NBER WORKING PAPER SERIES

A PRODUCT-QUALITY VIEW OF THE LINDER HYPOTHESIS

Juan Carlos Hallak

Working Paper 12712
http://www.nber.org/papers/w12712

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2006

I thank Alan Deardorff for numerous discussions and comments on the contents of this paper. I also thank Martin Daneri, Elhanan Helpman, Jim Levinsohn, Marc Melitz, Gary Solon, Walter Sosa Escudero, and seminar participants at Michigan State, Princeton, Syracuse and the NBER Summer Institute for helpful comments and suggestions. Tomas Castagnino, Isao Kamata, and Alejandro Molnar provided outstanding research assistance.

A Product-Quality View of the Linder Hypothesis
Juan Carlos Hallak
NBER Working Paper No. 12712
December 2006
JEL No. D12,F1,F12

ABSTRACT

The Linder hypothesis states that countries of similar income per capita should trade more intensely with one another. This hypothesis has attracted substantial research over decades, but the empirical evidence has failed to provide consistent support for it. This paper shows that the reason for the failure is the use of an inappropriate empirical benchmark, the gravity equation estimated using trade data aggregated across sectors. The paper builds a theoretical framework in which, as in Linder's theory, product quality plays the central role. A formal derivation of the Linder hypothesis is obtained, but this hypothesis is shown to hold only if it is formulated as a sector-level prediction. The "sectoral Linder hypothesis" is then estimated on a sample of 64 countries in 1995. The results support the prediction: after controlling for inter-sectoral determinants of trade, countries of similar per-capita income trade more intensely with one another. The paper also shows that a systematic aggregation bias explains the failure of the previous empirical literature to find support for Linder's theory.

Juan Carlos Hallak
Department of Economics
University of Michigan
351C Lorch Hall
Ann Arbor, MI 48109
and NBER
hallak@umich.edu

1 Introduction

There is ample evidence that the quality of goods that countries produce and consume varies systematically with their income levels. On the production side, export prices are strongly correlated with countries' income per capita, suggesting a positive relationship between per-capita income and quality production [Schott (2004), Hummels and Klenow (2005), Hallak (2006)]. On the consumption side, household data shows that quality demand is strongly correlated with household income [e.g. Bils and Klenow (2001)], suggesting that, in the aggregate, high income countries consume larger proportions of high quality goods. A positive relationship between countries' per-capita income and quality consumption is also found in studies based on international trade data.[1] The evidence of systematic supply-side and demand-side relationships between income per capita and product quality indicates a potentially important role for product quality as a determinant of bilateral trade patterns. Even though workhorse theories of international trade neglect this role, growing interest in product quality has spurred theoretical research on the interaction between quality and trade. This interaction is also a subject of interest in policy-oriented research. In particular, international organizations are concerned about the prevalence of quality standards imposed by firms and governments in developed countries that firms in developing countries cannot easily attain.[2]

Long before empirical evidence of systematic cross-country variation in quality production and consumption became available, Linder (1961) introduced the idea of quality as a determinant of the direction of trade. On the demand side of Linder's theory, countries with high income per capita spend a larger fraction of their income on high quality goods. On the supply side, countries develop a comparative advantage in the goods that are in high domestic demand. Thus, high (low) income countries develop a comparative advantage in the production of high (low) quality goods because those are the goods that are highly demanded in their domestic markets. Demand and supply are combined to argue that the overlap of production and consumption patterns between countries of *similar* income per capita should induce them to trade more intensely with one another,

[1] Brooks (2005) and Verhoogen (2006) provide evidence based on U.S. imports from Colombia and firm-level exports of Mexico, respectively. Hallak (2006) and Choi et al. (2006) provide evidence based on bilateral import patterns.

[2] See World Bank (1999) and WTO (2005). See also Maskus et al. (2005) and Chen et at. (2006).

a prediction commonly known as "the Linder hypothesis".

The Linder hypothesis has attracted the attention of scholars for decades due to its sharp contrast with the predictions of the Heckscher-Ohlin (or factor proportions) theory – the usual benchmark for most empirical work on determinants of trade patterns and effects of trade policies – which suggests more intense trade between countries of *dissimilar* income per capita. In light of the strong evidence confirming the empirical relevance of both sides of Linder's theory, however, it is puzzling that decades of testing the Linder hypothesis have failed to provide consistent support of its empirical validity.[3] This paper solves the puzzle and shows that product quality plays an important role in driving patterns of bilateral trade.

Several theoretical studies have formulated models capturing the key insights of Linder's theory.[4] In those models, high income countries have a comparative advantage in the production of high quality goods and consume those goods in greater proportions.[5] However, this literature does not yield the Linder hypothesis since all the models consider a world with only two countries while the Linder hypothesis involves at least four.

The paper presents a theoretical framework that captures the role of quality emphasized by Linder and delivers his conjecture as a formal prediction. However, while the Linder hypothesis is stated as a prediction for bilateral trade patterns at the aggregate level, this paper shows that it can only be expected to hold at the sector level. In contrast to the results obtained from standard estimates of the Linder hypothesis – using aggregate trade data – the empirical results of this paper provide support for the "sectoral Linder hypothesis".

In its standard (aggregate) formulation, the Linder hypothesis confounds the impact of two forces. First, product quality affects production and consumption patterns *within* sectors (e.g. high quality furniture versus low quality furniture). Second, other determinants of trade patterns, such as those at the center of traditional trade theories (e.g. the Ricardian and Heckscher-Ohlin models), operate and vary *between* sectors. Consequently, tests of the hypothesis at the aggregate

[3] Deardorff (1984), Leamer and Levinsohn (1995), and McPherson, Redfearn, and Tieslau (2001) survey the empirical literature.

[4] Falvey and Kierzkowski (1987), Flam and Helpman (1987), Stokey (1991), Copeland and Kotwel (1996), and Murphy and Shleifer (1997).

[5] The source of comparative advantage in quality supply varies across models, but high income countries are always assumed to possess a comparative advantage in the production of high quality goods.

level do not provide a meaningful test for the role of quality on bilateral trade patterns. The sectoral Linder hypothesis does capture the role of quality because its specification controls for inter-sectoral determinants of trade.

In this paper, countries are modeled as differing in their valuation of quality. In particular, a parameter governs the variation in countries' expenditure shares across varieties of different quality. On the supply side, since several plausible theories relate quality production to per-capita income, their common implications are captured by simply assuming that quality supply and per-capita income are systematically related. The interaction of demand for and supply of quality results in a gravity-type specification in which sectoral bilateral trade is a function of exporter and importer fixed effects, proxies for bilateral trade costs, and a "Linder term" measuring income (per-capita) dissimilarity between pairs of countries. When both quality demand and quality supply increase with income per capita – as postulated by Linder – the model predicts a negative coefficient on the Linder term. This prediction is the formal characterization of the sectoral Linder hypothesis.

The sectoral Linder hypothesis is estimated using a sample of bilateral trade flows at the 3-digit level among 64 countries during 1995. Based on Rauch (1999), I classify sectors as "Differentiated", "Reference-priced" and "Homogeneous", and use only the first group for the baseline estimates because those are the goods with characteristics that most closely match the theoretical assumptions. The empirical results support the prediction of the theory: countries with similar income per capita tend to trade more intensely with one another.

The unanimous empirical framework in the preceeding literature has been the gravity equation. This framework is inappropriate because it is specified at the aggregate level and is therefore unable to control for sector-specific determinants of trade. The omission of those controls is not innocuous. Since patterns of sectoral specialization within manufacturing are strongly correlated with income per capita, the inability to include such controls induces a systematic upward bias in the estimated coefficient of the Linder term. The bias increases with the level of aggregation at which the Linder hypothesis is estimated. In particular, when the Linder hypothesis is estimated aggregating all trade flows in differentiated products, the bias is sufficiently strong to reverse the estimated sign of the Linder effect.

A related theoretical literature also focuses on the aggregate Linder hypothesis but does not rely on product quality as its driving force. This literature emphasizes the role of inter-sectoral non-homotheticities in demand. Such non-homotheticities can generate higher trade intensity between countries of similar income per capita if per-capita income and comparative advantage in high-elasticity sectors are positively related.[6] While inter-sectoral non-homotheticities might generate trade flows consistent with the "aggregate" Linder hypothesis, the gravity-equation framework using aggregate data yields results that do not support the empirical prevalence of this effect.[7]

This paper simultaneously supports Linder's intuition on the link between product quality and the direction of trade and explains the failure to find consistent empirical support for the Linder hypothesis. The data bear out the premises of Linder's theory: countries of similar income have similar production and consumption patterns – they produce and consume goods of higher quality. But the hypothesized corollary that countries of similar income trade more with each other only holds at the sector level, where inter-sectoral determinants of trade are controlled for. At the aggregate level, the Linder hypothesis does not follow from the premises that are supposed to imply it. When it is properly formulated and estimated, it is empirically supported.

Section 2 of the paper follows the standard approach for estimating the Linder hypothesis (using aggregate trade data) and finds that it is strongly rejected. Section 3 develops a theoretical framework from which the sectoral Linder hypothesis is derived. Section 4 shows that the empirical estimates are consistent with the sectoral Linder hypothesis. Section 5 provides theoretical and empirical evidence on the existence of a systematic aggregation bias. Section 6 concludes.

2 The "aggregate" Linder hypothesis

The "Linder hypothesis" states that countries with similar income per capita trade more intensely with one another. The standard approach for testing this hypothesis relies on a gravity-equation framework, estimated using trade data aggregated across sectors. To obtain comparable results, this section follows the standard approach. Later sections demonstrate that the empirical framework is inappropriate, and explain why it fails to identify the role of product quality as a determinant of bilateral patterns of trade.

[6] See Markusen (1986), Bergstrand (1989), Deardorff (1998), and Matsuyama (2000). Hunter and Markusen (1988) and Hunter (1991) assess the empirical relevance of these non-homotheticities.

[7] While potentially relevant, this effect might be overshadowed by the (opposite) effect of other inter-sectoral determinants of trade.

Before presenting estimation results, this section describes some basic facts about the direction of world trade. Table 1 reports the volume of trade in 1995 between two groups of countries, "rich" and "poor", divided according to their income per capita. The cut-off value used in the table is a (PPP adjusted) GDP per-capita of $ 12000 U.S. dollars.[8] Panel (a) includes all countries with population above one million (107 countries), of which 26 countries are classified as "rich" and 81 countries are classified as "poor". Total trade (both exports and imports) among these 107 countries in 1995 is 5049 billion dollars, 97% of total world trade. High-income countries account for 81.1% of total exports ($ 4097 billion) and 78.8% of total imports ($ 3980 billion).

The Linder hypothesis predicts relatively more trade between countries of similar income per capita. On a first pass, this hypothesis can be assessed by comparing import and export shares between groups. Those shares are presented in the second set of columns. Rich countries buy 81.1% of their imports from other rich countries while they buy the remaining 18.9% from poor countries. Poor countries buy 81.5% of their imports from rich countries and 18.5% from other poor countries. Apparently at odds with the predictions of the Linder hypothesis, rich (poor) countries as a group do not import relatively more from other rich (poor) countries. The last set of columns provides the same message viewed from an exporter perspective. Rich countries export (slightly) relatively more to poor countries while poor countries export relatively more to rich countries.

The remainder of the paper focuses on a sample of 64 countries, those with population above 5 million and imports of "Differentiated" goods (Rauch 1999) above US$ 2 billion.[9] This sample includes 26 rich countries and 38 poor countries. The list of countries and their PPP GDP in 1995 is provided in Table A1. Panel (b) is analogous to panel (a), but is based only on trade flows between these 64 countries. As is the case in the larger sample, countries do not trade more with countries of similar per-capita income. For example, while poor countries' share of rich countries' imports is 16.8%, their share of poor countries' imports is 15.8%.[10]

Although informative about the structure of world trade, Table 1 cannot be used to assess

[8] This cut-off level is chosen as a threshold because it is located at the trough of the (bimodal) distribution of countries' per-capita income.

[9] I defer to Section 4 the description of data sources, classification of sectors into goods categories, and selection criteria for countries and sectors.

[10] The choice of an alternative cut-off value of US$ 15000, in which case Greece, Portugal, Taiwan, and South Korea switch to the group of poor countries, yields the same qualitative result.

the empirical validity of the Linder hypothesis. Countries of similar income per capita tend to be geographically close to each other. Since trade costs are not controlled for, more intense trade between countries of similar income per capita may be due to their lower bilateral transport costs rather than to income similarity. The empirical literature has long acknowledged this concern, and addressed it by adopting a gravity-equation framework in which bilateral trade is expressed as a function of the incomes of the trading partners and of proxies for their bilateral trade costs. The gravity specification is then augmented to include a "Linder term", a measure of income dissimilarity between pairs of countries. The specification that is typically estimated is

$$\ln(imports_{ij}) = \beta_0 + \beta_1 \ln(Y_i) + \beta_2 \ln(Y_j) + \beta_3 \ln(Dist_{ij}) + \boldsymbol{\beta}_4 \mathbf{I}_{ij} + \beta_5 Linder_{ij} + \varepsilon_{ij} \tag{1}$$

where i denotes the exporter, j denotes the importer, Y_i (Y_j) is the GDP of country i (j), $Dist_{ij}$ is the geographic distance between countries i and j, $\mathbf{I}_{ij}$ is a vector of dichotomous proxy variables for trade costs, and $Linder_{ij}$ is the Linder term. Several Linder terms have been proposed in the literature. The Linder term that will be used as baseline in this study is $Linder_{ij} = (\ln y_i - \ln y_j)^2$, where y_i (y_j) is the income per capita of country i (j). The Linder term is larger the more dissimilar are the two countries' incomes. Therefore, the prediction of the Linder hypothesis is that $\beta_5 < 0$. In its most simple form, this equation is estimated using ordinary least squares (OLS).

I follow the empirical literature by estimating equation (1) with OLS using the sample of 64 countries. The vector $\mathbf{I}_{ij}$ includes dummies indicating whether a country pair shares a common border, a common language, a preferential trade agreement (PTA), a colony-colonizer relationship, or a common colonizer. The results, presented in the first column of Table 2, strongly reject the Linder hypothesis. Instead of a negative coefficient on the Linder term, the estimated coefficient is positive and significant at the 1% level. The estimated coefficients on the other variables have the predicted sign in all cases. Larger GDP is associated with more trade, as are lower trade costs, induced by shorter distance or by sharing a border, common language, PTA, colonizer-colony relationship, or a common colonizer. The remaining columns of the table report results from estimating equation (1) including alternative Linder terms often used in the literature: $|y_i - y_j|$, $\ln |y_i - y_j|$, and $|\ln y_i - \ln y_j|$. The results are not sensitive to which Linder term is used. In contrast to the prediction of the Linder hypothesis, it is countries with *dissimilar* income per capita that appear to trade more with one another. In all cases, the coefficient on the Linder term is positive and significant at the 1% level.

Theoretical derivations of the gravity equation (e.g. Eaton and Kortum 2002, Anderson and Van Wincoop 2003) indicate that its standard formulation should be augmented to include exporter and importer fixed effects. For example, bilateral trade volumes are also influenced by the importer's "multilateral resistance", a common importer effect (across exporters) capturing trade costs between the importer and third countries. I thus re-estimate equation (1) including exporter and importer fixed effects.[11] The results are displayed in Table 3. Even though the magnitude and precision of the estimated coefficient on the Linder term decreases substantially with the inclusion of the fixed effects, the coefficient is still uniformly positive, contradicting the prediction of the Linder hypothesis.

The estimation exercises of this section have the sole objective of reproducing standard empirical tests of the Linder hypothesis. Since the literature fails to consistently support the hypothesis, the failure to find support for it here also is not surprising.[12] I will show next that this result is not due to the absence of the effects described by Linder, but due to a misspecified empirical framework to test their impact on trade.

3 The "sectoral" Linder hypothesis: a theoretical framework based on quality

Product quality plays a central role in Linder's theory. This section describes a theoretical framework that captures the essential aspects of that theory by allowing a systematic relationship between countries' income per capita and their supply of and demand for quality. Building on the results of previous theoretical work – which formalized Linder's ideas but did not generate testable predictions in a multilateral context due to a focus on a two-country equilibrium – this framework yields a formal derivation of the Linder hypothesis. In contrast to its standard formulation, the Linder hypothesis is shown to be valid only when formulated at the sector level, after controlling for inter-sectoral determinants of trade.

[11] Note that Y_i and Y_j can no longer be included as separate regressors after the inclusion of exporter and importer fixed effects.

[12] Thursby and Thursby (1987) finds suppoting results. However, since the study is based on a sample of only 17 countries, most of them OECD members, it does not exploit the most informative source of cross-country variation in per-capita income levels.

3.1 Demand for quality

The demand side of the model is based on Hallak (2006). Preferences are represented by a two-tier utility function. The upper tier is weakly separable into sectoral subutility indices $\{u_{jz}\}_{z=1,\dots,Z}$ and $\{u_{jg}\}_{g=Z+1,\dots,S}$, where j indexes countries, z indexes differentiated-good sectors, and g indexes homogeneous-good sectors. The utility derived from consuming goods in differentiated-good sector z is given by

$$u_{jz} = \left[\sum_{h \in \Omega_z} (\theta_h)^{\gamma_{jz}} q_{jh} \right]^{\frac{\sigma_z - 1}{\sigma_z}}, \qquad \theta_h, \gamma_{jz} > 0, \sigma_z > 1 \tag{2}$$

where Ω_z is the set of all varieties in sector z, θ_h is the quality of variety h, q_{jh} is the quantity consumed by country j of variety h, γ_{jz} is the intensity of country j's preference for quality (in sector z), and σ_z is the common elasticity of substitution between varieties in the sector. Demand in differentiated sectors is obtained in two stages. In the first stage, conditional on prices and income, the representative consumer chooses an allocation of expenditures across sectors $\{E_{jz}\}_{z=1,\dots,Z}$ and $\{E_{jg}\}_{g=Z+1,\dots,S}$. In the second stage, conditional on total expenditure E_{jz} in sector z, he chooses expenditure e_{jh} on variety h, which is given by

$$e_{jh} = s_{jh} E_{jz}, \qquad s_{jh} = \left(\frac{\frac{p_{jh}}{(\theta_h)^{\gamma_{jz}}}}{G_{jz}} \right)^{1-\sigma_z} \tag{3}$$

where s_{jh} is the share of variety h in sectoral expenditure E_{jz}, p_{jh} is the price of variety h faced by consumers in country j, and

$$G_{jz} = \left[\sum_{r \in \Omega_z} \left(\frac{p_{jr}}{(\theta_r)^{\gamma_{jz}}} \right)^{1-\sigma_z} \right]^{\frac{1}{1-\sigma_z}}$$

is an exact consumption price index for sector z in country j.

Product quality (θ_h) is modeled as a utility (or demand) shifter. This utility shifter captures all attributes of a product that consumers value. Conditional on price, higher quality increases the share spent on a given variety. The parameter γ measures intensity of preference for quality. Countries with higher γ are willing to pay more for high quality varieties. The impact of γ on expenditure shares is given by

$$\frac{\partial s_{jh}}{\partial \gamma_{jz}} = (\sigma_z - 1) s_{jh} \left[\ln \theta_h - \sum_{r \in \Omega_z} s_{jr} \ln \theta_r \right]. \tag{4}$$

Since $\sigma_z > 1$ and $s_{jh} > 0$, the sign of this derivative depends on the sign of the term in square brackets, i.e. on whether the quality of variety h is above or below the (weighted) average quality of the sector. When the quality of variety h is higher (lower) than the average quality, the term in brackets is positive (negative). In that case, expenditure shares increase (decrease) with the intensity of preference for quality.

Variation across countries in the value of γ provides this demand system with the flexibility to capture cross-country variation in quality demand. Countries with higher γ spend a larger proportion of their income on high quality goods. Such flexibility is convenient for capturing – in reduced form – the demand side of Linder's theory, which relates countries' demand for quality to income per capita. The demand system also embeds a relevant special case. When the intensity of preference for quality does not vary across countries ($\gamma_{jz} = \gamma_z$), the demand system is equivalent to the demand system generated by Dixit-Stiglitz preferences, which are standard for modeling demand for differentiated goods in international trade.

Exporter i produces N_{iz} symmetric varieties in sector z. Denote by s_{ijz} the share defined in equation (3) when applied to a typical variety from country i in sector z. Using this notation, denote by $\omega_{ijz} = N_{iz}s_{ijz}$ country i's total share in country j's sectoral expenditure. The amount of country j's imports from country i in sector z is then:

$$imports_{ijz} = \omega_{ijz} E_{jz}, \qquad \omega_{ijz} = N_{iz} \left(\frac{\frac{p_{iz}\tau_{ijz}}{(\theta_{iz})^{\gamma_{jz}}}}{G_{jz}} \right)^{1-\sigma_z}, \tag{5}$$

where the import price p_{ijz} is expressed as the product of the export price, p_{iz}, and the (iceberg) trade cost factor between countries i and j, τ_{ijz}. Equation (5) shows that the import share of country i decreases with the export price p_{iz} and increases with the quality of its varieties θ_{iz}.

To gain intuition about the role that cross-country differences in preference for quality play on the direction of trade, compare exporters i and i' shares in countries j and j''s imports in sector z, respectively, by focusing on the double ratio (or double difference if logarithms are taken on both sides):

$$r_{ii'jj'z} = \frac{\omega_{ijz}/\omega_{i'jz}}{\omega_{ij'z}/\omega_{i'j'z}}. \tag{6}$$

This ratio equals one when exporters i and i''s shares in country j's imports are proportional to their shares in country j''s imports, i.e. numerator and denominator are equal. The ratio is greater than one when exporter i's share relative to that of exporter i' is higher in country j 's imports than it is in country j''s imports. Abstracting from trade costs ($\frac{\tau_{ijz}/\tau_{i'jz}}{\tau_{ij'z}/\tau_{i'j'z}} = 1$) and using the definitions of ω_{ijz} and s_{ijz}, we can express this ratio as

$$r_{ii'jj'z} = \left[\frac{(\theta_{iz}/\theta_{i'z})^{\gamma_{jz}}}{(\theta_{iz}/\theta_{i'z})^{\gamma_{j'z}}} \right]^{\sigma_z - 1}. \tag{7}$$

Equation (7) shows that the existence of quality differences between countries i and i' is not sufficient to induce more intense trade between particular country pairs. In the benchmark case with no cross-country variation in γ_{jz}, numerator and denominator of equation (7) are equal even in the presence of quality differences ($\theta_{iz} \neq \theta_{i'z}$). Therefore, $r_{ii'jj'z} = 1$, implying that importers j and j' buy goods from exporters i and i' in the same proportions.[13] In contrast, when there is variation in γ_{jz}, quality differences do influence the direction of trade. For example, if exporter i produces goods of higher quality than exporter i' while importer j has a more intense preference for quality than importer j', then $r_{ii'jj'z} > 1$. In this case, the country with higher γ_{jz} (country j) imports relatively more from the country that produces higher quality goods (country i). Even though both importers prefer high quality rather than low quality goods, this preference is stronger for importers with higher γ.

Equation (5) is the basis of the empirical specification that will be used for estimation. I assume that trade costs are determined by $\tau_{ijz} = (Dist_{ij})^{\eta_z} e^{\boldsymbol{\zeta}_z \mathbf{I}_{ij}} e^{-\upsilon_{ijz}}$, where $\mathbf{I}_{ij}$ is the vector of dummy variables defined in section 2, and υ_{ijz} is a random disturbance to bilateral trade costs. Substituting τ_{ijz} for this expression in (5) and taking logarithms, we obtain

$$\ln imports_{ijz} = \varphi_{iz} + \psi_{jz} - \widetilde{\sigma}_z \eta_z \ln Dist_{ij} + \widetilde{\sigma}_z \boldsymbol{\zeta}_z \mathbf{I}_{ij} + \widetilde{\sigma}_z \gamma_{jz} \ln \theta_{iz} + \widetilde{\sigma}_z \upsilon_{ijz} \tag{8}$$

where φ_{iz} and ψ_{jz} are, respectively, exporter and importer fixed effects, and $\widetilde{\sigma}_z = \sigma_z - 1$ re-scales the elasticity of substitution for notational convenience.

The exporter fixed effect φ_{iz} controls for all determinants of country i's exports that are common across importers. In particular, it captures the impact of export price (p_{iz}) and number of varieties (N_{iz}) on this country's exports to any other country. The export price p_{iz} might reflect movements in the real exchange rate, but it primarily reflects (between-sector) comparative advantage in sector z, i.e. the ability of a country to produce goods at low cost, conditional on their quality.[14] The

[13] This prediction is also implied by standard Dixit-Stiglitz preferences. To see this, define $\widetilde{\theta}_{iz} = \theta_{iz}{}^{\gamma_z}$ as a preference parameter for goods from country i, which is common across importers.

[14] It is worth distinguishing here the notions of within-sector and between-sector comparative advantage. Within

number of varieties N_{iz} primarily reflects exporter size, but it also reflects comparative advantage in the sector, as sectors with lower export prices attract more resources relative to the size of the economy [Romalis (2004), Bernard, Redding and Schott (2005)]. The importer fixed effect ψ_{jz} controls for all determinants of country j's imports that are common across exporters. In particular, it captures the impact of country j's price index (G_{jz}) and sectoral expenditure (E_{jz}) on its imports from any other country. The importer price index G_{jz} captures inbound trade costs between importer j and every exporter, as well as comparative advantage. For example, a low price p_{jz} is given a disproportionate weight in G_{jz}, as domestic varieties do not pay trade costs. Sector expenditure E_{jz}, in turn, is primarily driven by size but can also capture differences in sector expenditure shares across countries. In sum, the exporter and importer fixed effects in equation (8) control for determinants of trade that operate *at the sector level*. Those controls are crucial for identifying the Linder effect.

The systematic relationship between per-capita income and demand for quality predicted by the theory can be formalized in this framework by postulating γ_{jz} as a function of importer per-capita income,

$$\gamma_{jz} = \gamma_z + \mu_z \ln y_j. \tag{9}$$

Equation (9) does not impose any parameter restriction on μ_z. For example, it allows for a benchmark case in which there are no cross-country differences in demand for quality ($\mu_z = 0$). in which case the distribution of import shares is common across importers – up to differences due to trade costs. However, the demand side of Linder's theory asserts that rich countries consume higher

sectors, countries have conventional comparative advantage in either low quality or high quality varieties. Either case is compatible with a between-sector comparative advantage or disadvantage in the sector as a whole. The country will have a between-sector comparative advantage if the cost of the goods it produces, relative to their quality, is sufficiently low to make it a net exporter in the sector. This will happen when a majority of world consumers find the price/quality relation offered by the country more appealing than the price/quality relation offered by other countries. For example, high income countries can be expected to have a within-sector comparative advantage in high quality apparel. However, the price-quality relation of their high quality varieties is presumably less appealing, for a majority of world consumers, than the price-quality relation of the low quality varieties offered by low income countries. The former are therefore net importers and the latter are net exporters in this sector. In Machinery, the same pattern of within-sector comparative advantage is consistent with the opposite pattern of between-sector comparative advantage, making high income countries net exporters and low income countries net importers.

quality goods. This prediction implies that $\mu_z > 0$.[15]

Combining equations (8) and (9) we obtain

$$\ln imports_{ijz} = \varphi_{iz} + \psi_{jz} - \widetilde{\sigma}_z \eta_z \ln Dist_{ij} + \widetilde{\sigma}_z \boldsymbol{\zeta}_z \mathbf{I}_{ij} + \widetilde{\sigma}_z \mu_z \ln \theta_{iz} \ln y_j + \widetilde{\sigma}_z \upsilon_{ijz}. \tag{10}$$

Since product quality (θ_{iz}) is not observable, equation (10) cannot be estimated as such. Hallak (2006) uses export unit values as indicators of product quality, which implies focusing only on the demand side of Linder's theory. However, since the Linder hypothesis is based on the interaction of demand and supply forces, a formal characterization of Linder's theory requires relating quality supply to income per capita.

3.2 Supply of quality

Linder argued that high income countries have a comparative advantage in the production of high quality goods. Several theories can explain a systematic relationship between per-capita income and quality production. For example, a Ricardian view of quality specialization predicts that richer countries will export high quality goods if their productivity advantage in those goods is relatively larger. Alternatively, a Heckscher-Ohlin view of quality specialization predicts that rich countries – which tend to be capital abundant – will export high quality goods if those goods are capital intensive. Linder, in turn, proposed the idea that closeness to demand for high quality products, which occurs disproportionately in high income countries, provides them with a comparative advantage in the production of those goods. Finally, Vernon's (1966) product cycle theory[16] can also explain high income countries' comparative advantage in high quality production if most innovations consist of quality upgrading of old varieties.[17] The Linder hypothesis' prediction on the direction of trade can hold under any of these theories about the supply-side determinants of quality production. Their

[15] Aggregate consumption of high quality goods should also depend on the distribution of income (see Choi et al. for recent empirical evidence). The connection between income distribution and the Linder hypothesis is explored empirically by Francois and Kaplan (1996) and by Dalgin et al. (2005) which, however, do not address the role of product quality.

[16] See also Flam and Helpman (1987), Stokey (1991) and Antràs (2005).

[17] The relationship between quality and income per capita could also be founded on subjective grounds. Evidence from the marketing literature finds that consumers take the origin country's development level as an extrinsic cue of product quality when quality is imperfectly observable. For a review of this literature, see Bilkey and Nes (1982) and Verlegh and Steenkamp (1999).

common implications are captured here by simply postulating a systematic relationship between quality supply and per-capita income,

$$\ln \theta_{iz} = \vartheta_z + \delta_z \ln y_i + \varsigma_{iz} \tag{11}$$

where ς_{iz} is a stochastic disturbance. The supply side of Linder's theory predicts that $\delta_z > 0$.

3.3 Demand and supply interaction: The sectoral Linder hypothesis

Substituting equation (11) into (10), we obtain:

$$\ln imports_{ijz} = \varphi_{iz} + \psi_{jz} - \beta_{Dz} \ln Dist_{ij} + \beta_{Iz} \mathbf{I}_{ij} + \beta_{yz} \ln y_i \ln y_j + \varepsilon_{ijz} \tag{12}$$

where $\beta_{Dz} = \widetilde{\sigma}_z \eta_z$, $\beta_{Iz} = \widetilde{\sigma}_z \boldsymbol{\zeta}_z, \beta_{yz} = \widetilde{\sigma}_z \mu_z \delta_z$, and $\varepsilon_{ijz} = \widetilde{\sigma}_z \mu_z \varsigma_{iz} + \widetilde{\sigma}_z v_{ijz}$.[18] The disturbances ς_{iz} and v_{ijz} are assumed to be uncorrelated with the regressors; therefore, so is ε_{ijz}.

The parameter of interest is the coefficient on the term interacting the trading partners' per-capita incomes, $\beta_{yz} = \widetilde{\sigma}_z \mu_z \delta_z$. Since $\widetilde{\sigma}_z > 0$, the sign of β_{yz} corresponds to the sign of $\mu_z \delta_z$. Linder's theory postulates that richer countries consume higher quality goods ($\mu_z > 0$) and also produce higher quality goods ($\delta_z > 0$). The interaction of these forces then implies that $\beta_{yz} > 0$. This prediction is in fact equivalent to the prediction of the Linder hypothesis, as typically specified in empirical exercises. The equivalence is easy to show; basic algebraic manipulation of the term $Linder_{ijz} = (\ln y_i - \ln y_j)^2$ yields

$$\ln y_i \ln y_j = \frac{1}{2} (\ln y_i)^2 + \frac{1}{2} (\ln y_j)^2 - \frac{1}{2} (\ln y_i - \ln y_j)^2 . \tag{13}$$

This expression, substituted back into equation (12), results in

$$\ln imports_{ijz} = \varphi_{iz} + \psi_{jz} - \beta_{Dz} \ln Dist_{ij} + \beta_{Iz} \mathbf{I}_{ij} + \beta_{Lz} (\ln y_i - \ln y_j)^2 + \varepsilon_{ijz} \tag{14}$$

where the exporter and importer fixed effects absorb, respectively, the first two terms of (13), and $\beta_{Lz} = -(1/2)\beta_{yz}$. The last equality implies the equivalence between the prediction that $\beta_{yz} > 0$ in equation (12) and the prediction that $\beta_{Lz} < 0$ in equation (14).

These results demonstrate that the Linder hypothesis can be derived from a theoretical framework that captures the supply and demand forces originally described by Linder. They also show

[18] The importer fixed effects absorb the interaction of the constant ϑ_z with $\ln y_j$.

that a gravity-equation specification, augmented with a Linder term, is appropriate for testing the role of quality described by the theory when it is formulated at the sector level. I henceforth refer to such formulation as the "sectoral Linder hypothesis".

4 The "sectoral" Linder hypothesis: Empirical results

4.1 Data

Estimating equation (14) requires bilateral trade flows at the sector level, which I obtain from the World Trade Flows data set (Feenstra 2000). This dataset breaks down trade flows up to the 4-digit SITC (Rev.2) level of aggregation. However, I define sectors at the 3-digit SITC level because information at the 4-digit level is often missing.

I follow Rauch's (1999) classification of 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 that do not satisfy either of the two previous criteria. I use Rauch's "liberal" classification 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 categories, the 3-digit sector is broken down accordingly, each part including only the relevant 4-digit sectors.

The criterion for selecting the countries in the sample attempts to balance two considerations. On the one hand, including more countries increases sample size and estimation precision. On the other hand, concentrating on relatively large countries decreases the proportion of bilateral country pairs with zero trade at the sector level, which prevents zero-trade observations from dominating the sample. The sample includes 64 countries, all of those with a population larger than 3 million, and with more than 2 billion-dollar imports of Differentiated goods. I also drop very small sectors, keeping only sectors with a volume of trade (among the 64 selected countries) above 2 billion dollars. The final sample includes 116 Differentiated sectors, 56 Reference-priced sectors, and 39 Homogeneous sectors. They are listed in Table A2.

The variable *distance* measures great circle distance between capital cities and comes from Shatz (1997). Dummies for border, common language, colonizer-colony relationship, and common-colonizer relationship were constructed using the CIA Factbook. Only "official" languages are taken into account in the construction of the common language variable, except for Malaysia-Singapore, which is coded here as sharing 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, and US-Israel. Data on PPP GDP come from the World Bank WDI.

4.2 Estimation issues

Equation (14) predicts bilateral trade at the sector level. Aggregation implicitly forces the parameters of this equation to be equal across sectors, which in particular precludes the use of sector-specific exporter and importer fixed effects that control for inter-sectoral determinants of trade. I thus estimate equation (14) by sector. I also estimate pooling the observations across sectors but allowing for cross-sector variation in all the parameters and fixed effects except for the "Linder" coefficient ($\beta_{Lz} = \beta_L$). While in the first case the estimation yields sector-by-sector estimates of the parameter of interest, in the second case the estimation yields a single estimate of this parameter.

I focus on Differentiated goods because those are the goods for which the assumptions of the theory most clearly apply. Equation (14) can also be estimated using intermediate-good sectors if we interpret equation (2) as a production function of a final good. The interpretation of the sectoral Linder hypothesis in that case is that richer countries consume a larger proportion of high-quality intermediate inputs, presumably as a requirement to produce high-quality final goods.

Estimation of gravity-type equations using OLS are known to suffer from a potential selection bias, as the OLS procedure drops bilateral pairs with zero trade. To deal with this problem, I use a generalized Tobit estimation with random and unobserved censoring value. The estimation strategy is based on the idea that zero-values of trade are due to the presence of fixed costs of exporting. The magnitude of fixed costs is unobservable but can be modeled as a function of observable variables. In particular, I postulate a censoring equation of the following form:

$$\ln c_{ijz} = \delta_{0z} + \delta_{dz} \ln Dist_{ij} + \boldsymbol{\delta}_{Iz} \mathbf{I}_{ij} + \delta_{xz} \ln Y_i + \delta_{mz} \ln Y_j + u_{ijz} \tag{15}$$

where the (unobserved) censoring value is determined by the level of trade that generates sufficient profits to cover the fixed costs, F_{ijz}. Given the constant elasticity of demand, the censoring value is proportional to those costs ($c_{ijz} = \sigma_z F_{ijz}$). The vector $\mathbf{I}_{ij}$ includes the same dummy variables as in (14), and u_{ijz} is a normally distributed random disturbance. Joint estimation of the parameters of equations (14) and (15), respectively the "Imports equation" and the "Fixed-cost equation", can be performed using maximum likelihood (ML) estimation under the assumption that the distribution of the two errors is bivariate normal. This estimation strategy is explained in more detail in Hallak (2006).[19]

4.3 Estimation results

Equation (14) is first estimated by OLS separately for each of the 116 Differentiated sectors. Panel A of Table 4 provides a summary report of the estimated coefficients, by sign (columns 1 and 2) and by significance level (columns 3 to 5). Standard errors are robust to heteroskedasticity in all cases. The results support the empirical validity of the "sectoral Linder hypothesis". The estimated coefficient on the Linder term is negative, as predicted, in more than 2/3 of the sectors (82), while it is positive in less than 1/3 of the sectors (34). The coefficient is negative and significant at the 5% level in approximately one half of the sectors (59) and it is positive and significant in less than 1/5 of the sectors (21). The variables that control for trade costs affect trade volumes in the predicted direction; shorter distance promotes trade, as does sharing a common border, common language, PTA, colonial relationship, or common colonizer. Column 6 shows the median magnitude of the estimated coefficients. The median has the predicted sign for all variables.

The last column of the table reports the parameter estimate and standard error for the coefficient on the Linder term when all the observations are pooled (but not aggregated) across sectors and the coefficient is constrained to be the same across sectors.[20] The magnitude of the estimated coefficient is substantially smaller than the median value of the sectoral estimates.[21] Nevertheless, the effect of the Linder term is still negative and it is significantly different from zero at the 1% level.

[19] Two more recent methodologies for dealing with zero-trade observations are proposed by Silva and Tenreyro (2005) and by Helpman, Melitz and Rubinstein (2005).

[20] The coefficients on all other variables are not constrained to be equal across sectors. Summary measures of those estimates are not reported to save space.

[21] All reported standard errors in pooled regressions are calculated allowing for clustering by country pair.

The term $\exp\left[\widehat{\beta}_{Lz}(\ln y_i - \ln y_j)^2\right]$ provides a measure of the impact of the Linder effect on the volume of bilateral trade. We can calculate the magnitude of this effect using the estimate $\widehat{\beta}_{Lz} = -0.0435$ obtained in the pooled regression. For example, as a consequence of income per-capita similarity, the U.S. is predicted to import from Switzerland 10% more than it imports from Colombia, 20% more than it imports from the Philippines, 40% more than it imports from Pakistan, and 70% more than it imports from Nigeria.

Panel B of Table 4 provides the results of the censoring model using ML estimation.[22] The censoring model yields qualitatively similar results, even though there is now a smaller number of sectors with negative and with negative and significant estimates of the Linder term. The median coefficient on this term is also substantially smaller (in absolute value), and is close in magnitude to the pooled-sample OLS estimate. The estimates for the other controls in the Imports Equation have the expected sign in most sectors. They also have the expected sign in the Fixed Cost Equation, except for common border and importer GDP, which are estimated to have a positive effect on the fixed costs of exporting.

Even though the theoretical framework of this paper suggests the use of a particular Linder term, alternative Linder terms can be used to assess the robustness of the results. Table 5 shows the results from estimating (14) with different Linder terms.[23] As was the case with the aggregate data, the results are not very sensitive to the choice of Linder term. Regardless of which one is used, the estimates support the sectoral Linder hypothesis: controlling for sectoral determinants of trade, countries with similar income per capita tend to trade more with one another.

The fact that estimated coefficients are positive and significant in a number of sectors (21 and 27 sectors when OLS and the censoring model are respectively used to estimate) raises the concern that forces correlated with the interaction of the country pair's per-capita incomes, but other than those highlighted by Linder's theory, might have a substantial influence on the results. This concern can be addressed by testing the sectoral Linder hypothesis using Reference-priced and Homogeneous sectors. Since the assumptions of the theory match the characteristics of those sectors in varying degrees, the results can be used to assess whether it is the Linder mechanism that drives the results for Differentiated goods.

[22] Estimating the censoring model on the pooled data is unfeasible due to the dimentionality of the computational problem. The pooled ML estimation requires over 10,000 sector-specific exporter and importer fixed effects.

[23] To conserve space, all subsequent tables only report results on the Linder term.

The sectoral Linder hypothesis is derived under the joint assumption that both quality supply and quality demand are systematically related to income per capita. On the supply side, despite the fact that quality differences often exist even among goods that are not classified as Differentiated, such differences are likely to be smaller for Reference-priced sectors and yet even smaller for Homogeneous sectors. Further, even if quality differences exist in those sectors, quality supply is less likely to be correlated with income per capita, mostly so in the case of Homogeneous sectors. It is reasonable to think, for example, that among manufactured goods (typically classified as Differentiated or Reference-priced) specialization of a country in the high quality end results from a high proportion of (human or physical) capital to labor. High income countries, which are usually abundant in capital, then produce high-quality capital-intensive manufactures. In contrast, among agricultural and mineral commodities (typically classified as Homogeneous), high quality production often results from a country's abundance of high quality natural resources, which is not necessarily related to the country's per-capita income. This suggests that the supply-side assumptions of the sectoral Linder hypothesis should weaken as we move from Differentiated to Reference-priced goods, and should weaken considerably more as we move from the latter goods to Homogeneous goods.

On the demand side, horizontal differentiation – as assumed in the theory – is a more appropriate characterization of Differentiated sectors than it is of Homogeneous sectors. To the extent that quality differences exist in the latter sectors, richer countries might still consume relatively more high quality products, and thus import relatively more from the countries that produce them. However, as there is no justification for a "love for variety" assumption in this case, they will be expected to import, among countries producing high quality, only from those with the lowest price (net of trade costs). Therefore, the bilateral predictions of equation (14) will only capture an average effect.

The combination of a weak relationship between per-capita income and quality supply and a weak relationship between per-capita income and quality demand implies that the theoretical results of the paper should not apply to Homogeneous sectors. In the case of Reference-priced sectors, it is not *a priori* obvious whether the supply and demand relationships predicted by the theory are sufficiently strong to induce bilateral patterns of trade consistent with the Linder hypothesis, leaving its validity as an open empirical question.

Table 6 compares the distribution of coefficient estimates for the baseline Linder term between the three groups of sectors (results for Differentiated goods reproduce those previously reported). Even though the theoretical prediction is ambiguous for Referenced-priced sectors, the results for those sectors are still consistent with the predictions of the sectoral Linder hypothesis. When OLS is used to estimate, the coefficient is negative in 4/5 of the sectors (43 out of 56) and is negative and significant in more than 1/3 of the sectors (21). When the censoring model is used to estimate, the coefficient is negative in 2/3 of the sectors (37) and is negative and significant in less than half of the sectors (24). Compared to the results for Differentiated sectors, a larger proportion of Reference-priced sectors display no significant Linder effect. The discrepancy between the OLS and ML results is not substantial here; for example, the median estimate is very similar in both cases. Finally, the estimation using the pooled sample yields a negative and significant coefficient estimate, close in magnitude to the estimate for Differentiated goods.

The results are radically different in the case of Homogeneous goods. The table shows that the estimated coefficient is more often positive than negative, more often positive and significant than negative and significant, and not significant for most sectors. In addition, both the median coefficient and the pooled coefficient are positive (in the latter case, significant at the 10% level). The failure to find Linder effects for Homogeneous sectors is consistent with the fact that the characteristics of those sectors violate the assumptions of the theory. Therefore, these results provide evidence that it is the interaction of supply and demand for quality stressed by the theory that drives the estimates of the Linder effect for the other categories of goods and not other unobserved factors related to per-capita income. The estimates are consistent with the Linder hypothesis only in those categories in which we expect quality supply and quality demand to be correlated with income per capita.

5 Aggregation Bias

This section shows that standard estimates of the Linder hypothesis, which use the gravity-equation framework at the aggregate level, suffer from a systematic bias. The bias explains the literature's failure to find empirical support for the hypothesis. The first part of the section estimates equation (14) at different levels of aggregation and shows that the estimated Linder coefficient increases with the aggregation level. The increase is sufficiently strong to reverse the sign of the coefficient at high levels of aggregation. The second part of the section shows that failure to control for inter-sectoral determinants of trade at aggregate levels generates a bias that can account for the reversal of the estimated Linder effect. The third part provides further evidence on the nature of the bias by comparing estimation results at different levels of aggregation in Differentiated, Reference-priced, and Homogeneous sectors.

5.1 Empirical evidence on the aggregation bias

Table 7 displays the results of estimating the Linder hypothesis using Differentiated sectors defined alternatively at the 3-digit level (116 sectors), 2-digit level (36 sectors), and 1-digit level (6 sectors). The table also reports estimates of the Linder hypothesis using aggregate trade flows for all Differentiated sectors. When trade flows are aggregated at the 2-digit level, the results are very similar to those obtained at the 3-digit level (reproduced in the first row), both in terms of magnitude and distribution of estimates across significance levels. However, when trade flows are aggregated at the 1-digit level, the results no longer support the Linder hypothesis. In particular, both the median of the sectoral estimates and the estimate of the pooled regression reverse signs, which are now estimated to be positive. Finally, when all trade in Differentiated sectors is aggregated into only one sector, the reversal of the estimated Linder effect is even starker. In that case, the Linder coefficient is positive and significant at the 1% level. The ML estimates in panel B show a similar pattern: the coefficient on the Linder term becomes more positive as we aggregate, with the sign reversing (significantly at the 1% level) when the estimation is performed aggregating all Differentiated sectors into one category.

5.2 The nature of the bias

To describe the bias, we will focus on the double ratio $r_{ii'jj'z}$ defined in equation (6) (countries i and i' are exporters and countries j and j' are importers). This double ratio can be obtained by double-differencing the estimating equation (14):

$$\begin{aligned} \ln r_{ii'jj'z} &= \ln imports_{ijz} - \ln imports_{i'jz} - \ln imports_{ij'z} + \ln imports_{i'j'z} \\ &= \beta_{Dz}\widetilde{d}_{ii'jj'} + \beta_{Iz}\widetilde{I}_{ii'jj'} + -2\beta_{Lz}\left(\ln y_i - \ln y_{i'}\right)\left(\ln y_j - \ln y_{j'}\right) + \widetilde{\varepsilon}_{ii'jj'z} \end{aligned} \quad (16)$$

where $\widetilde{d}_{ii'jj'} = \ln dist_{ij} - \ln dist_{i'j} - \ln dist_{ij'} + \ln dist_{i'j'}$, $\widetilde{I}_{ii'jj'} = I_{ij} - I_{i'j} - I_{ij'} + I_{i'j'}$, and $\widetilde{\varepsilon}_{ii'jj'z} = \varepsilon_{ijz} - \varepsilon_{i'jz} - \varepsilon_{ij'z} + \varepsilon_{i'j'z}$. Equation (16) is isomorphic to the baseline specification (14) with respect to β_{Lz}. Therefore, we will use this equation to describe the nature of the bias since it allows for a simpler exposition.

Consider two sectors: $z = 1, 2$. To describe the aggregation bias more transparently, we focus on a case in which the Linder effect is identical across sectors ($\beta_{L1} = \beta_{L2} = \beta_L$) and the effect of trade costs cancels out in the double difference ($\widetilde{d}_{ii'jj'} = \widetilde{I}_{ii'jj'} = 0$). Averaging equation (16) across the two sectors, we obtain

$$\frac{1}{2}\left(\ln r_{ii'jj'1} + \ln r_{ii'jj'2}\right) = -2\beta_L \left(\ln y_i - \ln y_{i'}\right)\left(\ln y_j - \ln y_{j'}\right) + \frac{1}{2}\left(\widetilde{\varepsilon}_{ii'jj'1} + \widetilde{\varepsilon}_{ii'jj'2}\right). \tag{17}$$

Define $r^*_{ii'jj'A} = \sqrt{r_{ii'jj'1} r_{ii'jj'2}}$, $\widetilde{y}_{ii'jj'} = \left(\ln y_i - \ln y_{i'}\right)\left(\ln y_j - \ln y_{j'}\right)$, and $\widetilde{\varepsilon}_{ii'jj'A} = \frac{1}{2}\left(\widetilde{\varepsilon}_{ii'jj'1} + \widetilde{\varepsilon}_{ii'jj'2}\right)$. Then, equation (17) can be written as

$$\ln r^*_{ii'jj'A} = -2\beta_L \widetilde{y}_{ii'jj'} + \widetilde{\varepsilon}_{ii'jj'A}. \tag{18}$$

If the OLS estimator of β_L in the baseline specification (14) is consistent, then so is the OLS estimator of β_L in (18). Without loss of generality, assume that $y_i > y_{i'}$ and $y_j > y_{j'}$. Then, the term $\widetilde{y}_{ii'jj'}$ is larger the larger are per-capita income differences between countries i and i' and countries j and j'. The ratios $r_{ii'jj'1}$ and $r_{ii'jj'2}$ are later interpreted as odds ratios. By construction, the "aggregate" odds ratio $r^*_{ii'jj'A}$ is bounded between the sector-level odds ratios $r_{ii'jj'1}$ and $r_{ii'jj'2}$.

Standard tests of the Linder hypothesis – as described and performed in section 2 – estimate equation (14) aggregating trade flows across sectors. Double-differencing that equation as is done above, we can interpret such estimates as the result of regressing (18) using a wrongly constructed dependent variable. In that case, the dependent variable is $\ln r_{ii'jj'A}$, where

$$\ln r_{ii'jj'A} = \ln imports_{ijA} - \ln imports_{i'jA} - \ln imports_{ij'A} + \ln imports_{i'j'A}$$

and subindex A indicates that trade flows are summed across sectors. Defining $u_{ii'jj'A} = \ln r_{ii'jj'A} - \ln r^*_{ii'jj'A}$ and substituting for $\ln r^*_{ii'jj'A}$ in (18), the standard specification for testing the Linder hypothesis can be expressed as

$$\ln r_{ii'jj'A} = -2\beta_L \widetilde{y}_{ii'jj'} + \widetilde{\varepsilon}_{ii'jj'A} + u_{ii'jj'A}. \tag{19}$$

Under conditions described below, $u_{ii'jj'A} < 0$, i.e. the aggregate odds ratio $r_{ii'jj'A}$ is systematically lower than $r^*_{ii'jj'A}$. Furthermore, $u_{ii'jj'A}$ is negatively correlated with $\widetilde{y}_{ii'jj'}$, inducing an upward bias in the Linder coefficient β_L. The form in which aggregation induces this systematic bias is illustrated next with a simple example. The Appendix demonstrates it in a more general setting.

Consider trade flows between two exporters, the U.S. and China, and two importers, Switzerland and South Africa, in two different 2-digit sectors, Industrial Machinery (SITC 72: Machinery Specialized for Particular Industries) and Apparel (SITC 84: Articles of Apparel and Clothing Accesories). Since U.S. per-capita income is higher than China's and Switzerland per-capita income is higher than South Africa's, U.S. quality is presumably higher in both sectors, as is Switzerland's intensity of preference for quality.

Reported trade flows between exporters U.S. (i) and China (i') and importers Switzerland (j) and South Africa (j') in each of the two sectors (in 1995) are displayed in the first set of rows of Table 8. The three vertical panels of this table can be thought of as frequency tables between an exporter variable $I = i,\ i'$ and an importer variable $J = j,\ j'$. The second set of rows reports exporter shares in import markets, where entry ij is the probability that \$1 of j's imports comes from i. The last row reports the double ratio, which in this context is an odds ratio. The odds ratio is higher than one in both sectors – 1.44 in Machinery and 3.32 in Apparel – indicating that Switzerland is more likely to import from the U.S. than is South Africa, as predicted by the theory. However, the last set of columns shows that, when we aggregate trade flows across the two sectors, the odds ratio is 0.29. In contrast to $r^*_{ii'jj'A}$, which is inside the bounds defined by the sectoral odds ratios, $r_{ii'jj'A}$ is not only outside those bounds but is also below 1, indicating a reversal in the direction of association between the exporter and importer variables. In the aggregate, Switzerland is less likely to import from the U.S. than is South Africa. This result is a typical example of Simpson's Paradox, a well-known case of association reversal due to aggregation [Simpson (1951), Samuels (1993)]. In this case, the paradox is driven by the fact that Switzerland (S. Africa) and the U.S. (China), the countries with similar income per capita, also have similar sectoral trade patterns. Therefore, Switzerland imports relatively more from China in the aggregate because it is a relatively large importer in the sector in which China is a large exporter (Apparel). The Appendix shows formally that the $r_{ii'jj'A} < r^*_{ii'jj'A}$ aggregation bias occurs whenever countries' patterns of inter-sectoral trade are related to per-capita income. It also shows that the bias increases with the strength of those patterns. As a result, $u_{ii'jj'A}$ is positively correlated with $\widetilde{y}_{ii'jj'}$, inducing an upward bias in $\widehat{\beta}_L$ obtained from estimating (19). This bias can explain the failure to find consistent support for the (aggregate) Linder hypothesis.

Empirical estimates at the 3-digit level and at the 2-digit level are very similar – as shown in Table 7 – because determinants of trade patterns across sectors are often similar for 3-digit sectors in the same 2-digit category. For example, countries that are net exporters of Apparel (SITC 84) are often net exporters of both Outer Garments, Mens, of Textile Fabrics (SITC 842) and of Outer Garments, Womens, of Textile Fabrics (SITC 843). When we aggregate at the 2-digit level, we are unable to control for determinants of trade at the 3-digit level, as only exporter and importer fixed effects at the 2-digit level can be included in the estimation. Despite this limitation, the similarity of trade determinants among 3-digit sectors in the same 2-digit category implies that fixed effects at the 2-digit level can still reasonably control for trade determinants at the 3-digit level, without inducing substantial bias.[24] This is not the case when we aggregate at higher levels, as determinants of trade can be drastically different across 2-digit sectors. In particular, when we aggregate all trade in Differentiated goods, the exporter and importer fixed effects that we include at such a high level of aggregation are unable to control for the wide heterogeneity in patterns of sectoral trade across the 2-digit sectors included in the aggregate.

5.3 The aggregation bias across different types of goods

A condition for the existence of systematic aggregation bias is that sectoral trade patterns are systematically related to countries' per-capita income. This is often the case in Differentiated sectors. For example, high income countries tend to be net exporters of Machinery and net importers of Apparel while low income countries tend to be net exporters of Apparel and net importers of Machinery).[25] The reversal of the estimated Linder coefficient when we aggregate Differentiated sectors suggests that inter-sectoral determinants of trade prevail over the intra-sectoral forces emphasized by Linder. As a result, aggregate trade flows in Differentiated products are more intensive

[24] A similar argument addresses concerns related to the fact that even trade flows at the 3-digit level aggregate more narrowly defined subsectors.

[25] A possible explanation for this pattern is that comparative advantage in those sectors might stem from the relative abundance of accumulable factors (e.g. physical and human capital), which is observed to be strongly correlated with countries' per-capita income.

between countries of *dissimilar* income per capita.

The systematic relationship between per-capita income and sectoral trade patterns is probably less pronounced for Reference-priced sectors. In those sectors, patterns of trade are sometimes related to the availability of natural resources (e.g. countries with abundant forests tend to be net exporters of paper products), in which case the correlation of these patterns with per-capita income are not as strong. Therefore, a weaker aggregation bias should be expected for those goods. This prediction is consistent with the results displayed in Table 9; even though there is a substantial aggregation bias, the Linder coefficient in this case is not significantly different from zero. In the case of Homogeneous sectors, aggregation induces a reversal of the estimated coefficient but in the opposite direction. However, this reversal is difficult to interpret in light of the fact that, as argued before, even the regressions at the sector level are not properly specified for Homogeneous goods.

Before concluding, it is interesting to note that Linder explicitly made the point that the connection between income and quality demand should operate within sectors instead of across sectors, even though he later used aggregate trade information to provide preliminary evidence in support of his theory. In particular, he argued that (Linder 1961, p. 95): "Qualitative product differences are not well brought out in empirical studies of consumer behavior along the lines first followed by Engel. The qualitative factor is submerged by taking broad groups of goods such as "food" or "clothing"." His own work and the work of subsequent empirical researchers did not follow this early lead.

6 Conclusions

Despite the persistent appeal of the Linder hypothesis, a large body of empirical work testing its validity has failed to find consistent support for it. This failure is the result of the unanimous use of a misspecified empirical benchmark, the gravity equation estimated using aggregate data. Building a theoretical framework that captures the main aspects of Linder's theory, this paper shows that the Linder hypothesis should be formulated at the sector level, where inter-sectoral determinants of trade can be controlled for. The "sectoral Linder hypothesis" is tested and confirmed empirically. Further, it is shown that aggregation across sectors induces a systematic bias against finding support for the Linder effect.

References

[1] Anderson, J. and E. Van Wincoop (2003), "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170-192.

[2] Antràs, P. (2005), "Incomplete Contracts and the Product Cycle," *American Economic Review*, 95 (3), 1054-1073.

[3] Bergstrand, J. (1989), "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory of International Trade," *The Review of Economics and Statistics*, 71 (1), 143-153.

[4] Bernard, A., S. Redding and P. Schott (2005), "Heterogenous Firms and Comparative Advantage," mimeo.

[5] Bilkey, W. and E. Nes (1982), "Country-Of-Origin Effects on Product Evaluations," *Journal of International Business Studies*, Spring/Summer, 89-100.

[6] Bils, M. and P. Klenow (2001), "Quantifying Quality Growth," *American Economic Review*, 91, 1006-1030.

[7] Brooks, E. (2005), "Why Don't Firms Export More? Product Quality and Colombian Plants," Forthcoming *Journal of Development Economics*.

[8] Chen, M., T. Otsuki, and J. Wilson (2006), "Do Standards Matter for Export Success?" World Bank Policy Research Working Paper 3809.

[9] Choi, Y., D. Hummels, and C. Xiang (2006), "Explaining Import Variety and Quality: the Role of Income Distribution, Purdue University, mimeo.

[10] Dalgin, M., Mitra, D., and V. Trindade (2005), "Inequality, Nonhomothetic Preferences, and Trade: A Gravity Approach," mimeo.

[11] Deardorff, A. (1984), "Testing Trade Theories and Predicting Trade Flows," in R.W.Jones and P.B.Kenen, ed., *Handbook of International Economics*, Vol.1, (Amsterdam, Elsevier).

[12] Deardorff, A. (1998), "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?," in J. Frankel, ed., *The Regionalization of the World Economy,* (Chicago, Chicago University Press).

[13] Eaton, J. and S. Kortum (2002), "Technology, Geography, and Trade," *Econometrica*, 70, 1741-1779.

[14] Falvey, R. and H. Kierzkowski (1987), "Product Quality, Intra-industry Trade and (Im)perfect Competition," in H. Kierzkowski, ed., *Protection and Competition in International Trade*, (Oxford: Basil Blackwell).

[15] Feenstra, R. (2000), "World Trade Flows, 1980-1997," Center for International Data, Institute of Governmental Affairs, UC Davis.

[16] Flam, H., and E. Helpman (1987), "Vertical Product Differentiation and North-South Trade," *American Economic Review*, 77 (5), 810-822.

[17] Francois, J. and S. Kaplan (1996), "Aggregate Demand Shifts, Income Distribution, and the Linder Hypothesis," *Review of Economics and Statistics*, 78, 244-250.

[18] Hallak, J. (2006), "Product Quality and the Direction of Trade," *Journal of International Economics*, 68(1), 238-265.

[19] Helpman, H., Melitz, M., Rubinstein, Y., (2005), "Trading Partners and Trading Volumes," mimeo.

[20] Hummels, D. and P. Klenow (2005), "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95: 704-723.

[21] Hunter, L. (1991), "The Contribution of Non-homothetic Preferences to Trade," *Journal of International Economics*, 30, 345-358.

[22] Hunter, L. and J. Markusen (1988), "Per-Capita Income as a Determinant of Trade, in R. Feenstra, ed., *Empirical Methods for International Trade*, (Cambridge, MA, MIT Press).

[23] Leamer, E. and J. Levinsohn (1995), "International Trade Theory: The Evidence," in G. Grossman and K. Rogoff, ed., *Handbook of International Economics*, Vol.3, (Amsterdam, Elsevier).

[24] Linder (1961), S., *An Essay on Trade and Transformation*, (Stockholm: Almqvist & Wiksell).

[25] Markusen, J. (1986), "Explaining the Volume of Trade: An Eclectic Approach," *American Economic Review*, 76 (5), 1002-1011.

[26] Maskus, K. (2005), "The Cost of Compliance with Product Standards for Firms in Developing Countries: An Econometric Study," World Bank Policy Research Working Paper 3590.

[27] 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 (6), 1093-1120.

[28] McPherson, M., M. Redfearn, and M. Tieslau (2001), "International Trade and Developing Countries: An Empirical Investigation of the Linder Hypothesis," *Applied Economics*, 33, 649-657.

[29] Murphy, K.M. and A. Shleifer (1997), "Quality and Trade," *Journal of Development Economics*, 53, 1-15.

[30] Rauch, J. (1999), "Networks versus Markets in International Trade," *Journal of International Economics*, 48, 7-35.

[31] Romalis, J. (2004), "Factor Proportions and the Structure of Commodity Trade", *American Economic Review*, 94, 67-97.

[32] Samuels, M. (1993), "Simpson's Paradox and Related Phenomena," *Journal of the American Statistical Association*, 88, 81-88.

[33] Santos Silva, J. and S. Tenreyro (2005), "The Log of Gravity," Forthcoming *The Review of Economics and Statistics.*

[34] Schott, P. (2004), "Across-Product versus Within-Product Specialization in International Trade," *Quarterly Journal of Economics*, 119 (2), 647-678.

[35] Shatz, H. (1997), "Note on Distance Data", mimeo.

[36] Simpson, E. (1951), "The Interpretation of Interaction in Contingency Tables," *Journal of the Royal Statistical Society Series B*, 13, 238-241.

[37] Stokey, N. (1991), "The Volume and Composition of Trade between Rich and Poor Countries." *Review of Economic Studies*, 58, 63-80.

[38] Thursby, J. and M. Thursby (1987), "Bilateral Trade Flows, The Linder Hypothesis, and Exchange Risk," *The Review of Economics and Statistics*, 69 (3), 488-495.

[39] Verhoogen, E. (2004), "Trade, Quality Upgrading and Wage Inequality in the Mexican Manufacturing Sector: Theory and Evidence from an Exchange Rate Shock," mimeo.

[40] Verlegh, P. and J-B. Steenkamp, "A Review and Meta-Analysis of Country-Of-Origin Research," *Journal of Economic Psychology*, 20, 521-546.

[41] Vernon, R. (1966), "International Investment and International Trade in the Product Cycle," *Quarterly Journal of Economics*, 80 (2), 190-207.

[42] World Bank (1999), *World Development Report: Knowledge for Development*, (Oxford University Press: Oxford).

[43] World Trade Organization (2005), *World Trade Report 2005: Exploring the Links Between Trade, Standards, and the WTO.*

A Appendix

This appendix derives necessary and sufficient conditions for the existence of aggregation bias. Consider a general case of the example in Table 8. Denote by x_{ijz} the exports of country i to country j in sector z. To facilitate interpretation of the results, we focus on the case with no Linder effect, i.e. the case in which $r_{ii'jj'z} = 1$ for $z = 1, 2$. When the odds ratio is 1 in both sectors,

$$\frac{x_{ij1}}{x_{i'j1}} = \frac{x_{ij'1}}{x_{i'j'1}} \Rightarrow x_{ij'1} = a x_{ij1}, x_{i'j'1} = a x_{i'j1}, a > 0 \tag{20}$$

and

$$\frac{x_{ij2}}{x_{i'j2}} = \frac{x_{ij'2}}{x_{i'j'2}} \Rightarrow x_{ij'2} = c x_{ij2}, x_{i'j'2} = c x_{i'j2}, c > 0. \tag{21}$$

Since the odds ratio equals 1 in both sectors, a negative bias will exist if the aggregate odds ratio is lower than 1, i.e. if

$$\frac{x_{ij1} + x_{ij2}}{x_{i'j1} + x_{i'j2}} < \frac{x_{ij'1} + x_{ij'2}}{x_{i'j'1} + x_{i'j'2}}.$$

Using again (20) and (21), this inequality can be expressed as

$$(a - c) x_{ij2} x_{i'j1} < (a - c) x_{ij1} x_{i'j2}. \tag{22}$$

The condition that generates bias in the aggregate odds ratio is that countries' between-sector pattern of trade is related to per-capita income. Without loss of generality, assume that higher (lower) income per-capita is associated with a comparative advantage in $z = 1$ ($z = 2$) and that $y_i > y_{i'}$ and $y_j > y_{j'}$. This condition is formally expressed here as

$$\begin{aligned} \frac{x_{ij1} + x_{ij'1}}{x_{ij2} + x_{ij'2}} &> \frac{x_{i'j1} + x_{i'j'1}}{x_{i'j2} + x_{i'j'2}} \\ \frac{x_{ij1} + x_{i'j1}}{x_{ij2} + x_{i'j2}} &< \frac{x_{ij'1} + x_{i'j'1}}{x_{ij'2} + x_{i'j'2}} \end{aligned}$$

The first inequality states that, since exporter i has a comparative advantage in sector 1, it exports relatively more in that sector compared to exporter i'. The second inequality states that, since importer j has a comparative advantage in sector 1, it imports relatively less in that sector compared to importer j'. Using (20) and (21), these inequalities can be simplified to

$$x_{ij1} x_{i'j2} > x_{ij2} x_{i'j1} \tag{23}$$

and

$$c < a. \tag{24}$$

Conditions (23) and (24) imply that (22) also holds. Therefore, they are sufficient to induce bias. Furthermore, the stronger are the differences in the pattern of inter-sectoral trade between i and i' and between j and j', i.e. the stronger the inequalities in (23) and (24), the stronger will be the bias. Therefore, to the extent that sectoral trade patterns are related to per-capita incomes, the bias will be stronger the larger is the difference in per-capita income levels between i and i' and between j and j'.

Table 1. Trade flows between Rich and Poor countries in 1995
(cut-off per-capita income level: US$ 12000)

	Trade volume (billion US$)			Exporter group share in imports			Importer group share in exports		
	Rich importers	Poor importers	Total	Rich importers	Poor importers	Total	Rich importers	Poor importers	Total
a. 107 countries[1]									
Rich exporters	3,226	871	4,097	81.1%	81.5%	81.1%	78.7%	21.3%	100.0%
Poor exporters	754	198	952	18.9%	18.5%	18.9%	79.2%	20.8%	100.0%
Total	3,980	1,069	5,049	100.0%	100.0%	100.0%	78.8%	21.2%	100.0%
b. 64 countries[2]									
Rich exporters	3,181	711	3,892	83.2%	84.2%	83.4%	81.7%	18.3%	100.0%
Poor exporters	641	133	774	16.8%	15.8%	16.6%	82.8%	17.2%	100.0%
Total	3,822	844	4,666	100.0%	100.0%	100.0%	81.9%	18.1%	100.0%

[1] Includes countries with population greater than 1 million

[2] Includes countries with population greater than 5 million and imports of Differentiated goods greater than US$ 2 billion

Table 2. The aggregate Linder hypothesis
Standard formulation

	Linder term			
	$(\ln y_i - \ln y_k)^2$	$\lvert y_i - y_k \rvert$	$\ln\lvert y_i - y_k \rvert$	$\lvert \ln y_i - \ln y_k \rvert$
Linder term	0.176 ***	0.289 ***	0.129 ***	0.394 ***
	(0.020)	(0.058)	(0.032)	(0.055)
Exporter GDP	1.461 ***	1.357 ***	1.365 ***	1.450 ***
	(0.045)	(0.046)	(0.046)	(0.046)
Importer GDP	1.110 ***	1.006 ***	1.013 ***	1.098 ***
	(0.047)	(0.044)	(0.044)	(0.048)
Distance	-0.797 ***	-0.796 ***	-0.808 ***	-0.794 ***
	(0.048)	(0.048)	(0.048)	(0.048)
Border	1.263 ***	1.183 ***	1.187 ***	1.249 ***
	(0.203)	(0.203)	(0.206)	(0.204)
Common language	0.352 ***	0.372 ***	0.343 ***	0.388 ***
	(0.116)	(0.118)	(0.118)	(0.117)
PTA	0.337 ***	0.381 ***	0.297 **	0.389 ***
	(0.123)	(0.126)	(0.124)	(0.124)
Colonial link	2.241 ***	2.454 ***	2.493 ***	2.336 ***
	(0.160)	(0.164)	(0.166)	(0.161)
Common Colony	1.357 ***	1.428 ***	1.479 ***	1.416 ***
	(0.240)	(0.233)	(0.233)	(0.235)
Number of obs.	3789	3789	3789	3789
R^2	0.415	0.406	0.404	0.410

Estimation of equation (1) by OLS using 64 countries in the sample. Standard errors in parenthesis
***, **, * significant at the 1%, 5%, and 10%, respectively

Table 3. The aggregate Linder hypothesis
(with exporter and importer fixed effects)

	Linder term			
	$(\ln y_i - \ln y_k)2$	$\lvert y_i - y_k \rvert$	$\ln\lvert y_i - y_k \rvert$	$\lvert \ln y_i - \ln y_k \rvert$
Linder term	0.027* (0.015)	0.082** (0.033)	0.036* (0.019)	0.055 (0.035)
Distance	-1.218*** (0.038)	-1.214*** (0.039)	-1.216*** (0.039)	-1.217*** (0.039)
Border	0.026 (0.151)	0.034 (0.151)	0.037 (0.151)	0.029 (0.151)
Common language	0.783*** (0.093)	0.804*** (0.094)	0.796*** (0.094)	0.790*** (0.094)
PTA	-0.198** (0.096)	-0.171* (0.097)	-0.197** (0.096)	-0.199*** (0.097)
Colonial link	0.820*** (0.184)	0.819*** (0.183)	0.827*** (0.184)	0.827*** (0.184)
Common colony	0.622*** (0.195)	0.604*** (0.195)	0.613*** (0.194)	0.615*** (0.194)
Exporter fixed eff.	Yes	Yes	Yes	Yes
Importer fixed eff.	Yes	Yes	Yes	Yes
Number of obs.	3789	3789	3789	3789
R^2	0.819	0.819	0.819	0.819

Estimation of equation (1) by OLS using 64 countries in the sample. Standard errors in parenthesis
***, **, * significant at the 1%, 5%, and 10%, respectively

Table 4. The sectoral Linder hypothesis
Differentiated goods: OLS and ML estimates

	Regressions by Sector						Pooled Regression
	Sign		Significance (5%)				
	Pos.	Neg.	Pos.	Not Sig.	Neg.	Median	
A. OLS Estimation							
Linder term	34	82	21	36	59	-0.0714	-0.0435*** (0.0111)
Distance	0	116	0	0	116	-1.0154	
Border	107	9	58	58	0	0.3827	
Common language	114	2	101	15	0	0.5766	
PTA	103	13	69	47	0	0.4048	
Colonial link	112	4	88	28	0	0.7742	
Common colony	98	18	29	86	1	0.4026	
B. ML Estimation							
<u>Imports Equation</u>							
Linder term	42	74	27	34	55	-0.0396	
Distance	0	116	0	0	116	-1.1375	
Border	108	8	58	58	0	0.3950	
Common language	115	1	107	9	0	0.6356	
PTA	90	26	57	55	4	0.2544	
Colonial link	114	2	101	15	0	0.8601	
Common colony	101	15	46	70	0	0.4859	
<u>Fixed Cost Equation</u>							
Distance	89	27	38	78	2	0.0901	
Border	98	18	23	93	0	0.4571	
Common language	10	106	1	79	36	-0.2687	
PTA	14	102	2	63	51	-0.4298	
Colonial link	31	85	2	99	15	-0.4789	
Common colony	43	73	0	105	11	-0.1643	
Exporter GDP	27	89	2	71	43	-0.0935	
Importer GDP	79	37	44	63	9	0.0598	

Notes: Estimation of equation (14) in Panel A. Estimation of censoring model (equations 14 and 15) in Panel B. The Linder term is $(\ln y_i - \ln y_j)^2$. Columns 1 and 2 provide a breakdown of the total number of sectors by sign of the estimated coefficient. Columns 3 to 5 provide a breakdown by sign and significance. Heteroskedasticity-robust standard errors in all regressions. Clustering by country pair across sectors in pooled regression. Standard errors in parentheses.

***, **, * Significant at the 1%, 5%, and 10% levels, respectively.

Table 5. The sectoral Linder hypothesis: Alternative Linder terms
Differentiated goods: OLS and ML estimates

	Regressions by Sector						Pooled Regression
	Sign		Significance (5%)				
	Pos.	Neg.	Pos.	Not Sig.	Neg.	Median	
A. OLS Estimation							
$(\ln y_i - \ln y_j)^2$	34	82	21	36	59	-0.0714	-0.0435*** (0.0111)
$\lvert y_i - y_k \rvert$	34	82	20	38	58	-0.0136	-0.010*** (0.0024)
$\ln(\lvert y_i - y_k \rvert)$	37	79	17	46	53	-0.0554	-0.0402*** (0.0129)
$\lvert \ln(y_i) - \ln(y_k) \rvert$	33	83	22	34	60	-0.1739	-0.1144*** (0.0267)
B. ML Estimation							
$(\ln y_i - \ln y_j)^2$	42	74	27	34	55	-0.0396	
$\lvert y_i - y_k \rvert$	43	73	27	28	61	-0.0102	
$\ln(\lvert y_i - y_k \rvert)$	45	71	24	37	55	0.0455	
$\lvert \ln(y_i) - \ln(y_k) \rvert$	39	77	25	31	60	0.1136	

Notes: Estimation of equation (14) in Panel A. Estimation of censoring model (equations 14 and 15) in Panel B. Only results for the Linder terms are displayed in the table. Columns 1 and 2 provide a breakdown of the total number of sectors by sign of the estimated coefficient. Columns 3 to 5 provide a breakdown by sign and significance. Heteroskedasticity-robust standard errors in all regressions. Clustering by country pair across sectors in pooled regression. Standard errors in parentheses

***, **, * Significant at the 1%, 5%, and 10% levels, respectively.

Table 6. The sectoral Linder hypothesis on different categories of goods
All goods: OLS and ML estimates

	Regressions by Sector						Pooled Regression
	Sign		Significance (5%)				
	Pos.	Neg.	Pos.	Not Sig.	Neg.	Median	
A. OLS Estimation							
Differentiated	34	82	21	36	59	-0.0714	-0.0435*** (0.0111)
Reference-priced	13	43	5	30	21	-0.0455	-0.0423*** (0.0121)
Homogeneous	24	15	9	26	4	0.0287	0.0250* (0.0144)
B. ML Estimation							
Differentiated	42	74	27	34	55	-0.0396	
Reference-priced	19	37	7	25	24	-0.0411	
Homogeneous	26	13	10	27	2	0.0349	

Notes: Estimation of equation (14) in Panel A. Estimation of censoring model (equations 14 and 15) in Panel B. Only results for the baseline Linder term are displayed in the table. Columns 1 and 2 provide a breakdown of the total number of sectors by sign of the estimated coefficient. Columns 3 to 5 provide a breakdown by sign and significance. Heteroskedasticity-robust standard errors in all regressions. Clustering by country pair across sectors in pooled regression. Standard errors in parentheses

***, **, * Significant at the 1%, 5%, and 10% levels, respectively.

Table 7. The sectoral Linder hypothesis at different levels of aggregation
Differentiated goods: OLS and ML estimates

	Regressions by Sector						Pooled Regression
	Sign		Significance (5%)				
	Pos.	Neg.	Pos.	Not Sig.	Neg.	Median	
A. OLS Estimation							
3-digit	34	82	21	36	59	-0.0714	-0.0435*** (0.0111)
2-digit	12	24	6	11	19	-0.0652	-0.0345*** (0.0113)
1-digit	3	3	1	3	2	0.0070	0.0078 (0.0116)
All Differentiated	1	0	1	0	0		0.0988*** (0.0148)
B. ML Estimation							
3-digit	42	74	27	34	55	-0.0396	
2-digit	14	22	1	27	8	-0.0343	
1-digit	3	3	2	2	2	0.0069	
All Differentiated	1	0	1	0	0		0.0950*** (0.0142)

Notes: Estimation of equation (14) in Panel A. Estimation of censoring model (equations 14 and 15) in Panel B. Only results for the baseline Linder term are displayed in the table. Columns 1 and 2 provide a breakdown of the total number of sectors by sign of the estimated coefficient. Columns 3 to 5 provide a breakdown by sign and significance. Heteroskedasticity-robust standard errors in all regressions. Clustering by country pair across sectors in pooled regression. Standard errors in parentheses

***, **, * Significant at the 1%, 5%, and 10% levels, respectively.

Table 8. The aggregation bias: An illustrative example using actual trade flows in 1995

		Industrial Machinery (SITC 72)			Apparel (SITC 84)			Total		
		Importer			Importer			Importer		
		Switz. (j)	S.Africa (j')	Total	Switz. (j)	S.Africa (j')	Total	Switz. (j)	S.Africa (j')	Total
Value of exports (US$ millions)										
	U.S. (i)	116.0	260.0	376.1	32.7	4.7	37.4	148.8	264.7	413.5
Exporter:	China (i')	3.0	9.7	12.7	176.9	84.0	260.9	179.9	93.6	273.5
	Total	119.0	269.7	388.7	209.6	88.6	298.3	328.7	358.3	687.0
Exporter shares										
	U.S. (i)	97.5%	96.4%	96.7%	15.6%	5.3%	12.5%	45.3%	73.9%	60.2%
Exporter:	China (i')	2.5%	3.6%	3.3%	84.4%	94.7%	87.5%	54.7%	23.1%	39.8%
	Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Ratios										
	Ratio i/i'	38.76	26.92		0.18	0.06		0.83	2.83	
	Odds Ratio $(i/i')^j/(i/i')^{j'}$	1.44			3.32			0.29		

Table 9. The sectoral Linder hypothesis at different levels of aggregation
Reference-priced and Homogeneous goods: OLS and ML estimates

	Regressions by Sector						Pooled Regression
	Sign		Significance (5%)				
	Pos.	Neg.	Pos.	Not Sig.	Neg.	Median	
1. Reference-priced							
A. OLS Estimation							
3-digit	13	43	5	30	21	-0.0455	-0.0423*** (0.0121)
2-digit	6	22	3	13	12	-0.0471	-0.0531*** (0.0122)
1-digit	3	4	0	4	3	-0.0248	-0.0304** (0.0132)
All Ref.-priced	1	0	0	1	0		0.0157 (0.0190)
B. ML Estimation							
3-digit	19	37	7	25	24	-0.0411	
2-digit	8	20	3	12	13	-0.0315	
1-digit	3	4	1	4	2	-0.0199	
All Ref.-priced	1	0	0	1	0		0.0258 (0.0161)
2. Homogeneous							
A. OLS Estimation							
3-digit	24	15	9	26	4	0.0287	0.0250* (0.0144)
2-digit	8	12	2	16	2	-0.0231	-0.0092 (0.0143)
1-digit	3	3	0	5	1	-0.0010	-0.0163 (0.0158)
All Homogeneous	0	1	0	0	1		-0.0488** (0.0222)
B. ML Estimation							
3-digit	26	13	10	27	2	0.0349	
2-digit	11	9	3	16	1	0.0156	
1-digit	4	2	1	4	1	0.0313	
All Homogeneous	0	1	0	1	0		-0.0126 (0.0201)

Notes: Estimation of equation (14) in Panel A. Estimation of censoring model (equations 14 and 15) in Panel B. Only results for the baseline Linder term are displayed in the table. Columns 1 and 2 provide a breakdown of the total number of sectors by sign of the estimated coefficient. Columns 3 to 5 provide a breakdown by sign and significance. Heteroskedasticity-robust standard errors in all regressions. Clustering by country pair across sectors in pooled regression. Standard errors in parentheses

***, **, * Significant at the 1%, 5%, and 10% levels, respectively.

Table A1. List of countries and PPP GDP per capita in 1995

Country	GDP per Capita (PPP) US$	Country	GDP per Capita (PPP) US$
USA	27,395	Libya	7,180
Switzerland	25,475	Mexico	7,061
Norway	24,693	Poland	6,605
Japan	23,211	Brazil	6,572
Canada	23,085	Romania	6,430
Denmark	22,947	Thailand	6,217
Belgium-Lux.	22,700	Colombia	6,151
Singapore	22,270	Venezuela	5,979
Hong Kong	22,166	Costa Rica	5,940
Austria	22,089	Turkey	5,803
Germany	21,478	Bulgaria	5,608
Australia	21,267	Iran	4,968
Netherlands	20,812	Tunisia	4,870
Italy	20,512	Algeria	4,697
France	20,492	Paraguay	4,598
Sweden	20,030	Peru	4,329
United Kingdom	19,465	Dominican Rep.	3,997
Finland	18,764	Lebanon	3,964
New Zealand	17,705	Philippines	3,518
Israel	17,394	Guatemala	3,444
Ireland	17,264	Syria	3,211
Spain	15,163	Ecuador	3,162
Portugal	13,613	Morocco	3,052
South Korea	13,502	Indonesia	2,869
Taiwan	13,335	Egypt	2,869
Greece	13,147	Sri Lanka	2,741
Saudi Arabia	10,766	China	2,560
Argentina	10,736	India	1,877
South Africa	8,581	Pakistan	1,733
Malaysia	8,145	Vietnam	1,478
Uruguay	7,831	Bangladesh	1,253
Chile	7,544	Nigeria	832

Table A2. List of 3-digit SITC sectors

Differentiated goods sample (116 sectors)									
034	048	056	073	098	248	291	292	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	712	713
714	716	718	721	722	723	724	725	726	727
728	736	741	742	743	744	745	749	751	752
759	761	763	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 goods sample (56 sectors)									
011	014	034	036	037	054	057	058	081	112
122	233	251	266	273	278	322	323	334	335
341	511	512	513	514	515	516	522	523	524
533	541	562	582	583	584	592	634	641	642
651	652	653	661	662	671	672	673	674	677
678	682	684	693	699	778				
Homogeneous goods sample (39 sectors)									
001	011	022	023	024	041	042	043	044	054
057	058	061	071	081	222	232	246	247	251
263	268	281	282	287	288	333	334	423	424
522	634	667	681	682	683	684	686	689	