods depends on the intensity of preference for quality of consumers—or firms in the case of intermediate goods. Even though this is much harder to assess, it is interesting to note that, for example, while high-income consumers might be unwilling to buy low-quality furniture (a differentiated product) or producers of high-quality stationery might require high-quality paper (a reference-priced product) as intermediate input, consumers of several homogeneous products, such as oil refineries consuming crude oil or steel mills consuming iron ore or scrap iron, do not always need to match the quality of their products with the quality of the

[25] Dispersion measures are calculated taking the logarithm of the indices. Thus, their magnitude does not depend on which particular country is used to normalize.

intermediate inputs they use. This suggests that the second premise might be more relevant for differentiated or reference-priced than for homogeneous sectors.

In addition to the premises that support the intuitive idea that richer countries import relatively more from countries that produce higher quality goods, the particular empirical specification that is estimated – Eqs. (12) or (17) – is derived under the assumption that there are both vertical and horizontal components of product differentiation. Absent the horizontal component, goods of similar quality would be perfect substitutes. In that case, sectoral bilateral trade flows would tend to have corner solutions, as is the case with the predictions of the Heckscher–Ohlin model with transport costs. Richer countries would still import relatively more from countries that produce high-quality goods, but this would only be true on average, as they would import from only a few such countries. The estimating equations would then only capture this average effect. Since the classification into the three categories of goods distinguishes sectors mostly according to their degree of horizontal differentiation, we can expect the extent of this problem to increase as we move from differentiated to homogeneous goods. In sum, while we expect the theory to work for differentiated sectors, we expect it not to work for homogeneous sectors. For reference-priced sectors, the theoretical prediction is ambiguous, as we do not know the extent to which the premises of the theory apply to them.

5.1. Estimation results

Table 5 shows the results of estimating (17) by OLS and 2SLS for reference-priced and homogeneous sectors. For reference-priced sectors, the distribution of the estimated coefficient on the price interaction by sign and significance level resembles that obtained for differentiated sectors (shown in Table 3). When estimated with OLS, the coefficient is positive in 32 sectors. When estimated with 2SLS, it is positive in 35 sectors. In both cases, the proportion of positive coefficients is approximately 2/3. The coefficient is positive and significant in 14 and 15 sectors, respectively, less than 1/3 of the sectors. The median magnitude of the estimated coefficient is surprisingly high; it is almost double the median for differentiated goods. This result is confirmed by the pooled regression estimates, where the coefficient is always positive and significant at the 1% level. Even though the estimated sign is consistent with the theory, the fact that the quality effect is stronger for reference-priced than for differentiated sectors appears at odds with our previous discussion. One possible explanation is that the coefficient that is estimated is ζ'_{1z}, which contains $\tilde{\sigma}$. Therefore, if the elasticity of substitution is higher for reference-priced sectors, then ζ'_{1z} can be higher even though μ_z is lower.

The coefficient on the income interaction term is now more often negative than positive. The median magnitude is also negative, indicating that higher income is associated with comparative disadvantage in a higher proportion of reference-priced sectors than of differentiated sectors. Finally, the distance interaction gives stronger support than before to the Alchian–Allen conjecture.

When (17) is estimated on homogeneous sectors, the results strongly accord with the predictions. In particular, there is no quality effect on trade. The price interaction term is now more often negative than positive, and in not one sector is it positive and significant. In addition, the median of the OLS and 2SLS estimates is negative, and the estimated coefficient in the pooled regression is negative in three out of four cases, in none of which it is significant. Since these are the goods with characteristics most clearly at odds with the assumptions of the theoretical framework, these results provide additional evidence that it is the interaction between quality supply and quality demand postulated by the theory that explains the results for differentiated sectors.

Finally, Table 6 shows the estimates using the censoring model. Here, the puzzling results for reference-priced goods disappear. In particular, the median estimate of the price interaction for those sectors is no longer higher than the median estimate for differentiated goods. In the case of homogeneous goods, the estimates strongly resemble those obtained under OLS and 2SLS.

6. Conclusions and further comments

A substantial amount of theoretical work predicts that product quality plays an important role as a determinant of the direction of trade. This paper provides evidence supporting the empirical relevance of this prediction. First, it builds a theoretical

Table 5
Price indices as indicators of quality reference-priced goods (51 sectors) and homogeneous goods (38 sectors)—OLS and 2SLS estimates

		Regressions by sector						Pooled regression	
		Sign		Significance (5%)			Median	Coefficient (unweighted)	Coefficient (weighted)
		Pos.	Neg.	Pos.	Not Sig.	Neg.			
Reference-priced goods									
Price, income, and distance to the US									
OLS:	$\ln(P_i)*\ln(y_k)$	32	19	15	34	2	0.2843	0.1728*** (0.0429)	0.2716*** (0.0557)
	$\ln(y_i)*\ln(y_k)$	23	28	8	36	7	−0.0324		
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	19	32	5	35	11	−0.0544		
2SLS:	$\ln(P_i)*\ln(y_k)$	35	16	14	33	4	0.3608	0.3066*** (0.0679)	0.3868*** (0.0771)
	$\ln(y_i)*\ln(y_k)$	23	28	9	32	10	−0.0291		
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	19	32	5	35	11	−0.0436		
Homogeneous goods									
Price, income, and distance to the US									
OLS:	$\ln(P_i)*\ln(y_k)$	17	21	0	34	4	−0.0648	−0.0211 (0.0857)	0.0698 (0.0936)
	$\ln(y_i)*\ln(y_k)$	11	27	0	31	7	−0.1528		
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	11	27	1	32	5	−0.0711		
2SLS:	$\ln(P_i)*\ln(y_k)$	17	21	0	37	1	−0.1355	−0.0839 (0.1494)	−0.0709 (0.1519)
	$\ln(y_i)*\ln(y_k)$	12	26	0	36	2	−0.1369		
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	10	28	1	35	2	−0.1186		

Estimation of Eq. (17). Estimation by 2SLS uses lagged export price index (93–94) as instrument for current index (95–96). See Table 1 for description of columns, estimation of standard errors, and weights.
***, **, *Significant at the 1%, 5%, and 10% level, respectively.

framework that yields predictions on the effect of quality on bilateral trade flows. In the empirical specification, this effect is captured by a term interacting exporter quality and importer income per capita. A key aspect of the paper is its strategy for identifying unobserved cross-country differences in product quality, which is measured by a price index based on cross-country variation in export unit values. The empirical specification is then tested using a cross-section of bilateral trade flows between 60 countries.

Concerns about forces other than quality driving the results are addressed in two ways. First, it is shown that the interaction term between exporter price and importer per capita income retains most of its explanatory power even after a similar interaction term in which exporter income substitutes for exporter price is also included. Second, the estimation is also performed using reference-priced and homogeneous sectors. For reference-priced sectors, where the empirical specification is expected to capture the effect of quality only on average, the results resemble those obtained for differentiated sectors. Surprisingly, they appear to be even stronger. This is not true when a censoring model, which accounts for zero-valued trade flows, is used for estimation. For homogeneous sectors, where the theory should not apply, no quality effect on trade is identified in the data. An overall assessment of the evidence presented in the paper provides a compelling case for quality as a significant factor explaining global patterns of bilateral trade.

Further research will hopefully improve many aspects of the theoretical framework and empirical strategy used here. First, an advantage of the proposed demand system is its simplicity for capturing income effects on quality choice. However, it could be extended to allow for more flexible substitution patterns, in particular, closer substitutability between goods of similar quality. Second, as suggested by the theoretical work on trade with vertical differentiation, it is not merely the mean but the entire distribution of income that matters as a determinant of international trade. Third, the accuracy and reliability of the export price indices used in this study are as of yet untested. Since unit values are likely to be the best, although indirect, available source of information on cross-country differences

Table 6
Price indices as indicators of quality reference-priced (51 sectors) and homogeneous goods (38 sectors)—ML estimates

		Regressions by sector					
		Sign		Significance (5%)			Median
		Pos.	Neg.	Pos.	Not Sig.	Neg.	
Reference-priced goods							
Price, income, and distance to US							
MLE:	$\ln(P_i)*\ln(y_k)$	30	21	16	30	5	0.1144
	$\ln(y_i)*\ln(y_k)$	28	23	13	31	7	0.0099
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	19	32	5	40	6	−0.0236
Homogeneous goods							
Price, income, and distance to US							
MLE:	$\ln(P_i)*\ln(y_k)$	17	21	2	30	6	−0.0554
	$\ln(y_i)*\ln(y_k)$	8	30	1	26	11	−0.1458
	$\ln(\mathrm{Dist}_i^{\mathrm{US}})*\ln(y_k)$	9	29	3	25	10	−0.1676

Estimation of censoring model (Eqs. (14) and (17)). Heteroskedasticity-robust standard errors in all regressions.

in quality levels covering a broad range of goods, further research focused on these indices seems necessary and promising.

Acknowledgements

I am grateful to Elhanan Helpman and Marc Melitz for their guidance and support, to Alan Deardorff for very helpful discussions and comments, and to Peter Morrow for excellent research assistance. I also thank Robert Barro, Francesco Caselli, Edwin Diewert, Doireann Fitzgerald, Daniel Hamermesh, Jim Levinsohn, Rosa Matzkin, Jim Rauch, Gary Solon, Silvana Tenreyro, anonymous referees, and seminar participants at Columbia, Di Tella, Duke, Harvard, Michigan, Michigan State, Princeton, Purdue, San Andrés, Texas-Austin, Toronto, University of British Columbia, the World Bank, Yale, the EIIT Conference, and the NBER Summer Institute for helpful comments and suggestions.

Appendix A. Methodology of construction of export price indices

A representative 2-digit sector z includes 10-digit categories $n=1, \ldots, N$. Dividing the value by the quantity of imports in each of these categories, we obtain the unit value (average export price) p_{in} for country i in category n. Based on 10-digit-level prices p_{in}, I will construct multilateral price indices p_{iz} at the 2-digit level. Country i is "active" in category n if p_{in} is non-missing. Country i is active in sector z if it has at least two active categories. Otherwise, p_{iz} takes a missing value.

There is considerable measurement error in the source database. I use the following procedure to detect and remove outliers. For each category n, I calculate the geometric mean of unit values across countries, excluding the observations with maximum and minimum values. I then remove observations with unit values 4 times above or below the mean. Since observations with extreme unit values show disproportionately low export quantities, I remove observations with quantity below the minimum of 50 units or a quarter of the average quantity for the category.

I then take country j as a numeraire. For each other country i, I calculate the bilateral Fisher price index P_i^j between i and j, using only their common active categories (P_i^j takes a missing value if i and j do not have any common active category). As a result, I obtain a vector of bilateral price indices P_j, with country j as the numeraire. I repeat this procedure taking alternatively all countries as numeraire, and obtain vectors P^j, $j=1, \ldots, C$, where C is the number of active countries in the sector. These vectors are separated into three groups. The first group contains vectors $j=1, \ldots, C_1$, those with no missing values. The second group contains vectors $j=C_1+1, \ldots, C_2$, with at most 5 missing values. The third group contains the remaining vectors $j=C_2+1, \ldots C$.

I then follow a three step procedure. I take the first group and normalize each vector to sum up to 1. Denote normalized vectors by $\tilde{P}^j$. The geometric weighted average of these vectors is:

$$P_z^1 = \prod_{j=1}^{C} \left(\tilde{P}^j\right)^{w^j}, \tag{18}$$

where w^j is the number of active categories of country j in sector z. In the second step, I take each vector P^j in the second group and normalize it to sum up to $1-m^j$, where m^j is the sum

Table A1
List of 3-digit SITC sectors included in each sample (some 4-digit sectors may be excluded)

Differentiated sample (114 sectors)									
034	048	056	073	098	111	248	431	533	541
551	553	554	591	598	611	612	621	625	628
635	642	651	652	653	654	655	656	657	658
659	661	662	663	665	666	667	673	678	679
691	692	694	695	696	697	699	711	713	714
716	721	722	723	724	725	726	727	728	736
741	742	743	744	745	749	751	752	759	761
762	764	771	772	773	774	775	776	778	781
782	783	784	785	786	791	792	793	812	821
831	842	843	844	845	846	847	848	851	871
872	874	881	882	884	885	892	893	894	895
896	897	898	899						
Reference-priced sample (51 sectors)									
011	014	034	036	037	054	057	058	062	081
112	122	233	266	278	334	335	341	511	512
513	514	515	516	522	523	524	533	541	562
582	583	592	634	641	642	651	652	653	661
662	671	672	673	674	677	678	682	684	699
778									
Homogeneous sample (38 sectors)									
011	022	023	024	041	042	043	044	054	057
058	061	071	121	222	232	247	251	263	268
281	282	287	288	333	334	423	424	522	634
651	667	681	682	684	686	689	971		

of the entries in P_z^1 corresponding to the missing elements in P_j. I then impute the values in P_z^1 to the missing elements in P^j, thus obtaining a normalized (to 1) vector $\tilde{P}^j$. Using normalized vectors $\tilde{P}^j, j=1, \ldots C_2$, I recalculate (18) and obtain P_z^2. In the third step, I repeat the procedure for the remaining vectors. I finally calculate $p_z = \Pi_{j=1}^{C} \left(\tilde{P}^j\right)^{w^j}$, where each element is p_{iz}.

References

Bergstrand, J., 1989. The generalized gravity equation, monopolistic competition, and the factor–proportions theory of international trade. Review of Economics and Statistics 71, 143–153.

Brooks, E., 2003. Why Don't Firms Export More? Product Quality and Colombian Plants UC Santa Cruz, mimeo.

Cogan, J., 1981. Fixed costs and labor supply. Econometrica 49, 945–963.

Dalgin, M., Mitra, D., Trindade, V., 2004. Inequality, nonhomothetic preferences, and trade: a gravity approach. Mimeo.

Deardorff, A., 1984. Testing trade theories and predicting trade flows. In: Jones, R., Kenen, P. (Eds.), Handbook of International Economics, vol. 1. Elsevier, Amsterdam, pp. 467–518.

Deardorff, A., 1998. Determinants of bilateral trade: does gravity work in a neoclassical world? In: Frankel, J. (Ed.), The Regionalization of the World Economy. Chicago University Press, Chicago.

Diewert, W., 1993. Index numbers. In: Diewert, W., Nakamura, A. (Eds.), Essays in Index Number Theory, vol. 1. Elsevier Science Publishers B.V., Amsterdam.

Falvey, R., Kierzkowski, H., 1987. Product quality, intra-industry trade and (Im)perfect competition. In: Kierzkowski, H. (Ed.), Protection and Competition in International Trade. Basil Blackwell, Oxford, pp. 143–161.

Feenstra, R., 2000. World Trade Flows, 1980–1997. Center for International Data, Institute of Governmental Affairs, UC Davis.

Feenstra, R., Romalis, J., Schott, P., 2002. US Imports, exports, and tariff data, 1989–2001. NBER Working Paper 9387.

Flam, H., Helpman, E., 1987. Vertical product differentiation and North–South trade. American Economic Review 77, 810–822.

Francois, J., Kaplan, S., 1996. Aggregate demand shifts, income distribution, and the Linder hypothesis. Review of Economics and Statistics 78, 244–250.

General Accounting Office, 1995. US Imports: Unit Values Vary Widely for Identically Classified Commodities. Report GAO/GGD-95-90.

Hallak, J., 2003. The Effect of Cross-country Differences in Product Quality on the Direction of International Trade. RSIE Working Paper 493.

Helpman,H., Melitz, M., Rubinstein, Y., 2004. Trading Partners and Trading Volumes. Mimeo.

Hummels, D., Klenow, P., 2005. The Variety and Quality of a Nation's Exports. American Economic Review 95, 704–723.

Hummels, D., Skiba, A., 2004. Shipping the good apples out? An empirical confirmation of the Alchian–Allen conjecture. Journal of Political Economy 112, 1384–1402.

Hunter, L., 1991. The contribution of non-homothetic preferences to trade. Journal of International Economics 30, 345–358.

Hunter, L., Markusen, J., 1988. Per-capita income as a determinant of trade. In: Feenstra, R. (Ed.), Empirical Methods for International Trade. The MIT Press, Cambridge, MA, pp. 89–109.

Leamer, E., Levinsohn, J., 1995. International trade theory: the evidence. In: Grossman, G., Rogoff, K. (Eds.), Handbook of International Economics, vol. 3. North-Holland, Amsterdam, pp. 1339–1394.

Linder, S., 1961. An Essay on Trade and Transformation. Almqvist & Wiksell, Stockholm.

Markusen, J., 1986. Explaining the volume of trade: an eclectic approach. American Economic Review 76, 1002–1011.

Matsuyama, K., 2000. A Ricardian model with a continuum of goods under nonhomothetic preferences: demand complementarities, income distribution, and North–South trade. Journal of Political Economy 108, 1093–1120.

McPherson, M., Redfearn, M., Tieslau, M., 2001. International trade and developing countries: an empirical investigation of the Linder hypothesis. Applied Economics 33, 649–657.

Murphy, K., Shleifer, A., 1997. Quality and Trade. Journal of Development Economics 53, 1–15.

Rauch, J., 1999. Networks versus markets in international trade. Journal of International Economics 48, 7–35.

Romalis, J., 2004. Factor proportions and the structure of commodity trade. American Economic Review 94, 67–97.

Schott, P., 2004. Across-product versus within-product specialization in international trade. Quarterly Journal of Economics 119, 647–678.

Shatz, H., 1997. Note on Distance Data. Harvard University, mimeo.

Stokey, N., 1991. The volume and composition of trade between rich and poor countries. Review of Economic Studies 58, 63–80.

Verhoogen, E., 2004. Trade, Quality Upgrading and Wage Inequality in the Mexican Manufacturing Sector: Theory and Evidence from an Exchange-Rate Shock. Columbia University, mimeo.

Wooldridge, J., 2002. Econometric analysis of cross section and panel data. The MIT Press, Cambridge, MA.