Estimating Cross-Country Differences in Product Quality*

Juan Carlos Hallak†
University of Michigan & NBER

Peter K. Schott‡
Yale School of Management & NBER

PRELIMINARY AND INCOMPLETE

July 15, 2005

Abstract

This paper develops a methodology for identifying countries' unobserved product quality from observed information on their trade flows. In contrast to a naive approach that equates export price with quality, our methodology accounts for cross-country variation in product prices induced by other factors, e.g., comparative advantage or currency misalignment. It also accounts for the impact that horizontal differentiation and consumer love of variety have on enhancing the demand for imports from countries that produce many varieties. We infer quality by combining information about export prices with data on net trade. Holding prices constant, our framework implies that countries exhibiting trade surpluses must be offering higher quality than countries running trade deficits. We implement the methodology by estimating the evolution of manufacturing product quality among the top 45 U.S. trading partners from 1980 to 1997.

Keywords: Import Prices; Export Quality; Revealed Preferences; Paasche; Laspeyres

JEL classification: F1; F2; F4

*Special thanks to Alan Deardorff for many fruitful discussions. We also thank Keith Chen, Rob Feenstra, Justin McCrary, Peter Neary, Serena Ng, Ben Polak, and seminar participants at the LSE, Maryland, Michigan, San Andrés, and Penn State for their comments and insight.

†611 Tappan Street, Ann Arbor, MI 48109, *tel*: (734) 763-9619, *fax*: (734) 764-2769, *email*: hallak@umich.edu

‡135 Prospect Street, New Haven, CT 06520, *tel*: (203) 436-4260, *fax*: (203) 432-6974, *email*: peter.schott@yale.edu

1. Introduction

A growing body of empirical research suggests that variation in product quality is an important determinant of international trade patterns and economic development.[1] Failure to account for this variation can result in erroneous conclusions about both the usefulness of trade theory and the effectiveness of trade policy. Unfortunately, reliable data on product quality are unavailable for most countries, industries and years. We address this need by developing a methodology for identifying countries' unobserved product quality from observed variation in their export unit values and quantities.

Export prices (unit values) vary considerably across countries even within narrowly defined product categories. It is often assumed that this variation is driven by quality, which we define here to be any tangible or intangible attribute – such as durability or cachet – that increases consumers' valuation of a product. Shirt varieties from Italy that are twice as expensive as shirt varieties from China, for example, might be considered to have double the quality, perhaps because they are two times as "fashionable". Similar assumptions are invoked in studies of intra-industry trade, where horizontal and vertical trade flows are differentiated according to the magnitude of their underlying unit value ratios.[2] In policy research, estimates of countries' "quality competitiveness" are often derived from cross-country comparisons of export unit values.[3]

International variation in export prices, however, may be influenced by factors other than product quality. In particular, they may be subject to variation in quality-adjusted prices due, for example, to comparative advantage: shirt prices may vary across exporters because of differences in countries' relative production efficiency or factor costs. The goal of this paper is to devise a technique for decomposing prices into quality versus quality-adjusted components.

Our focus on cross-sectional variation in quality differentiates this study from a very large literature on index number theory that constructs price indexes that adjust prices for time-series variation in product quality. Here, we have a different aim: rather than measure quality changes in bundles of products purchased over time, we seek to identify quality

[1] Recent research suggests that export quality is influential in determining the direction of trade (Hallak 2005), the skill premium in developing countries (Verhoogen 2004), and export success among firms (Brooks 2003, Verhoogen 2004). Data aggregation that obscures cross-country variation in product quality also helps explain previous poor empirical support for the Heckscher-Ohlin model (Schott 2003, 2004).

[2] Abed-el-Rahman (1991), for example, suggests that exports with unit values ratios between (outside) 0.85 to 1.15 are horizontally (vertically) differentiated. Using this rule of thumb, Aturupane et al. (1999) find a positive association between vertical intra-industry trade and product differentiation, economies of scale, the labor intensity of production and foreign direct investment. Aiginger (1997) proposes further differentiating vertical intra-industry trade flows according to whether they are accompanied by a trade surplus or deficit. We rely on a similar intuition below.

[3] See, for example, Aiginger (OECD 1998), Verma (ICRIER 2002) and Ianchovichina et al. (IADB/World Bank 2003).

variation over simultaneously purchased bundles from different sources of supply.[4] We also operate under different information constraints. In particular, we assume no knowledge of products' underlying attributes. As a result, we are unable to make use of standard strategies – such as hedonic pricing – that exploit additional information on product characteristics to adjust price indexes for quality.[5] Our methodology complements such efforts, however, because its use of readily available trade data permits estimation of product quality across a broad range of countries, products and years for which surveys of product characteristics may be unavailable or prohibitively expensive to collect.

Although we do not exploit hedonics, our methodology does rely upon information about consumer demand, captured in countries' net trade, to extract estimates of relative quality from observed relative prices. The intuition for our technique is straightforward: because consumers care about price in relation to quality in choosing among products, countries that export products with high relative prices are estimated to possess higher relative quality if they run a trade surplus than a trade deficit. That is, countries would not be able to run a trade surplus with high export prices unless their quality was high, i.e., their price per unit of quality was low.

Our methodology generalizes this intuition by deriving a theoretical relationship between quality-adjusted prices and sectoral net trade in the presence of both horizontal differentiation and consumer taste for variety.[6] These two features are desirable given the large body of theoretical and empirical research in international trade that emphasizes the importance of love of variety and horizontal differentiation in determining bilateral trade flows, the evolution of countries' terms of trade, and exchange rates.[7] Nevertheless, they complicate the analysis considerably: allowing countries to produce an unobserved number of horizontal varieties within each product category introduces an additional factor besides quality that can increase consumer demand for a product. All else equal, consumer taste for variety implies that countries exporting a larger number of varieties in a product category will export larger quantities and therefore have higher net trade. This increase in net trade

[4] Aw and Roberts (1986) and Feenstra (1988) identify quality upgrading over time – defined as an across-product shift from lower- to higher-priced exports – in response to quantitative trade restrictions. In addition to focusing on time-series changes, that conceptualization of quality change differs from ours in that it does not take into account potential within-product changes in product quality.

[5] Feenstra (1995), for example, demonstrates how information on product attributes can be used to establish bounds on the exact hedonic price index. Data constraints also prevent us from adopting the approach of the International Price Program of the U.S. Bureau of Labor Statistics, which constructs import and export price indexes by combining survey data on firm prices with firms' assessments about changes in their products' quality over time. See www.bls.gov/mxp/ for more detail.

[6] Absent horizontal differentiation, prices solely reflect quality variation, as any difference in quality-adjusted prices would be arbitraged away. Perfect substitutability among products of varying quality is assumed in typical models of quality ladders, such as Grossman and Helpman (1991).

[7] Seminal theoretical contributions include Krugman (1980, 1981), Helpman and Krugman (1985) and Krugman (1989).

will manifest, erroneously, as higher quality unless we control for the number of varieties countries produce. Our methodology accounts for cross-country differences in the number of varieties by tying them to variation in countries' quality-adjusted prices.

We link unobserved bilateral product quality to observed export prices and net trade in four steps. First, we construct an index of (unobservable) quality-adjusted prices in a sector, and show that net trade (in the sector) is a function of that index. Second, we define an index of unadjusted prices, which can be decomposed into the aforementioned quality-adjusted index and an index of quality. Even though the index of unadjusted prices is composed of observed prices, it is unobservable because it incorporates unknowns such as the number of varieties countries produce. However, using revealed preference analysis, we show that this unobserved index is bounded by observable Paasche and Laspeyres price indexes defined over countries' common exports to a third country. Third, we outline a strategy for using a large set of Paasche-Laspeyres bounds across country pairs to estimate an unadjusted price index number for each country relative to a numeraire country. Finally, we estimate countries' product quality relative to the numeraire by correcting the estimated index of unadjusted prices to account for variation in quality-adjusted prices. Exploiting our first result, we use net trade information to infer this variation. More detailed intuition for our methodology is provided in Section 2.

Our approach to estimating quality in this paper is most closely related to a recent study by Hummels and Klenow (2005). Even though their aim is not to derive quality estimates for each country, they infer the cross-sectional elasticity of quality with respect to per-capita income and country size under various assumptions about the number of varieties countries produce as well as elasticities of substitution. For example, they calculate that countries with twice the per-capita income have product quality that is 9 to 23 percent higher. Here, we provide a methodology that allows for explicit estimation of relative quality levels by country, sector and year.

Reliable estimates of product quality are obviously useful for testing trade models. Another promising and related application of our methodology is the measurement of national accounts aggregates at internationally comparable prices. Current estimates of "real GDP", such as the Penn World Tables, deflate nominal GDP using a purchasing power parity (PPP) deflator based on final expenditure data. Though helpful to gauge demand, this measure of real GDP may not be optimal for capturing changes in countries' *production* over time. A recent contribution to the literature by Feenstra et al. (2004) suggests using a PPP index constructed from output prices rather than expenditure prices to deflate nominal GDP. This new deflator would include a terms-of-trade correction estimated from import and export price indexes based on unit values. An ability to net quality out of these indexes before performing the terms-of-trade correction would enhance their accuracy.

The remainder of the paper is structured as follows. Section 2 provides an intuitive overview of our methodology. Section 3 derives the relationship between net trade and quality-adjusted prices. Section 4 derives observable bounds on the (unobservable) index of unadjusted prices. Section 5 describes the empirical implementation of the methodology and the results. Section 6 concludes.

2. A Road Map for the Paper

In this section we describe the nature of our identification problem and provide an intuitive discussion of our methodology.

Suppose that the prices of two export varieties of the same product are observed, for example two different men's shirts originating from two different countries. We want to identify the extent to which price variation across these varieties reflects differences in product quality versus differences in price for the same quality. That is, we want to decompose the observed relative "impure" price p into two elements: relative product quality, λ, and relative quality-adjusted, or "pure", price, $\widetilde{p}$, where $p = \lambda\widetilde{p}$. Because $\widetilde{p}$ and λ are not observed, we need to impose assumptions that allow them to be inferred from observables. This paper is an effort to determine how unobserved product quality can be recovered from observed international trade data.

We identify quality by combining information about relative export quantities and prices (obtained from observed export unit values) with inferences about consumer valuation drawn from countries' trade balances. Under common assumptions about consumer preferences, price and quality are not evaluated independently. Rather, the key variable determining consumer demand is a good's pure price, i.e. its price per-unit of quality ($\widetilde{p} = p/\lambda$). Because $\widetilde{p}$ is unobservable, we make inferences about it via some observable aspect of consumer demand. Trade flows are the only expression of consumer valuation that is available for a wide range of countries and sectors: the higher a country's trade surplus, the greater the consumer demand for its products. Once we infer $\widetilde{p}$ from consumer demand, we can back out relative quality, λ, from the observable impure price, p. Intuitively, countries exporting goods with high impure prices and a trade surplus must be exporting goods of higher quality than countries exporting goods with the same impure prices and a trade deficit.

The use of net trade to infer consumer demand imposes a practical constraint on our methodology, limiting its application to the level of aggregation at which trade balances are observed. Currently, the most reliable time-series information on sectoral net trade for a large sample of countries, the World Trade Flows database compiled by Feenstra et al. (2004), is restricted to a relatively small number of industries compared to the relatively large number of products at which trade for particular countries (e.g. the United States) is tracked. In light of this practical constraint, we develop a methodology for identifying export quality by sector rather than by product. In doing so, we assume that each country's export quality is constant across all products in a sector. This assumption, coupled with another requirement that products in the same sector share a common elasticity of substitution, highlights an "aggregation tension" in our methodology. Greater disaggregation is more consistent with assuming common quality and a common elasticity of substitution, but is at odds with our ability to observe net trade for a large sample of diverse countries. On the other hand, for sufficiently small deviations from these assumptions, aggregation increases sample size and therefore heightens our estimates' robustness to measurement error.

Given our need to aggregate, our challenge *vis a vis* the simple example above is to define an "index" of observed relative product prices between a pair of countries, c and d, in sector s, denoted P_s^{cd}, that can be decomposed into relative quality, λ_s^{cd}, and an index of relative pure prices, $\widetilde{P}_s^{cd}$, such that $P_s^{cd} = \widetilde{P}_s^{cd}\lambda_s^{cd}$. To be useful, this impure price index must possess two properties. First, its pure-price component should play the same role in sectoral consumer demand, and hence the sectoral trade balance, that the single relative pure price played in the motivating example above. Second, it must be either observable or feasible to estimate.

In Section 3 we demonstrate that countries' net trade can be expressed as a linear function of $\widetilde{P}_s^{co}$ and trade costs. Because $P_s^{cd} = \widetilde{P}_s^{cd}\lambda_s^{cd}$, this relationship satisfies the first property required of our impure price index. In Section 4 we show that P_s^{cd}, though unobservable, is bounded by observable bilateral Paasche and Laspeyres indexes defined over the country pairs' common exports to a third country. The existence of these bounds satisfies our second requirement for an impure price index. We illustrate in Section 5 how an index number for each country relative to a numeraire country o, P_s^{co}, can be estimated from the full set of bilateral Paasche-Laspeyres intervals defined by all country pairs. The intuition for this estimation comes from the transitivity of the impure price index: because $P_s^{cd} = P_s^{co}/P_s^{do}$, the bounds placed on the impure index between countries c and d impose implicit constraints on P_s^{co} and P_s^{do}.

Armed with estimates of P_s^{co} and using net trade as indicator of $\widetilde{P}_s^{co}$, we are able to identify elasticity parameters, in particular the elasticity of net trade with respect to pure prices and trade costs. Combining P_s^{co}, $\widetilde{P}_s^{co}$, and the necessary elasticity parameters, we back out the evolution of product quality of the top 45 U.S. trading partners between 1980 and 1997.

3. Net Trade as Indicator of Pure Price Variation

This section derives the theoretical relationship between net trade and an index of "pure", i.e., "quality-adjusted", prices.

3.1. Net trade as a function of pure prices

Goods are classified into product categories, which are in turn classified into sectors. Sectors are indexed by $s = 1, ..., S$, while product categories (within sectors) are indexed by $z = 1, ..., Z_s$.[8] There are C countries, indexed by $c = 1, ..., C$.

Preferences are common across countries, and are represented by a two-tier utility function. The upper tier is Cobb-Douglas, with expenditure shares b_s for each sector s. The lower tier has the following CES form[9]

$$u_s = \left[\sum_c \sum_z n_z^c \left(\xi_z \lambda_s^c x_z^c \right)^{\varphi_s} \right]^{1/\varphi_s} \qquad \varphi_s \epsilon (0, 1). \tag{1}$$

In the subutility function (1), n_z^c is the number of horizontally differentiated varieties of product z produced by country c, and x_z^c is the quantity consumed per variety.[10] This function includes a utility shifter or preference parameter, ξ_z, which is standard for allowing asymmetric preferences across different classes of products (e.g. tables versus chairs). This shifter captures consumers' common valuation of the essential characteristics that define heterogeneous varieties in a particular product category. The utility shifter ξ_z varies across product categories, but is constant across countries for a particular product category. The subutility function includes a second utility shifter, λ_s^c, which we interpret as product quality. Product quality captures the combined effect of all attributes of a good, other than price and the common attributes already captured by ξ_z, on consumers' valuation of the good. Product quality then represents both physical characteristics (e.g. durability) and intangible characteristics (e.g. product image due to advertising) of a good. Product quality varies across countries and sectors, but is constant across products within a particular country and sector. These assumptions can be formalized as follows:

Assumption 1: $\xi_z^c = \xi_z, \qquad \forall c = 1, ..., C.$

Assumption 2: $\lambda_z^c = \lambda_s^c, \qquad \forall z = 1, ..., Z_s.$

[8] In our empirical investigation below, product categories correspond to seven-digit Tariff System of the United States (TSUSA) and ten-digit Harmonized System (HS) categories, the finest possible level of aggregation.

[9] To simplify notation, subindexes on summations refer to all members of a set unless otherwise noted, e.g. $\sum_c$ and $\sum_{c'}$ both sum over all countries $c = 1, ..., C$ while $\sum_{c' \neq c}$ sums over all countries except c. For product categories, $\sum_z$ denotes the sum across all product varieties in sector s, $z = 1, ..., Z_s$.

[10] Note that by indexing product categories instead of varieties, we implicitly assume symmetry across varieties in the same product category.

Let p_z^c be the export price of a typical variety of product z produced in country c, and let $\widetilde{p}_z^c = \frac{p_z^c}{\xi_z \lambda_s^c}$ be the "pure" price of that variety. Note that in contrast to the pure price defined in the motivating example in Section 2, we now define the pure price to include the utility shifter ξ_z. The pure price is a quality-adjusted price. It is also divided here by ξ_z for notational compactness, but none of the results or their interpretation is affected by this choice. Exporting goods from country c to country c' requires paying trade costs of $\tau_s^{cc'}$. Therefore, $p_z^c \tau_s^{cc'}$ is the import price of product z in country c'. Given the preference structure assumed in (1), we can derive country c''s import demand from country c (in sector s) using well-known results associated with CES prefernces. Summing over $c' \neq c$, the value of country c's exports, X_s^c, is

$$X_s^c = \sum_{c' \neq c} \left[\sum_z \frac{n_z^c \left(\widetilde{p}_z^c \tau_s^{cc'} \right)^{1-\sigma_s}}{\left(G_s^{c'} \right)^{1-\sigma_s}} \right] b_s Y^{c'} \tag{2}$$

where $Y^{c'}$ is the income of country c', $\sigma_s = \frac{1}{1-\varphi_s} > 1$ is the elasticity of substitution,

$$\left(G_s^{c'} \right)^{1-\sigma_s} = \sum_{c''} \sum_z n_z^{c''} \left(\widetilde{p}_z^{c''} \tau_s^{c''c'} \right)^{1-\sigma_s} \tag{3}$$

is a price aggregator measuring country c's trade resistance, and the expression in brackets in equation (2) is country c's share in country c''s sectoral expenditure, $b_s Y^c$. This share does not depend on prices and quality levels independently of one another, but only on the ratio of the two, $\widetilde{p}_z^c$.[11]

In a similar manner, we can obtain the value of country c's imports:

$$M_s^c = \sum_{c' \neq c} \left[\sum_z \frac{n_z^{c'} \left(\widetilde{p}_z^{c'} \tau_s^{c'c} \right)^{1-\sigma_s}}{\left(G_s^c \right)^{1-\sigma_s}} \right] b_s Y^c = \left[1 - \sum_z \frac{n_z^c \left(\widetilde{p}_z^c \right)^{1-\sigma_s}}{\left(G_s^c \right)^{1-\sigma_s}} \right] b_s Y^c. \tag{4}$$

We can now use equations (2) and (4) to calculate country c's net trade in sector s, T_s^c, as a proportion of its expenditure in the sector:

$$\frac{1}{b_s} \frac{T_s^c}{Y^c} = -1 + \sum_{c'} \sum_z \frac{n_z^c \left(\widetilde{p}_z^c \tau_s^{cc'} \right)^{1-\sigma_s}}{\left(G_s^{c'} \right)^{1-\sigma_s}} \frac{Y^{c'}}{Y^c}. \tag{5}$$

Since a similar expression holds for a numeraire country o, we can express the difference in net export positions between countries c and o as

$$\frac{1}{b_s} \left[\frac{T_s^c}{Y^c} - \frac{T_s^o}{Y^o} \right] = \sum_{c'} \sum_z \left[\frac{n_z^c \left(\widetilde{p}_z^c \tau_s^{cc'} \right)^{1-\sigma_s}}{\left(G_s^{c'} \right)^{1-\sigma_s}} \frac{Y^{c'}}{Y^c} - \frac{n_z^o \left(\widetilde{p}_z^o \tau_s^{oc'} \right)^{1-\sigma_s}}{\left(G_s^{c'} \right)^{1-\sigma_s}} \frac{Y^{c'}}{Y^o} \right]. \tag{6}$$

[11] We can associate and infinite price $\widetilde{p}_z^c$ with a product z that is not produced in country c. Since pure prices are elevated to a negative exponent, this product will have no effect on the volume of trade or the price aggregator.

Equation (6) shows that net trade is a function of all pure prices at the product category level. Our objective next is to simplify this expression by relating differences in net trade between countries c and o in sector s to a summary measure of pure price differences between the two countries in that sector.

3.2. *The Pure Price Index*

Define $\overline{n}_s^c$ to be the average number of varieties across product categories produced by country c (in sector s),

$$\overline{n}_s^c = \frac{1}{Z_s}\sum_z n_z^c \qquad \forall c = 1, ...C. \tag{7}$$

Define $\overline{n}_z$ to be the ("country o–normalized") world average number of varieties of product z,

$$\overline{n}_z = \frac{1}{C}\sum_c n_z^c \frac{\overline{n}_s^o}{\overline{n}_s^c} \qquad \forall z = 1, ...Z_s. \tag{8}$$

The normalization in (8) re-scales the number of varieties of each country into common, country-o units, according to the ratio of the average number of varieties between o and c.

Define $\widetilde{n}_z^c$ to be country c's "excess variety" in product z relative to the world average,

$$\widetilde{n}_z^c = n_z^c \frac{\overline{n}_s^o}{\overline{n}_s^c} - \overline{n}_z. \tag{9}$$

Note that excess variety has the convenient property $\sum_z \widetilde{n}_z^c = 0, \forall c = 1, ..., C$.

Equipped with this notation, we can now define a pure price "aggregator"[12]

$$\widetilde{P}_s^c = \left[\sum_z \overline{n}_z \left(\widetilde{p}_z^c\right)^{1-\sigma_s}\right]^{\frac{1}{1-\sigma_s}} \tag{10}$$

Based on this aggregator, we finally define the Pure Price Index between countries c and d as

$$\widetilde{P}_s^{cd} = \frac{\widetilde{P}_s^c}{\widetilde{P}_s^d} = \left[\frac{\sum_z \overline{n}_z \left(\widetilde{p}_z^c\right)^{1-\sigma_s}}{\sum_z \overline{n}_z \left(\widetilde{p}_z^d\right)^{1-\sigma_s}}\right]^{\frac{1}{1-\sigma_s}}. \tag{11}$$

The Pure Price Index is a summary measure of pure price variation between countries. The index has the desirable property of transitivity, so that $\widetilde{P}_s^{cd}\widetilde{P}_s^{do} = \widetilde{P}_s^{co}$. Therefore, choosing country o as the numeraire, or base country, we can associate an index number, $\widetilde{P}_s^{co}$, with each country c, noting that we can always obtain $\widetilde{P}_s^{cd}$ from the ratio $\widetilde{P}_s^{co}/\widetilde{P}_s^{do}$. In particular, the value of this ratio is independent of which country is chosen as the numeraire.

[12] This type of price aggregator is often called a price "index" in the trade literature (e.g. Anderson and van Wincoop, 2004). We reserve the term "index" here for price comparisons between countries, in accordance with terminology employed in the index number literature.

3.3. Net trade as a function of the Pure Price Index

To express equation (6) as a function of the Pure Price Index, we must impose structure on the relationship between pure prices and number of varieties countries produce. Note, however, that our methodology does not require that we identify the economic forces that determine pure prices in equilibrium. Variation in pure prices can be driven by traditional sources of comparative advantage, or it can be the result of macroeconomic conditions, such as over- or under-valued currencies.

Theoretical models of international trade with product differentiation that do not assume factor price equalization (e.g., Romalis 2004, Bernard et al. 2004) find that, across sectors, the relative number of varieties between two countries is a negative function of the countries' relative prices. This finding supports the intuitive notion that countries should have a relatively higher (lower) number of firms in sectors in which they are relatively more (less) competitive, i.e. those sectors with relatively lower (higher) prices. It is possible to reformulate these models in terms of quality-adjusted variables. Thus reinterpreted, these models predict that the relative number of varieties in a sector is a negative function of relative pure (or quality-adjusted) prices in that sector. Based on the results of these models, we postulate the following negative relationship between the average number of varieties and the Pure Price Index:

Assumption 3: $\frac{\overline{n}_s^c/Y^c}{\overline{n}_s^o/Y^o} = \left(\widetilde{P}_s^{co}\right)^{-\eta_s}, \quad \forall c = 1, ..., C, \qquad \eta_s > 0.$

Assumption 3 focuses on the relationship between pure prices and number of varieties across sectors. However, pure prices and number of varieties also vary across product categories within sectors. Here, we consider the sample covariance between excess variety and pure price relative to the aggregator $\widetilde{P}_s^c$. This covariance can be expressed as the sum of a common component across countries (V_s), and a mean-zero, country-specific idiosyncratic component:

$$cov\left[\widetilde{n}_z^c, \left(\frac{\widetilde{p}_z^c}{\widetilde{P}_s^c}\right)^{1-\sigma_s}\right] = V_s + \theta_s^c, \quad \forall c = 1, ..., C. \tag{12}$$

Based on the same theoretical results underlying Assumption 3, we expect this covariance to be positive: products with a lower price (relative to the pure price aggregator) should have on average positive excess variety. Nevertheless, we do not need to impose a particular sign for this covariance, leaving it open as an empirical question.

The objective of this section is to derive an expression relating net trade at the sectoral level to the value of the Pure Price Index. Since net trade also depends on trade costs, we also want this expression to depend on summary measures of trade costs in the sector. To that end, we define some additional variables. Let $g^c = Y^c / \sum_{c'} Y^{c'}$ be the share of country c in world income, and let $r_s^c = \frac{1}{G_s^{1-\sigma_s}} \sum_z n_z^c \left(\widetilde{p}_z^c\right)^{1-\sigma_s}$ be the share of country c in the price aggregator $G_s^{1-\sigma_s} = \sum_{c'} \sum_z n_z^{c'} \left(\widetilde{p}_z^{c'}\right)^{1-\sigma_s}$, which is common for all countries under free trade ($G_s^{c'} = G_s$), and is thus denoted omitting the country superscript. In the free-trade equilibrium with those pure prices and number of varieties, r_s^c is also the share of country c in world expenditure (in sector s). We can now define summary measures of "inbound" and "outbound" trade costs for country c, respectively, as

$$\tau_s^{\overline{c}c} = g^c \sum_{c' \neq c} r_s^{c'} \left(\tau_s^{c'c} - 1\right) \tag{13}$$

$$\tau_s^{c\overline{c}} = (1 - r_s^c) \sum_{c' \neq c} g^{c'} \left(\tau_s^{cc'} - 1\right) \tag{14}$$

The inbound average trade cost, defined in equation (13), is a weighted average, across countries, of the bilateral costs of exporting from other countries to country c. The weights are the shares of each country in the price aggregator $G_s^{1-\sigma_s}$, and capture the importance of a country as a producer in sector s. The adjustment of this average by the term g^c is discussed below. Similarly, the outbound average trade cost, defined in equation (14), is a weighted average, across countries, of the bilateral costs of exporting from country c to other countries. In this case, the weights are the shares in world income. As with the inbound average, the adjustment by the term $(1 - r_s^c)$ is discussed below.

The following Proposition describes the main result of this section.

Proposition 1 *Under Assumption 3, country c's sectoral net trade relative to that of country o can be approximated (via a Taylor expansion) as a linear function of the Pure Price Index and these countries' inbound and outbound average trade costs,*

$$\left[\frac{T_s^c}{Y^c} - \frac{T_s^o}{Y^o}\right] \simeq \gamma_s \ln \widetilde{P}_s^{co} - \gamma_s \mu_s \left[\tau_s^{\overline{c}c} - \tau_s^{\overline{c}o}\right] + \gamma_s \mu_s \left[\tau_s^{c\overline{c}} - \tau_s^{o\overline{c}}\right] - \gamma_s \theta_s^{co}, \tag{15}$$

$$\gamma_s = \frac{(1-\sigma_s-\eta_s) b_s \Psi_s^o \left(\widetilde{P}_s^o\right)^{1-\sigma_s}}{Y^o}, \quad \Psi_s^o = \left[1 + Z_s \left(V_s + \theta_s^o\right)\right] \sum_{c'} \frac{Y^{c'}}{\left(G_s^{c'}\right)^{1-\sigma_s}} \left(\tau_s^{oc'}\right)^{1-\sigma_s},$$

$$\mu_s = \frac{(\sigma_s - 1)}{(\sigma_s + \eta_s - 1)}, \qquad \theta_s^{co} = \frac{Z_s}{(\sigma_s + \eta_s - 1)} \left(\theta_s^c - \theta_s^o\right)$$

Proof. See Appendix A. ■

In equation (15), we expect $\gamma_s < 0$, as we expect $\Psi_s^o > 0$ and $1 - \sigma_s - \eta_s < 0$. The term Ψ_s^o can only be negative if the covariance in equation (12) is sufficiently negative for the numeraire country. Since we expect this not be true in general, we continue the analysis under the assumption that $\gamma_s < 0$, but later confirm empirically the validity of this assumption. In turn, $\mu_s > 0$, as $\sigma_s > 1$ and $\eta_s > 0$.

Proposition 1 provides a simple and intuitive expression for the relationship between net trade, pure prices and trade costs. It formalizes the simple idea that the surplus in a country's net trade position should be larger the lower are its pure prices, as summarized by the Pure Price Index. In addition to pure prices, trade costs also influence net trade. In particular, inbound trade barriers for country c, as summarized by $\tau_s^{\bar{c}c}$, have a positive effect on net trade, while outbound trade barriers, as summarized by $\tau_s^{c\bar{c}}$, have a negative effect on net trade. The average inbound trade cost is multiplied by g^c to capture the fact that the effect of trade barriers on net trade, at a free-trade equilibrium, is proportional to country size.[13] The multiplication of the average outbound trade cost by $(1 - r_s^c)$ incorporates the opposite effect, i.e., the (negative) influence of outbound tariffs is decreasing in the relative importance of country c in sector s. Finally, we know that the effects of trade costs on net trade characterized in Proposition 1 are "conditional on pure prices". This implies that, while they appropriately adjust the relationship between net trade and pure prices, they do not provide a comparative statics assessment of the impact of inbound and outbound trade costs on net trade. Changes in those costs will typically affect pure prices in general equilibrium, implying an indirect effect on net trade not captured in equation (15).

Equation (15) can be interpreted as a relative demand function, where net trade is the "quantity" variable, the Pure Price Index is the "price" variable, and the trade costs are demand shifters. The first term captures movements along the demand curve: higher pure prices of country c in sector s (relative to those of country o) are associated with a worsening of this country's net trade position in that sector (relative to that of country o). The second and third terms capture movements of the demand curve. Conditional on pure prices, inbound trade costs shift the demand curve to the right, while outbound trade costs shift this curve to the left.

4. Using Revealed Preferences to Bound an "Impure" (i.e., Non-Quality-Adjusted) Price Index

There exists an index of observed prices that can be decomposed into the Pure Price Index and an index of quality. Even though this index depends on observed prices, it also depends on unobservables, such as the number of varieties and the elasticity of substitution. Therefore, the index is itself unobservable. However, we show in this section that, under plausible assumptions, it can be bounded between observable Paasche and Laspeyres indexes. Overlapping bilateral bounds across country pairs then allows identification of the value of this unobservable index for a set of exporters.

[13] This result is analogous to Implication 1 in Anderson and van Wincoop (2003).

4.1. Definitions and Notation

Define the Impure Price Index between countries c and d as

$$P_s^{cd} = \frac{P_s^c}{P_s^d} = \left[\frac{\sum_z \overline{n}_z \xi_z^{\sigma_s - 1} \left(p_z^c\right)^{1-\sigma_s}}{\sum_z \overline{n}_z \xi_z^{\sigma_s - 1} \left(p_z^d\right)^{1-\sigma_s}}\right]^{\frac{1}{1-\sigma_s}}. \tag{16}$$

The Impure Price Index is a summary measure of observed price differences between countries c and d in sector s. The index is "impure" in the sense that it is defined over prices that are "contaminated" by quality. The value of the index rises the higher are the prices of country c relative to those of country d. As was the case with the Pure Price Index, the Impure Price Index is also transitive.

Define also a "Quality Index" between countries c and d as

$$\lambda_s^{cd} = \frac{\lambda_s^c}{\lambda_s^d}. \tag{17}$$

Since quality is assumed to be constant across products within a sector, the Quality Index simply measures the ratio of quality levels between two countries.

The Impure Price Index can be expressed as the product of the Pure Price Index and the Quality Index:

$$P_s^{cd} = \widetilde{P}_s^{cd} \lambda_s^{cd} \tag{18}$$

where the decomposition highlights the fact that, as opposed to the usual approach of inferring quality from differences in observed prices (export unit values), price differences across countries can also result from differences in pure prices, i.e., price differences not associated with quality differences.

In this section, we focus on countries' exports to a single "common importer", which we refer to as the United States given the focus of our empirical examination below. Nevertheless, the analysis would be identical were it to be applied to any other common importer. We define a country as "active" in product z if it reports positive exports to the United States in that category. Let I_s be the set of all product categories in sector s, and let I_s^c be the subset of active categories in country c. Define vector $\mathbf{p}_s$ to include the U.S. import prices of all active categories in sector s from all countries. Define analogously vectors $\mathbf{q}_s, \mathbf{n}_s, \boldsymbol{\lambda}_s$, and $\boldsymbol{\xi}_s$. A vector of per-variety consumption $\mathbf{x}_s$ is implicitly defined by $\mathbf{q}_s$ and $\mathbf{n}_s$. Finally, stack these vectors across sectors to form vectors $\mathbf{p}$, $\mathbf{q}$, $\mathbf{n}$, $\boldsymbol{\lambda}$, $\boldsymbol{\xi}$, and $\mathbf{x}$.

Since our methodology is based on comparing import prices (as measured by unit values) across pairs of U.S. trading partners, we need to use notation specific to country pairs. Index countries in a pair of U.S. trading partners by c and d. Denote by I_s^{cd} the set of active categories common to c and d in sector s. Z_s^{cd} is the number of such categories.

Denote also by $I_s^{c,-d}$ the set of products in which c is active but not d, by $I_s^{d,-c}$ the set of products in which d is active but not c, and by U_s^{cd} the union of these two sets. Finally, $\varnothing_s^{cd}$ is the set of products in which neither of the two countries is active. The set I_s can be partitioned into I_s^{cd}, U_s^{cd}, and $\varnothing_s^{cd}$. We can use I_s^{cd} to break each of vectors $\mathbf{p}$ and $\mathbf{q}$ into two components. First, alternatively for each $i = c, d$, $\mathbf{p}_{s(cd)}^i$ and $\mathbf{q}_{s(cd)}^i$ include prices and quantities, respectively, of exports by i in products $z \in I_s^{cd}$. The remaining parts of $\mathbf{p}$ and $\mathbf{q}$ are denoted by $\mathbf{p}_{s(cd)}^{-i}$ and $\mathbf{q}_{s(cd)}^{-i}$. These vectors include categories $z \in I_s^{cd}$ exported by all countries other than i, and also categories $z \notin I_s^{cd}$ exported by all countries (including i).[14]

4.2. The Conditional Expenditure Function

For a pair of exporting countries c and d, we now define the conditional expenditure (or import) function $m_{s(cd)}^c(\mathbf{p}_s^i, \mathbf{q}_s^{-c}, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U)$. This function represents the minimum expenditure that the representative consumer in the U.S. would be required to spend on varieties exported by country c in categories $z \in I_s^{cd}$ in order to attain utility level U when import prices of those varieties are $\mathbf{p}_s^i$, if this consumer is constrained to consume quantities $\mathbf{q}_s^{-c}$ of all other products, and the number of varieties, quality, and product shifters are, respectively, $\mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}$. The conditional expenditure function solves the problem

$$\min_{q_s^c} \ \mathbf{p}_s^i \mathbf{q}_s^c \qquad s.t. \quad U(\mathbf{q}_s^c, \mathbf{q}_s^{-c}, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}) = U, \qquad i = c, d \tag{19}$$

where $U(.)$ is the representative consumer utility function.[15]

By revealed preference, the minimum import expenditure on products produced by country c in categories $z \in I_s^{cd}$, when import prices of those products are $\mathbf{p}_s^c$ and $\mathbf{q}_s^{-c}, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}$, and U take their equilibrium values, is the observed amount of imports:

$$m_{s(cd)}^c(\mathbf{p}_s^c, \mathbf{q}_s^{-c}, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U) = \mathbf{p}_s^c \mathbf{q}_s^c. \tag{20}$$

However, when prices are $\mathbf{p}_s^d$ instead of $\mathbf{p}_s^c$, the minimum import expenditure is equal to or lower than $\mathbf{p}_s^d \mathbf{q}_s^c$, because the amount $\mathbf{p}_s^d \mathbf{q}_s^c$ is sufficient to attain utility U but $\mathbf{q}_s^c$ is not necessarily optimal given $\mathbf{p}_s^d$. Hence

$$m_{s(cd)}^c(\mathbf{p}_s^d, \mathbf{q}_s^{-c}, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U) \leq \mathbf{p}_s^d \mathbf{q}_s^c. \tag{21}$$

[14] The term in parenthesis in the subindex denotes the subset of products within sector s in which countries c and d export in common to the U.S., i.e. $\{z : z \in I_s^{cd}\}$. As we focus on one country pair at a time, we henceforth omit the parenthesis for notational simplicity, simply denoting these vectors $\mathbf{p}_s^i$, $\mathbf{q}_s^i$, $\mathbf{p}_s^{-i}$, and $\mathbf{q}_s^{-i}$.

[15] The conditional expenditure function is a variation of Neary's (1980) "constrained expenditure function", which he uses to analyze consumption choices under rationing in a competitive context. In contrast to the latter, the conditional expenditure function that we define here only takes into account expenditures on the unconstrained goods. In addition, it also depends on the number of varieties for each country-product, as well as on quality levels and the product-specific shifters.

Taking the ratio of (20) over (21), we obtain

$$M^c_{s(cd)} = \frac{m^c_{s(cd)}(\mathbf{p}^c_s, \mathbf{q}^{-c}_s, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U)}{m^c_{s(cd)}(\mathbf{p}^d_s, \mathbf{q}^{-c}_s, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U)} \geq \frac{\mathbf{p}^c_s \mathbf{q}^c_s}{\mathbf{p}^d_s \mathbf{q}^c_s} = H^{cd}_s. \tag{22}$$

The left hand side of (22), $M^c_{s(cd)}$, captures the change in minimum expenditure on country c's varieties (in categories $z \in I^{cd}_s$) that would be necessary to maintain utility U, if import prices of those varieties changed from $\mathbf{p}^d_s$ to $\mathbf{p}^c_s$, holding constant their number and characteristics (including quality), and the number, characteristics and quantity consumed of all other goods. The right hand side of (22), H^{cd}_s, is a Paasche price index defined over the observed prices of the country pair's common exports to the US in sector s.

Similarly, we can focus on imports from country d to obtain

$$M^d_{s(cd)} = \frac{m^d_{s(cd)}(\mathbf{p}^c_s, \mathbf{q}^{-d}_s, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U)}{m^d_{s(cd)}(\mathbf{p}^d_s, \mathbf{q}^{-d}_s, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U)} \leq \frac{\mathbf{p}^c_s \mathbf{q}^d_s}{\mathbf{p}^d_s \mathbf{q}^d_s} = L^{cd}_s, \tag{23}$$

where L^{cd}_s is a Laspeyres price index defined over the country pair's common exports to the US in sector s.[16]

Given that the Cobb-Douglas form assumed for the upper tier of the utility function is separable into sectoral CES subutility indexes u_s, the constraint in problem (19) can be rewritten as

$$U(\mathbf{q}^c_s, \mathbf{q}^{-c}_s, \mathbf{n}, \boldsymbol{\lambda}, \boldsymbol{\xi}) = \prod_{s'} u_{s'}^{b_{s'}} = U. \tag{24}$$

The value of all subutility indexes for sectors other than s are constant, as their arguments are held constant in problem (19). Therefore, constraint (24) determines the minimum value of u_s that is required to attain utility U, conditional on the (fixed) value of the subutility indexes for the other sectors:

$$u_s = \frac{U}{\prod_{s' \neq s} \overline{u}_{s'}^{b'_s}} \tag{25}$$

Since we focus on expenditure (imports) only on varieties produced by country c in categories $z \in I^{cd}_s$, it is convenient to rewrite the subutility function for sector s as

$$u_s = \left[\sum_{z \in I^{cd}_s} n^c_z \left(\xi_z \lambda^c_s x^c_z \right)^{\varphi_s} + \widehat{u}_s \right]^{1/\varphi_s} \tag{26}$$

[16] Paasche and Laspeyres indexes are typically defined in a time series context, where there is a natural ordering of time periods. Since there is no natural ordering of countries in a multilateral context, calling one of these indexes Paasche and the other one Laspeyres or *vice versa* is arbitrary.

where

$$\widehat{u}_s = \sum_{z \notin I_s^{cd}} n_z^c \left(\xi_z \lambda_s^c x_z^c\right)^{\varphi_s} + \sum_{k \neq c} \sum_{z \in I_s} n_z^k \left(\xi_z \lambda_s^k x_z^k\right)^{\varphi_s}.$$

The first term on the right-hand side of this expression represents the utility from categories imported from country c in sector s that are not also imported from country d. The second term captures the utility from goods imported from all other countries (including d) in any category in sector s. Substituting (26) into (25), and after some algebra, we obtain

$$\left[\sum_{z \in I_s^{cd}} n_z^c \left(\xi_z \lambda_s^c x_z^c\right)^{\varphi_s}\right]^{1/\varphi_s} = \left[\left(\frac{U}{\prod_{s' \neq s} u_{s'}^{b'_s}}\right)^{\varphi_s} - \widehat{u}_s\right]^{1/\varphi_s} \equiv u_s^*.$$

Then, we can rewrite the problem in equation (19) that defines the constrained expenditure function as

$$\min_{x_z^c} \quad \sum_{z \in I_s^{cd}} n_z^c p_z^i x_z^c \quad s.t \quad \left[\sum_{z \in I_s^{cd}} n_z^c \left(\xi_z \lambda_s^c x_z^c\right)^{\varphi_s}\right]^{1/\varphi_s} = u_s^*, \quad i = c, d.$$

The solution to this problem is the product between a CES aggregator measuring the unit cost of utility and the target level of utility, u_s^*[17]

$$m_{s(cd)}^c(\mathbf{p}_s^i, \mathbf{q}_s^{-c}, \boldsymbol{\lambda}, \boldsymbol{\xi}, U) = \left[\sum_{z \in I_s^{cd}} n_z^c \left(\widetilde{p}_z^i \frac{\lambda_s^i}{\lambda_s^c}\right)^{1-\sigma_s}\right]^{\frac{1}{1-\sigma_s}} u_s^*. \tag{27}$$

We can now obtain an explicit expression for $M_{s(cd)}^c$ in equation (22):

$$M_{s(cd)}^c = \left[\frac{\sum_{z \in I_s^{cd}} n_z^c \left(\widetilde{p}_z^c\right)^{1-\sigma_s}}{\sum_{z \in I_s^{cd}} n_z^c \left(\widetilde{p}_z^d \frac{\lambda_s^d}{\lambda_s^c}\right)^{1-\sigma_s}}\right]^{\frac{1}{1-\sigma_s}} = \widetilde{P}_s^{cd} \lambda_s^{cd} \left[\frac{\sum_{z \in I_s^{cd}} n_z^c \left(\frac{\widetilde{p}_z^c}{\widetilde{P}_s^c}\right)^{1-\sigma_s}}{\sum_{z \in I_s^{cd}} n_z^c \left(\frac{\widetilde{p}_z^d}{\widetilde{P}_s^d}\right)^{1-\sigma_s}}\right]^{\frac{1}{1-\sigma_s}} \tag{28}$$

Taking logarithms on both sides of (28) and using the fact that $P_s^{cd} = \widetilde{P}_s^{cd} \lambda_s^{cd}$, we can combine this equation with (22) to obtain

$$\ln H_s^{cd} \leq \ln M_{s(cd)}^c = \ln P_s^{cd} + \ln \phi_s^c, \qquad \phi_s^c = \left[\frac{\sum_{z \in I_s^{cd}} n_z^c \left(\frac{\widetilde{p}_z^c}{\widetilde{P}_s^c}\right)^{1-\sigma_s}}{\sum_{z \in I_s^{cd}} n_z^c \left(\frac{\widetilde{p}_z^d}{\widetilde{P}_s^d}\right)^{1-\sigma_s}}\right]^{\frac{1}{1-\sigma_s}}. \tag{29}$$

[17] It is here where Assumptions 1 and 2 are critical. In equation (27) we use these assumptions to derive $\frac{p_z^i}{\lambda_z^c \xi_z^c} = \frac{p_z^i}{\lambda_z^i \xi_z^i} \frac{\lambda_z^i \xi_z^i}{\lambda_z^c \xi_z^c} = \widetilde{p}_z^i \frac{\lambda_s^i}{\lambda_s^c}$, $i = c, d$.

Similarly, an expression analogous to (28) can be obtained for $M^d_{s(cd)}$, which combined with (23) yields[18]

$$\ln L^{cd}_s \geq \ln M^d_{s(cd)} = \ln P^{cd}_s + \ln \phi^d_s, \qquad \phi^d_s = \left[\frac{\sum_{z \in I^{cd}_s} n^d_z \left(\frac{\tilde{p}^c_z}{\tilde{P}^c_s} \right)^{1-\sigma_s}}{\sum_{z \in I^{cd}_s} n^d_z \left(\frac{\tilde{p}^d_z}{\tilde{P}^d_s} \right)^{1-\sigma_s}} \right]^{\frac{1}{1-\sigma_s}}. \tag{30}$$

Equations (29) and (30) relate the implications of consumer cost minimization to cross-sectional Paasche and Laspeyres price indexes. Although both $M^c_{s(cd)}$ and $M^d_{s(cd)}$ capture the change in minimum expenditure induced by changes in prices from $\mathbf{p}^c_s$ to $\mathbf{p}^d_s$, their magnitudes are typically different, as the varieties that change prices in each case, in particular the number of them – respectively n^c_z and n^d_z – are different. This implies that ϕ^c_s and ϕ^d_s are also different. The fact that $M^c_{s(cd)}$ is in general not equal to $M^d_{s(cd)}$ highlights an important difference between the results of this section and the well-known result that Paasche and Laspeyres price indexes bound the ideal (cost of utility) price index when preferences are homothetic and the environment is competitive.[19] In that case, the relevant change in cost-of-utility is unambiguously defined. An analogous result would be easy to obtain in our love-of-variety framework if $M^c_{s(cd)} = M^d_{s(cd)}$, but this is not true in general.[20] Under plausible assumptions that we outline below, however, we can show that $\ln \phi^c_s < 0$ and $\ln \phi^d_s > 0$ and therefore that the Impure Price Index is bounded by Paasche and Laspeyres indexes, as

$$\ln H^{cd}_s \leq \ln M^c_{s(cd)} \leq \ln P^{cd}_s \leq \ln M^d_{s(cd)} \leq \ln L^{cd}_s. \tag{31}$$

4.3. Paasche and Laspeyres Bounds on the Impure Price Index

Before describing the main result of this section, we develop additional notation specific to country pairs. For each pair of countries c and d, define the pair's (o–normalized) average number of varieties in product category z:

$$\widehat{n}^{cd}_z = \frac{1}{2} \left(\frac{\overline{n}^o_s}{\overline{n}^c_s} n^c_z + \frac{\overline{n}^o_s}{\overline{n}^c_s} n^d_z \right), \tag{32}$$

[18]Note that all prices (observed and pure) considered up to now in this section are import prices. Since trade costs are assumed constant across product categories within a sector, the indexes $M^c_{s(cd)}, M^d_{s(cd)}, H^{cd}_s, L^{cd}_s$ can be alternatively defined in terms of export prices, if they are appropriately scaled by the factor $\frac{\tau^{cUS}_s}{\tau^{dUS}_s}$. As a result, the inequalities in equations (29) and (30) also hold if the indexes are defined over export prices. We use the latter definition for the indexes in the remainder of the paper.

[19]See, for example, Appendix 1 of Feenstra (2003).

[20]$M^c_{s(cd)}$ and $M^d_{s(cd)}$ would be equal, for example, in the unlikely case that the number of varieties in countries c and d are proportional to one another for every product category.

and the country pair's (o–normalized) "multilateral excess variety" in product z relative to the world average:

$$\widetilde{\widetilde{n}}_z^{cd} = \widehat{n}_z^{cd} - \overline{n}_z. \tag{33}$$

Multilateral excess variety measures the extent to which the average number of varieties in countries c and d is above or below the world average.

Also, for each country $i = c, d$ in the country pair, define i's (o-normalized) "bilateral excess variety" in product z relative to the country-pair average,

$$\widetilde{n}_z^{i,cd} = \frac{\overline{n}_s^o}{\overline{n}_s^i} n_z^i - \widehat{n}_z^{cd}. \tag{34}$$

Bilateral excess variety measures the extent to which the number of varieties in a country is above or below the bilateral average. These measures of excess variety possess three convenient properties:

$$\sum_z \widetilde{n}_z^{i,cd} = 0, \qquad \sum_z \widetilde{\widetilde{n}}_z^{cd} = 0, \qquad \widetilde{n}_z^{c,cd} = -\widetilde{n}_z^{d,cd} \tag{35}$$

The first and second properties indicate that, across product categories within country i, both bilateral and multilateral excess variety sum to zero. The third property reveals that two countries cannot both have positive bilateral excess variety in the same category.

Define the bilateral difference in two countries' pure prices in product category z relative to their countries' pure price aggregator as

$$\Delta\widetilde{p}_z^{cd} = \left(\frac{\widetilde{p}_z^c}{\widetilde{P}_s^c}\right)^{1-\sigma_s} - \left(\frac{\widetilde{p}_z^d}{\widetilde{P}_s^d}\right)^{1-\sigma_s}. \tag{36}$$

A positive $\Delta\widetilde{p}_z^{cd}$ indicates that country c has a lower pure price of z (relative to the price aggregator) than country d. A lower pure price may arise, for example, due to comparative advantage, i.e., variation in exporters' relative production efficiency or factor costs.

Finally, for set of products A, define the sample covariance over that set of products as $cov_A(x, y) = \sum_{z \in A} (x_z - \overline{x})(y_z - \overline{y})$.

We now lay out a set of sufficient conditions for the Impure Price Index to be bounded by observable Paasche and Laspeyres price indexes.

Assumption 4: $cov_{I_s^{cd}}\left(\widetilde{n}_z^{c,cd}, \Delta\widetilde{p}_z^{cd}\right) = cov_{I_s^{cd}}\left(\widetilde{n}_z^{d,cd}, \Delta\widetilde{p}_z^{dc}\right) \geq 0$

Assumption 4 states that country c will tend to have a positive bilateral excess variety (relative to country d) in those products in which it has a lower relative price. This covariance is a bilateral version of the covariance in (12), but in this case we assume that it is positive. The motivation for this assumption is analogous to that supporting Assumption 3.

Assumption 5: $cov_{I_s}\left(\widetilde{\widetilde{n}}_z^{cd}, \Delta\widetilde{p}_z^{cd}\right) = 0$

Assumption 5 imposes the restriction that there is no correlation between the country-pair's multilateral excess variety and bilateral differences in pure relative prices. This assumption is not very strong, as there is no obvious relationship between the country pair's excess variety relative to the world average and relative comparative advantage among countries *within* the pair.

Assumption 6: $\delta_s^{cd} = \delta_s^{dc} = 0$,

$$\delta_s^{cd} = \frac{\sum\limits_{z\in U_s^{cd}} \widetilde{n}_z^{c,cd} \frac{1}{Z_s^{cd}} \sum\limits_{z\in I_s^{cd}} \Delta\widetilde{p}_z^{cd} + \sum\limits_{z\in U_s^{cd}} \widehat{n}_z^{cd} \Delta\widetilde{p}_z^{cd}}{\sum\limits_{z\in I_s^{cd}} n_z^c \left(\frac{\widetilde{p}_z^d}{\widetilde{P}_s^d}\right)^{1-\sigma_s}},$$

δ_s^{dc} is defined analogously.

The magnitude of the terms δ_s^{cd} and δ_s^{dc} depends on the extent to which countries c and d are "similarly active". Assumption 6 requires that these terms are zero. A sufficient condition that implies assumption 6 is that the two countries are active in the same categories. In that case, the numerator in the expression for δ_s^{cd} is zero, as it sums over elements of an empty set, U_s^{cd}. Since the sums in the numerator involve positive and negative terms, it is still possible that the numerator is zero even if U_s^{cd} is non-empty. More generally, δ_s^{cd} and δ_s^{dc} will tend to be smaller (in absolute magnitude) the smaller is the number of mismatched active categories (in the numerator) relative to the number of matched active categories (in the denominator). Also, since $\Delta\widetilde{p}_z^{cd} > 0$ and $\widetilde{n}_z^{c,cd} > 0$ for $z \in I_s^{c,-d}$, and $\Delta\widetilde{p}_z^{cd} < 0$ and $\widetilde{n}_z^{c,cd} < 0$ for $z \in I_s^{d,-c}$, these terms will tend to be smaller the more similar is the number of products in $I_s^{c,-d}$ to the number of products in $I_s^{d,-c}$.

Proposition 2 *Under Assumptions 1, 2, 4, 5, and 6, for any two countries c and d, the (unobservable) Impure Price Index is bounded by the (observable) Paasche and Laspeyres indexes:*

$$\ln H_s^{cd} \leq \ln P_s^{cd} \leq \ln L_s^{cd}$$

Proof. See Appendix B. ■

This finding in combination with the result in Proposition 1 provides the basis of our emprical strategy.

5. Empirical Implementation and Results

In this section we use the results of Propositions 1 and 2 to estimate product quality for the United States' top trading partners. Our estimation strategy proceeds in two stages. In the first stage, based on the results of Section 4, we use data on export unit values and quantities to derive an estimate of each country's Impure Price Index. In the second stage, using the results of Section 3, we use information on countries' net trade and trade costs to extract estimates of product quality from the first-stage results. We begin by describing our data sources and outlining our estimation strategy. We then present quality estimates for all manufacturing products as well as for a subset of these products, manufactured materials.

5.1. Data

The first stage requires product-level export prices for every country. These prices are derived from product-level U.S. import data available from the U.S. Census Bureau and compiled by Feenstra et al. (2002). The database records the customs value of all U.S. imports by source country from 1972 to 2001. Imports are recorded according to thousands of finely detailed seven-digit Tariff System of the United States (TSUSA) categories (1974 to 1988) and ten-digit Harmonized System (HS) categories (1989 to 2001). Our estimates for All Manufacturing include products in SITC aggregates 5 through 8. Our estimates for Manufactured Materials include products in SITC 6.

The U.S. trade data include information on both quantity and value for many goods. We compute the unit value, or "price", of product z from country c, p_z^c, by dividing import value (v_z^c) by import quantity (q_z^c), $p_z^c = v_z^c/q_z^c$.[21] Examples of the units employed to classify products include dozens of shirts in apparel, square meters of carpet in textiles and pounds of folic acid in chemicals.

Product-level trade data are noisy due to both aggregation bias and measurement error.[22] Aggregation bias is minimized by using detailed data, but is likely to remain. We therefore trim the data along two dimensions before using them to compute Paasche and Laspeyres indexes. The first trim involves dropping country-year-product observations with value less than \$10,000 or quantity equal to 1. The second trim eliminates country-pair-year-product observations when the relative quantity or the relative price of the country-pair-product is either below the 2nd percentile or above the 98th percentile of all country-pair-product observations in that year. The first trim gets rid of unusual and unrealistic imports while the second trim discards unreliable country comparisons.

The second stage requires measures of trade balance and trade costs at the sectoral level.

[21] Availability of unit values averages about 80 percent over the years in our sample.
[22] See, for example, GAO (1995) and Schott (2004).

We measure countries' sectoral trade balance relative to GDP by dividing nominal dollar-denominated trade flow data from the World Trade Flows database compiled by Feenstra et al. (2004) with GDP data from the World Bank's World Development Indicators database. For the real exchange rate we rely on version 6.1 of the Penn World Tables.[23]

Ideally, our estimates of trade costs between countries would include measures of transportation costs, tariffs and non-tariff barriers as well as other costs due to language barriers, etc. Here, due to data constraints, we focus on the former.[24] We measure bilateral transport costs using the U.S. trade data. Most records in the U.S. trade data report both the customs-insurance-freight (cif) and free-on-board (fob) value of the import flow. We estimate *ad valorem* transport costs per mile for industry s in year t by regressing the relative value spent on customs, insurance and freight on imports from country c on the distance the exports have travelled,

$$\frac{cif_{st}^{c} - fob_{st}^{c}}{fob_{st}^{c}} = \delta_{st} D^{c,US} + \in_{st}^{c} \tag{37}$$

where $D^{c,US}$ represents the great circle distance in miles between the United States and country c. In our estimations below, we set τ_{st}^{cd} equal to $\widehat{\delta}_{st} D^{cd}$.

We report quality estimates for the top 45 non-OPEC U.S. trading partners for the period 1980 to 1997. This sample was chosen to yield a relatively long and balanced panel. We exclude years prior to 1980 because trade is dominated by a relatively small group of high-income countries. We exclude years after 1997 because of significant outliers in the trade balance data between 1998 and 2001.[25]

5.2. Estimation Strategy

5.2.1. First Stage: Estimation of the Impure Price Index

In the first stage of the estimation strategy, we use the results of Proposition 2 to estimate each country's Impure Price Index, $\widehat{P}_s^{co}$, where country o is the numeraire country.[26] The idea of the identification strategy is as follows. For generic country pair c and d, the estimated indexes $\widehat{P}_s^{co}$ and $\widehat{P}_s^{do}$ implicitly determine a bilateral index $\widehat{P}_s^{cd} = \widehat{P}_s^{co}/\widehat{P}_s^{do}$. This index should satisfy the Paasche and Laspeyres bounds for that country pair, as

[23] A country's real exchange rate is found by dividing its purchasing-power-parity relative to the United States (PPP) into its local currency per U.S. dollar exchange rate (XRAT).

[24] Going forward, our technique will benefit from the ongoing development of trade cost datasets such as TRAINS. We are currently exploring the use of the TRAINS in our estimation, but the sparseness of data in TRAINS prior to the late 1990s severely restricts the sample size of the second stage of our estimation.

[25] We are currently investigating these outliers and plan to extend the analysis to 2001 once they are verified.

[26] The choice of numeraire is made without loss of generality. In the results presented below, Germany (DEU) is the numeraire.

outlined in Proposition 2. Similarly, for C trading partners, the estimation of $C-1$ Impure Price Indexes, $\widehat{P}_s^{co}\ \forall c \neq o$, implicitly determine $C(C-1)$ bilateral indexes, $\widehat{P}_s^{cd}\ \forall c, d$, which should satisfy the bilateral Paasche and Laspeyres price index bounds for all country pairs. If the Paasche and Laspeyres bounds were observed without error, estimation would entail searching for an interior solution to the set of restrictions imposed by the bounds across country pairs. Here, in light of evidence that import prices are in fact mis-recorded on customs documents[27], we instead allow for the possibility that the true Paasche and Laspeyres indexes are observed with error.

Denote the "true" Paasche and Laspeyres indexes by H_s^{*cd} and L_s^{*cd}, respectively. We assume that the observed indexes, H_s^{cd} and L_s^{cd}, vary from the true indexes by a multiplicative error, $\ln H_s^{cd} = \ln H_s^{*cd} + \zeta_{h,s}^{cd}$ and $\ln L_s^{cd} = \ln L_s^{*cd} + \zeta_{l,s}^{cd}$. We also assume that each error is distributed normally, $\zeta_{h,s}^{cd} \sim N(0, \psi/w_s^{cd})$ and $\zeta_{l,s}^{cd} \sim N(0, \psi/w_s^{cd})$, and that the errors for each bound are independent both of each other and of error terms for other bilateral pairs.[28] Note that we weight the standard deviation of the error distribution by w_s^{cd}. In the results below, this weight is set equal to the square root of the number of categories that countries c and d export in common to the United States. This weight is meant to increase the contribution to the likelihood of country pairs with a relatively large number of exports in common.

Satisfying the inequality constraints of Proposition 2 for a given pair of countries implies:

$$\ln P_s^{cd} \geq \ln H_s^{*cd} \Longrightarrow \zeta_{h,s}^{cd} \geq \ln H_s^{cd} - \ln P_s^{cd} \tag{38}$$

$$\ln P_s^{cd} \leq \ln L_s^{*cd} \Longrightarrow \zeta_{l,s}^{cd} \leq \ln L_s^{cd} - \ln P_s^{cd}. \tag{39}$$

We estimate the set of index numbers $\widehat{P}_s^{co}$, $\forall c \neq o$, and the variance parameter $\widehat{\psi}$, for a given year t, by maximizing the likelihood that the "true" Paasche and Laspeyres bounds contain the estimates. The likelihood for a single pair of country bounds is

$$l_s^{cd} = \ln \Phi \left(\frac{\ln\left(\frac{P_s^{co}}{P_s^{do}}\right) - \ln H_s^{cd}}{\psi / w_s^{cd}} \right) + \ln \Phi \left(\frac{\ln L_s^{dc} - \ln\left(\frac{P_s^{co}}{P_s^{do}}\right)}{\psi / w_s^{cd}} \right) \tag{40}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

[27] See U.S. General Accounting Office (1995) for an in-depth study of price variation within eight product categories.

[28] Our assumptions about the normality and independence of the errors represent a potentially strong simplification. Errors across country pairs with one country in common are likely to be correlated as they are constructed using similar information. The within-country-pair Paasche and Laspeyres errors are also likely to be correlated: a high negative Paasche error will coincide with a high positive Laspeyres error. We are currently working on relaxing these assumptions.

5.2.2. Second Stage: Estimation of Product Quality

Variation in estimates of countries' Impure Price Indexes contains information about pure prices and product quality. Proposition 1 demonstrates that countries' pure prices, as summarized by the Pure Price Index, determine their sectoral trade balance. In the second stage, we use the results of that proposition to extract the pure-price component of the Impure Price Index. In particular, we infer quality via estimation of equation (15). Incorporating $\ln \widetilde{P}_s^{cd} = \ln P_s^{cd} - \ln \lambda_s^{cd}$ from equation (18), and neglecting the error arising from the linear approximation described in the proof of Proposition 1, we can rewrite equation (15) as

$$\left[\frac{T_{st}^c}{Y_t^c} - \frac{T_{st}^o}{Y_t^o}\right] = \gamma_s \ln \widehat{P}_{st}^{co} - \gamma_s \ln \lambda_{st}^{co} + \gamma_s \mu_s \tau_{st}^{c,IN} - \gamma_s \mu_s \tau_{st}^{c,OUT} + \gamma_s \kappa_{st}^{co} - \gamma_s \theta_{st}^{co} \tag{41}$$

where $\tau_{st}^{co,IN} = \tau_{st}^{\bar{c}c} - \tau_{st}^{\bar{c}o}$ and $\tau_{st}^{c,OUT} = \tau_{st}^{c\bar{c}} - \tau_{st}^{o\bar{c}}$ are the inbound and outbound trade costs, respectively, $\kappa_s^{co} = \ln P_s^{co} - \ln \widehat{P}_s^{co}$ is the estimation error in the first-stage estimates, and subscript t indexes time periods. Equation (41) highlights the fact that countries' unobserved product quality relative to the numeraire country (λ_{st}^{co}) is part of a compound error term that also includes the estimation error in the first stage (κ_{st}^{co}) and the idiosyncratic components of the covariance between excess variety and pure prices (θ_{st}^{co}) from equation (12). We assume that both κ_{st}^{co} and θ_{st}^{co} are uncorrelated with $\widehat{P}_s^{co}$. However, the assumption that the quality component of the error term ($\ln \lambda_{st}^{co}$) is uncorrelated with the regressor $\ln \widehat{P}_{st}^{co}$ is untenable. Developed countries, which tend to have higher export prices, are also likely to produce higher quality. (This presumption is confirmed later by our results.)

To deal with this endogeneity problem, we first specify a time path for the evolution of product quality relative to the base country:

$$\ln \lambda_{st}^{co} = \alpha_{0s}^{co} + \alpha_{1s}^{co} t + \varepsilon_{st}^{co} \tag{42}$$

where α_0^{co} and α_1^{co} are a country fixed effect and the slope of a country-specific time trend, respectively, and ε_{st}^{co} represents deviations of quality from this trend. Incorporating this (country-specific) linear trend for quality into equation (41), we obtain

$$\left[\frac{T_{st}^c}{Y_t^c} - \frac{T_{st}^o}{Y_t^o}\right] = \gamma_s \ln \widehat{P}_{st}^{co} - \gamma_s \left(\alpha_{0s}^{co} + \alpha_{1s}^{co} t\right) + \gamma_s \mu_s \tau_{st}^{c,IN} - \gamma_s \mu_s \tau_{st}^{c,OUT} + \gamma_s \upsilon_{st}^{co} \tag{43}$$

where $\upsilon_{st}^{co} = \varepsilon_{st}^{co} + \kappa_{st}^{co} + \theta_{st}^{co}$. This equation can be more transparently written as

$$\left[\frac{T_{st}^c}{Y_t^c} - \frac{T_{st}^o}{Y_t^o}\right] = \gamma_s \ln \widehat{P}_{st}^{co} + \beta_{0s}^{co} + \beta_{1s}^{co} t + \beta_{2s} \tau_{st}^{c,IN} + \beta_{3s} \tau_{st}^{c,OUT} + \iota_{st}^{co} \tag{44}$$

where

$$\begin{aligned} \alpha_{0s}^{co} &= -\beta_{0s}^{co}/\gamma_s \\ \alpha_{1s}^{co} &= -\beta_{1s}^{co}/\gamma_s \\ \mu_s &= \beta_{2s}/\gamma_s = -\beta_{3s}/\gamma_s \\ \upsilon_{st}^{co} &= \iota_{st}^{co}/\gamma_s \end{aligned} \quad (45)$$

The inclusion of country fixed effects in (44) eliminates the most obvious source of endogeneity, i.e. the cross-sectional correlation between the time-invariant components of countries' prices and quality levels. The inclusion of country-specific time trends further reduces the remaining correlation between regressor and error term, as the latter term now only includes deviations of quality from country-specific trends. However, correlation between ε_{st}^{co} and $\widehat{P}_{st}^{co}$ may still persist, as shocks to quality are likely to be accompanied by increases in prices. To address this potential endogeneity problem, we use the real exchange rate as an instrument for $\widehat{P}_{st}^{co}$. As usual, the instrument needs to satisfy two conditions. First, since the estimating equation includes country-specific fixed effects and time trends, the instrument has to be (partially) correlated with $\widehat{P}_{st}^{co}$, *after controlling for the fixed effects and time trends*. In other words, deviations of the real exchange from its own time trend have to be correlated with similar deviations of $\widehat{P}_{st}^{co}$. Macroeconomic conditions typically determine periods of over- and under-valuation of countries' real exchange rate around long-run trends. These periods also determine changes in the international competitiveness of a countries' exports, captured in our model by $\widetilde{P}_{st}^{co}$. Since $\widetilde{P}_{st}^{co}$ is a component of $\widehat{P}_{st}^{co}$, periods of over- or under-valuation will also be associated with movements of $\widehat{P}_{st}^{co}$. Second, the instrument has to be uncorrelated with the error term ε_{st}^{co}, which requires that shocks to quality around the trend in sector s are not correlated with the real exchange rate. While we cannot rule out that such a correlation exists, we think that it is unlikely to be important. Shocks to quality in sector s might be accompanied by exactly offsetting increases in prices, leaving pure prices – and hence net trade in that sector – unchanged. Even if these shocks affect pure prices, they might have a negligible effect on the real exchange rate. This is more likely to be true if the shocks are temporary deviations around a trend, and if they are specific to sector s, i.e. not correlated with shocks to quality in other sectors.

We estimate equation (44) in first differences using two-stage least squares.[29] Our estimation of countries' trend in export quality over the sample period is

$$\ln \widehat{\lambda}_{st}^{co} = \widehat{\alpha}_{0s}^{co} + \widehat{\alpha}_{1s}^{co} t \quad (46)$$

[29] We report results for first differences because residuals in levels are autocorrelated while there is no evidence of autocorrelation of residuals in first differences. In any case, estimation in levels or in second differences yields similar results.

where t indexes years starting in 1980, $\widehat{\alpha}_{1s}^{co} = -\widehat{\beta}_{1s}^{co}/\widehat{\gamma}_s$, and $\widehat{\alpha}_{0s}^{co}$ is recovered using the parameter estimates.[30] Note that we can only identify the linear trend in quality. Deviations of quality from the trend are confounded with the other two components of the error term. For comparison, in some of our results we also report

$$\ln \widehat{\lambda}_{st}^{co} + \widehat{\upsilon}_{st}^{co} = \widehat{\alpha}_{0s}^{co} + \widehat{\alpha}_{1s}^{co} t + \widehat{\upsilon}_{st}^{co} \quad (47)$$

which is the estimate of quality plus the compund error term from equation (43).

5.3. *Estimation Results*

In this section we report preliminary estimates of export quality. We begin by examining product quality in All Manufacturing. While we intend for our methodology to be applied primarily to more disaggregate sectors, we prefer to start with a relatively aggregate sector in order to focus on the fundamental aspects of the methodology while abstracting from sector-specific nuances we discuss below. Examination of aggregate manufacturing is also useful for assessing how our priors about manufacturing prowess compare to our methodology's estimates of quality. After analyzing manufacturing as a whole, we turn to a subset of that sector, manufactured materials (SITC 6).

5.3.1. *All Manufacturing*

Table 1 summarizes the results of the first stage maximum likelihood results. The log likelihood and point estimates for $\widehat{\psi}$ are reported for each year. The log likelihood declines with time while $\widehat{\psi}$ is relatively constant.[31] Table 1 also displays the average number of products country pairs export in common to the United States each year. The maximum number of products in common ranges from 2302 in 1980 to 5107 in 1997.

Figure 1 reports first stage estimates of $\ln \widehat{P}_{st}^{co}$ for each trading partner, along with their associated 95 percent confidence intervals for four years, 1980, 1985, 1990 and 1997. These results, at roughly five-year intervals, are representative of the variation observed more generally in the estimates over time, and we report just these four years to conserve space. The horizontal axis sorts countries from low to high while the vertical axis reports each country's $\ln \widehat{P}_{st}^{co}$ relative to Germany, where $\ln P_{st}^{DEU} = 0$. The ordering of countries

[30] The recovered country fixed effect $\widehat{\alpha}_{0s}^{co}$ is equal to

$$\overline{\left[\frac{T_{st}^c}{Y_t^c} - \frac{T_{st}^o}{Y_t^o}\right]} - \widehat{\gamma}_s \overline{\widehat{P}_{st}^{co}} - \widehat{\alpha}_{1s}^{co}\overline{t} - \widehat{\beta}_{2s}\overline{\tau_s^{c,IN}} - \widehat{\beta}_{3s}\overline{\tau_s^{c,OUT}},$$

where a bar over a variable denotes the average for each country over the sample period.

[31] Recall that observations are weighted by $1/w_s^{cd}$, the inverse of the square root of the number of products in which countries c and d are active. As indicated in Table 1, this overlap – as well as the number of product categories – increases with time.

accords with level of development, with higher-income countries generally having relatively higher Impure Price Indexes than lower-income developing countries. Bangladesh (BGD) and Pakistan (PAK) are among the lowest ranked countries in every year, while Switzerland (CHE) is among the highest.

Figure 2 provides an alternate view of the first-stage results by tracing four countries' $\ln \widehat{P}_{st}^{co}$ across the entire sample period. The four countries are Argentina (ARG), China (CHN), Ireland (IRL), and Malaysia (MYS). Estimated Impure Price Indexes are relatively high and increasing in the 1990s for Ireland and relatively low and increasing in the 1990s for China and Malaysia. All countries exhibit important movements in their $\ln \widehat{P}_{st}^{co}$ over time. These movements might be induced by various factors, including movement in the base country's impure prices. As is clear from our discussion in the previous section, a generally rising or falling trend in $\ln \widehat{P}_{st}^{co}$ over time cannot be attributed to quality until variation in pure prices $\ln \widetilde{P}_{st}^{co}$ is netted out.

Table 2 reports the second-stage 2SLS estimates of γ_s from first differencing equation (44), where the real exchange rate is used to instrument for the first-stage estimates of the Impure Price Indexes and transport costs are used as estimates of more general trade costs. Outbound trade costs are weighted by the trading partners' share of world GDP. Inbound trade costs are weighted by trading partners' share of world exports in industry s.[32] Three sets of coefficients are reported, accompanied by standard errors that are robust to heteroskedasticity and are clustered at the country level. The first column reports results for OLS, while the second and third columns report results for 2SLS excluding and including trade costs, respectively. The OLS estimate for γ_s, while negative, is close to zero and statistically insignificant. The 2SLS estimates of γ_s are both negative and significant and slightly higher for the specification that controls for trade costs. A negative estimate for γ_s implies that positive sectoral trade balances adjust the Paasche and Laspeyres indexes upwards, contributing toward higher assessments of relative quality. Estimates of the coefficients on outbound and inbound tariffs have the expected sign but are of marginal significance; as indicated in the second to last row of the table, the null hypothesis that they sum to zero, implied by the model, cannot be rejected. Finally, the first stage F statistics in the final row of the table, at 102 and 79 for the two 2SLS estimations, respectively, support our choice of instrument.[33]

[32] In the construction of trade cost variables, we omit pre-multiplication by g^c in equation (13) and by $(1 - r_s^c)$ in equation (14). These adjustment factors are, to a large extent, an artifact of the linear approximation (around a free-trade equilibrium), and strongly over-predict the importance of size as a determinant of the impact of trade costs on trade volumes away from that equilibrium. Our examination of this issue is ongoing.

[33] Results for the first-stage of the 2SLS estimation are omitted to conserve space. Note that we cannot perform a test of over-identifying restrictions, such as the Sargan test, because we have only a single instrument.

Figure 3 reports the second-stage estimates implied by the third column of our 2SLS results for the four countries examined earlier.[34] Each panel compares the first-stage Impure Price Index estimate to second-stage quality estimates that both exclude (equation 46) and include (equation 47) the error term. All four countries exhibit gains in quality over time, though the upward trend for Ireland, for example, is sharper than that of Argentina. Both China and Ireland exhibit a crossing of their first and second stage estimates that coincides with their running a trade surplus in manufacturing: a high quality-to-impure-price-index ratio implies relatively low pure prices that increase the attractiveness of the countries' goods on world markets and drives up the trade surplus.

Figure 4 compares the first and second-stage estimates for all countries in 1997, the final year of the sample period. The figure contains two panels: the top panel sorts countries by their first-stage estimate while the lower panel sorts countries according to their second-stage estimates. As indicated in the bottom panel, Ireland and Finland (FIN) are estimated to have the highest manufacturing export quality while Guatemala (GTM) and El Salvador (SLV) have the lowest.

Comparison of second- versus first-stage estimates reveals that China, Taiwan (TWN) and Ireland experience substantial upward adjustments while Guatemala and El Salvador exhibit two of the biggest downward adjustments. These "corrections" of the Impure Price Index highlight the intuition of our approach. China, for example, runs a relatively large manufacturing trade surplus in 1997 relative to other countries. This surplus indicates strong demand for Chinese goods, which means that Chinese pure prices must be relatively low, and hence its quality must be high relative to the observed prices. Note that the validity of this inference does not depend on why pure prices are low: China's low pure prices could be due to a comparative advantage in manufacturing as well as to an under-valuation of its currency. El Salvador, on the other hand, runs a substantial trade deficit. As a result, its pure prices must be relatively high and consequently its export quality must be relatively low. Countries with relatively balanced manufacturing trade, such as Italy (ITA) have second-stage quality indexes roughly equal to their first-stage impure price indexes.

5.3.2. Manufactured Materials

As noted above, implementation of our methodology on more disaggregate industries has the virtue of greater consistency with our assumptions of constant quality and elasticity of substitution for product categories in the same industry. In this section, we provide preliminary results for a sub-industry of manufacturing, manufactured materials (SITC 6),

[34] Displays of second-stage results both with and excluding tariffs reveals that virtually all of the adjustment occurs through γ_s rather than μ_s. As a result, we do not provide a decomposition of these adjustments in our reported results. We are currently investigating whether this result persists when richer measures of trade costs that include tariff and non-tariff barriers are employed in the second stage.

which includes leather goods, textiles, paper products and steel. We discuss the results for this industry relatively briefly, placing emphasis on how they, in comparison with results for all of manufacturing, provide intuition for our approach.

First- and second-stage maximum likelihood and 2SLS results for SITC 6, respectively, are reported in Tables 3 and 4. As above, estimates of γ_s are negative, but they are lower in absolute magnitude than the corresponding estimates for all manufacturing. These results highlight the fact that the magnitude of the estimated coefficients should be related to the size of the sector (as the sectoral trade balance is normalized by total GDP). Coefficients on inbound and outbound trade costs have the expected sign and have similar magnitude to one another. This magnitude, as expected, is smaller in absolute value than for All Manufacturing.

Figures 5 and 6 display the first-stage estimates for SITC 6. *Vis a vis* All Manufacturing, the impure price indexes generally have larger standard errors and, for the four highlighted countries in Figure 6, rise somewhat less with time. In China's case, impure prices indexes decline between 1980 and 1997. These features of the first-stage results carry over to the second-stage quality index estimates displayed in Figures 7 and 8. The upper-right panel of Figure 7, for example, indicates a moderate decline in the quality of Chinese manufactured materials over the sample period. In addition, because China and Taiwan do not run a very large surplus in Manufactured Materials, the wedge between their first- and second-stage estimates is lower here than in All Manufacturing. As a result, their quality ranking in 1997 is lower in Figure 7 than in Figure 3. Finland and Pakistan, on the other hand, run relatively large trade surpluses in SITC 6. Consequently, they are estimated to have relatively high-quality exports compared with manufacturing as a whole. Some countries run a trade deficit in SITC 6 mainly as a result of running a trade deficit in SITC 65 (defined to be "textile yarn, fabrics, made-up articles, not elsewhere specified, and related products"). Many of these countries also run a surplus in SITC 84 ("articles of apparel and clothing accessories"), which suggests that imports in SITC 65 are used as intermediate inputs for re-export as SITC 84. In these cases, it appears more appropriate to aggregate industries with strong input-output relationships. The identification of these industries as well as the proper method for such aggregation is subject to ongoing research.

6. Conclusion

A vast literature associates cross-country variation in export unit-values with variation in product quality. This paper calls attention to the strength of the assumptions necessary to justify this approach and develops a methodology for exploiting information on countries' sectoral trade balance to identify pure price variation not associated with quality.

References

Abed-el-Rahman, K., 1991. Firms' Competitive and National Comparative Advantages as Joint Determinants of Trade Composition. Weltwirtschaftliches Archiv, 127:83-97.

Aiginger, Karl, 1998. Unit Values to Signal the Quality Position of CEECs. In The Competitiveness of Transition Economies. OECD Proceedings, 1998(10):1-234.

Aiginger, Karl, 1997. The Use of Unit Values to Discriminate Between Price and Quality Competition. Cambridge Journal of Economics, 21(5):571-592.

Anderson, James and Eric van Wincoop, 2003. Gravity with Gravitas: A Solution to the Border Puzzle, American Economic Review, 93:170-192.

Anderson, James and Eric van Wincoop, 2004. Trade Costs. Journal of Economic Literature, XLII(3): 691-751.

Aturupane, Chonira, Simeon Djankov and Bernard Hoekman, 1999. Horizontal and Vertical Intra-Industry Trade Between Eastern Europe and the EU. Weltwirtschaftliches Archiv, 135(1):62-81.

Aw, Bee Yan and Mark J. Roberts, 1986. Measuring Quality Change in Quota-Constrained Import Markets. Journal of International Economics, 21(1): 45-60.

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

Bernard, Andrew B., Stephen Redding and Peter K. Schott, 2004. Heterogenous Firms and Comparative Advantage. NBER Working Paper 10668.

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

Feenstra, Robert C., 1988. Quality Change Under Trade Restraints in Japanese Autos. Quarterly Journal of Economics, 103:131-146.

Feenstra, Robert C., 1995. Exact Hedonic Price Indexes. Review of Economics and Statistics, 77:634-653.

Feenstra, Robert C., 2004. Advanced International Trade, Princeton University Press, Princeton, NJ.

Feenstra, Robert C., John Romalis and Peter K. Schott, 2002. U.S. Imports, Exports, and Tariff Data, 1989-2001, NBER Working Paper 9387.

Feenstra, Robert C, Alan Heston, Marcel Timmer and Haiyan Deng, 2004. Estimating Real Production and Expenditures Across Nations: A Proposal for Improving Existing Practice. UC Davis, mimeo.

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

Grossman, Gene and Elhanan Helpman, 1991. Quality Ladders and Product Cycles. Quarterly Journal of Economics, 106(2): 557-586.

Hallak, Juan C., 2005. Product Quality and the Direction of Trade. Forthcoming in Journal of International Economics.

Helpman, Elhanan and Paul Krugman, 1985. Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy, MIT Press, Cambridge, MA.

Hummels, David and Peter Klenow, 2005. The Variety and Quality of a Nation's Exports. American Economic Review, 95: 704-723.

Ianchovichina, Elena, Sethaput Suthiwart-Narueput and Min Zhao, 2003. Regional Impact of China's WTO Accession. In Krumm, Kathie and Homi Kharas (eds), East Asia Integrates: A Trade Policy Agenda for Shared Growth (Washington DC: The International Bank for Reconstruction and Development / The World Bank).

Krugman, Paul R., 1979. Increasing Returns, Monopolistic Competition, and International Trade. Journal of International Economics IX:469-479.

Krugman, Paul R., 1980. Scale Economies, Product Differentiation and the Pattern of Trade. American Economic Review LXX:950-959.

Krugman, Paul R., 1989. Differences in Income Elasticities and Trends in Real Exchange Rates. European Economic Review 33:1055-85.

Neary, J. Peter, 1980. The Theory of Household Behaviour Under Rationing. European Economic Review 13:25-42.

Romalis, John, 2004. Factor Proportions and the Structure of Commodity Trade. American Economic Review, 94: 67-97.

Schott, Peter K, 2003. One Size Fits All? Heckscher-Ohlin Specialization in Global Production. American Economic Review, 93: 686-708.

Schott, Peter K, 2004. Across-Product versus Within-Product Specialization in International Trade. Quarterly Journal of Economics, 119(2): 647-678.

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

Verma, Samar, 2002. Export Competitiveness of Indian Textile and Garment Industry. Indian Council for Research on International Economic Relations, Working Paper 94.

A Proof of Propostition 1

We start by reproducing equation (6):

$$\frac{1}{b_s}\left[\frac{T_s^c}{Y^c}-\frac{T_s^o}{Y^o}\right]=\sum_{c'}\sum_z\left[\frac{n_z^c\left(\widetilde{p}_z^c\tau_s^{cc'}\right)^{1-\sigma_s}}{(G_s^{c'})^{1-\sigma_s}}\frac{Y^{c'}}{Y^c}-\frac{n_z^o\left(\widetilde{p}_z^o\tau_s^{oc'}\right)^{1-\sigma_s}}{(G_s^{c'})^{1-\sigma_s}}\frac{Y^{c'}}{Y^o}\right] \tag{48}$$

Solving for n_z^c and n_z^o using (9), and substituting into equation (48), we can rewrite the right hand side of the latter equation as

$$\left[\begin{array}{c}\frac{\overline{n}^c}{\overline{n}^o}\frac{1}{Y^c}\left(\sum_{c'}\frac{Y^{c'}}{(G_s^{c'})^{1-\sigma_s}}(\tau_s^{cc'})^{1-\sigma_s}\right)\left(\sum_z\overline{n}_z(\widetilde{p}_z^c)^{1-\sigma_s}+\left(\widetilde{P}_s^c\right)^{1-\sigma_s}\sum_z\widetilde{n}_z^c\left(\frac{\widetilde{p}_z^c}{(\widetilde{P}_s^c)^{1-\sigma_s}}\right)^{1-\sigma_s}\right)-\\ -\frac{1}{Y^o}\left(\sum_{c'}\frac{Y^{c'}}{(G_s^{c'})^{1-\sigma_s}}(\tau_s^{oc'})^{1-\sigma_s}\right)\left(\sum_z\overline{n}_z(\widetilde{p}_z^o)^{1-\sigma_s}+\left(\widetilde{P}_s^o\right)^{1-\sigma_s}\sum_z\widetilde{n}_z^o\left(\frac{\widetilde{p}_z^o}{(\widetilde{P}_s^o)^{1-\sigma_s}}\right)^{1-\sigma_s}\right)\end{array}\right]$$

Using the definition of $\widetilde{P}_s^c$ in (10) and the fact that, since $\sum_z\widetilde{n}_z^c=0$, $\sum_z\widetilde{n}_z^c(\widetilde{p}_z^c)^{1-\sigma_s}=Z_s cov\left[\widetilde{n}_z^c,(\widetilde{p}_z^c)^{1-\sigma_s}\right]$, this expression can be further rewritten as

$$\frac{\left(\widetilde{P}_s^o\right)^{1-\sigma_s}}{Y^o}\left[\begin{array}{c}\frac{\overline{n}^c}{\overline{n}^o}\frac{Y^o}{Y^c}\left(\frac{\widetilde{P}_s^c}{\widetilde{P}_s^o}\right)^{1-\sigma_s}\left(\sum_{c'}\frac{Y^{c'}}{(G_s^{c'})^{1-\sigma_s}}(\tau_s^{cc'})^{1-\sigma_s}\right)\left(1+Z_s cov\left[\widetilde{n}_z^c,(\widetilde{p}_z^c)^{1-\sigma_s}\right]\right)-\\ -\left(\sum_{c'}\frac{Y^{c'}}{(G_s^{c'})^{1-\sigma_s}}(\tau_s^{oc'})^{1-\sigma_s}\right)\left(1+Z_s cov\left[\widetilde{n}_z^o,(\widetilde{p}_z^o)^{1-\sigma_s}\right]\right)\end{array}\right]$$

Using Assumption 3 and equation (12), we can substitute the latter expression for the right hand side of (48):

$$\frac{1}{b_s}\left[\frac{T_s^c}{Y^c}-\frac{T_s^o}{Y^o}\right]=\frac{\left(\widetilde{P}_s^o\right)^{1-\sigma_s}}{Y^o}\left[\left(\widetilde{P}_s^{co}\right)^{1-\eta_s-\sigma_s}\left(\sum_{c'}\frac{Y^{c'}}{(G_s^{c'})^{1-\sigma_s}}(\tau_s^{cc'})^{1-\sigma_s}\right)\left[1+Z_s\left(V_s+\theta_s^c\right)\right]-\Psi_s^o\right] \tag{49}$$

where $\Psi_s^o=\left(\sum_{c'}\frac{Y^{c'}}{(G_s^{c'})^{1-\sigma_s}}(\tau_s^{oc'})^{1-\sigma_s}\right)\left[1+Z_s\left(V_s+\theta_s^o\right)\right]$.

We can divide both sides of (49) by $\frac{\Psi_s^o\left(\widetilde{P}_s^o\right)^{1-\sigma_s}}{Y^o}$, sum 1 to both sides, and then take logarithms, to obtain

$$\begin{aligned}\ln\left\{1+\frac{1}{\gamma_s'}\left[\frac{T_s^c}{Y^c}-\frac{T_s^o}{Y^o}\right]\right\}&=\ln\left(\widetilde{P}_s^{co}\right)^{1-\eta_s-\sigma_s}+\\&+\ln\left(\sum_{c'}\frac{Y^{c'}}{(G_s^{c'})^{1-\sigma_s}}(\tau_s^{cc'})^{1-\sigma_s}\right)+\ln\left[1+Z_s\left(V_s+\theta_s^c\right)\right]-\ln(\Psi_s^o)\end{aligned} \tag{50}$$

where $\gamma_s' = \frac{b_s \Psi_s^o \left(\widetilde{P}_s^o\right)^{1-\sigma_s}}{Y^o}$. Using $\ln(1+x) \simeq x$, and abstracting from the approximation error, we can express equation (50) as

$$\frac{1}{\gamma_s'}\left[\frac{T_s^c}{Y^c} - \frac{T_s^o}{Y^o}\right] = \left[\begin{array}{c} (1-\eta_s-\sigma_s)\ln \widetilde{P}_s^{co} + \ln\left(\sum_{c'} \frac{Y^{c'}}{\left(G_s^{c'}\right)^{1-\sigma_s}} (\tau_s^{cc'})^{1-\sigma_s}\right) + Z_s\left(V_s + \theta_s^c\right) - \\ -\ln\left(\sum_{c'} \frac{Y^{c'}}{\left(G_s^{c'}\right)^{1-\sigma_s}} (\tau_s^{oc'})^{1-\sigma_s}\right) - Z_s\left(V_s + \theta_s^o\right) \end{array}\right] \quad (51)$$

We will perform a first-order Taylor expansion of the logarithm terms in equation (51). We describe here the expansion of the first of these terms (for country c). The case for country o is analogous. Using the definition of $G_s^{c'}$ in (3), we can rewrite this term as

$$\ln\left[\frac{Y^c}{\sum_{c'}\sum_z n_z^{c'}\left(\widetilde{p}_z^{c'}\tau_s^{c'c}\right)^{1-\sigma_s}} + \sum_{c'\neq c}\left(\frac{Y^{c'}}{\sum_{c''}\sum_z n_z^{c''}\left(\widetilde{p}_z^{c''}\tau_s^{c''c'}\right)^{1-\sigma_s}}(\tau_s^{cc'})^{1-\sigma_s}\right)\right] \quad (52)$$

The first term of this expression is a function of the "inbound" trade costs $\tau_s^{c'c}, c' = 1, ..., C$, the trade costs that c' has to bear to export to country c. The second term is a function of the "outbound" trade costs $\tau_s^{cc'}, c' = 1, ..., C$, the trade costs that country c has to bear to export to country c'. We will perform the Taylor expansion around a free-trade equilibrium, i.e. a point at which $\tau_s^{c'c} = \tau_s^{cc'} = 1, \forall c, c'$. Note that under free trade, the term in the denominator is the same for every country, $G_s^{c'} = G_s$. Define $r_s^{c'} = \frac{1}{G_s^{1-\sigma_s}} \sum_z n_z^{c'} \left(\widetilde{p}_z^{c'}\right)^{1-\sigma_s}$ and $g^{c'} = Y^{c'} / \sum_{c'} Y^{c'}$. A first-order Taylor expansion of (52) around the free-trade point results in an expression of the following form:

$$\ln\left[\sum_{c'}\frac{Y^{c\prime}}{G_s^{1-\sigma_s}}\right] + \sum_{c'\neq c}\left(\frac{\partial \ln[.]}{\partial \tau_s^{c'c}} \mid \tau_s^{c'c} = 1\right)\left(\tau_s^{c'c} - 1\right) + \sum_{c'\neq c}\left(\frac{\partial \ln[.]}{\partial \tau_s^{cc'}} \mid \tau_s^{cc'} = 1\right)\left(\tau_s^{cc'} - 1\right) \quad (53)$$

where

$$\left(\frac{\partial \ln[.]}{\partial \tau_s^{c'c}} \mid \tau_s^{c'c} = 1\right) = \frac{G_s^{1-\sigma_s}}{\sum_{c'} Y^{c'}} Y^c (\sigma_s - 1)\left(\frac{\sum_z n_z^{c'}\left(\widetilde{p}_z^{c'}\right)^{1-\sigma_s}}{\left(G_s^{1-\sigma_s}\right)^2}\right) = (\sigma_s - 1) g^c r_s^{c'}$$

and

$$\left(\frac{\partial \ln [.]}{\partial \tau_s^{cc'}} \mid \tau_s^{cc'} = 1\right) = \frac{G_s^{1-\sigma_s}}{\sum_{c'} Y^{c'}} Y^{c'} \left[\frac{(1-\sigma_s) G_s^{1-\sigma_s} - (1-\sigma_s) \sum_z n_z^c \left(\widetilde{p}_z^c\right)^{1-\sigma_s}}{\left(G_s^{1-\sigma_s}\right)^2}\right] =$$

$$= (1-\sigma_s) \frac{Y^{c'}}{\sum_{c'} Y^{c'}} \left[1 - \frac{\sum_z n_z^c \left(\widetilde{p}_z^c\right)^{1-\sigma_s}}{G_s^{1-\sigma_s}}\right] = -(\sigma_s - 1)\left(1 - r_s^c\right) g^{c'}$$

Substituting these results into equation (53), we obtain:

$$\ln\left(\sum_{c'} \frac{Y^{c'}}{\left(G_s^{c'}\right)^{1-\sigma_s}} \left(\tau_s^{cc'}\right)^{1-\sigma_s}\right) = \ln\left(\sum_{c'} \frac{Y^{c'}}{G_s^{1-\sigma_s}}\right) + (\sigma_s - 1) g^c \sum_{c' \neq c} r_s^{c'} \left(\tau_s^{c'c} - 1\right) -$$

$$-(\sigma_s - 1)\left(1 - r_s^c\right) \sum_{c' \neq c} g^{c'} \left(\tau_s^{cc'} - 1\right)$$

This expression can be further simplified by using the adjusted weighted averages defined in equations (13) and (14). Taking also into account that analogous results apply to country o, we obtain

$$\left[\ln\left(\sum_{c'} \frac{Y^{c'}}{\left(G_s^{c'}\right)^{1-\sigma_s}} \left(\tau_s^{cc'}\right)^{1-\sigma_s}\right) - \ln\left(\sum_{c'} \frac{Y^{c'}}{\left(G_s^{c'}\right)^{1-\sigma_s}} \left(\tau_s^{oc'}\right)^{1-\sigma_s}\right)\right] = \tag{54}$$

$$= (\sigma_s - 1)\left[\left(\tau_s^{\overline{c}c} - \tau_s^{\overline{c}o}\right) - \left(\tau_s^{c\overline{c}} - \tau_s^{o\overline{c}}\right)\right]$$

Finally, substituting (54) into (51), we obtain

$$\left[\frac{T_s^c}{Y^c} - \frac{T_s^o}{Y^o}\right] = \gamma_s \ln \widetilde{P}_s^{co} - \mu_s \gamma_s \left[\tau_s^{\overline{c}c} - \tau_s^{\overline{c}o}\right] + \mu_s \gamma_s \left[\tau_s^{c\overline{c}} - \tau_s^{o\overline{c}}\right] - \gamma_s \theta_s^{co} \tag{55}$$

where $\gamma_s = (1 - \sigma_s - \eta_s)\gamma_s'$, $\mu_s = \frac{(\sigma_s - 1)}{(\sigma_s + \eta_s - 1)} > 0$, and $\theta_s^{co} = \frac{Z_s}{(\sigma_s + \eta_s - 1)} \left(\theta_s^c - \theta_s^o\right)$.

B Proof of Proposition 2

We have already shown that $\ln H_s^{cd} \leq \ln P_s^{cd} + \ln \phi_s^c$. Here, we need to show that $\ln \phi_s^c \leq 0$, which will then imply that $\ln H_s^{cd} \leq \ln P_s^{cd}$. A similar proof shows that $\ln P_s^{cd} \leq \ln L_s^{cd}$. We rely extensively the fact that $\sum_{z \in I^j} x_z y_z = Z_j cov_{I^j}(x_z, y_z) + \frac{1}{Z^j} \sum_{z \in I^j} x_z \sum_{z \in I^j} y_z$.

The central part of the proof is to show that

$$\sum_{z \in I_s^{cd}} n_z^c \Delta \widetilde{p}_z^{cd} \geq - \sum_{z \in U_s^{cd}} \widetilde{n}_z^c \frac{1}{Z_s^{cd}} \sum_{z \in I_s^{cd}} \Delta \widetilde{p}_z^{cd} - \sum_{z \in U_s^{cd}} \widehat{n}_z^c \Delta \widetilde{p}_z^{cd}$$

This is done first:

$$\begin{aligned}
\sum_{z\in I_s^{cd}} n_z^c \Delta\widetilde{p}_z^{cd} &= \frac{\overline{n}^c}{\overline{n}^o}\left[\sum_{z\in I_s^{cd}} \widetilde{n}_z^{c,cd}\Delta\widetilde{p}_z^{cd} + \sum_{z\in I_s}\widehat{n}_z^{cd}\Delta\widetilde{p}_z^{cd} - \sum_{z\in U_s^{cd}}\widehat{n}_z^{cd}\Delta\widetilde{p}_z^{cd}\right] = \\
&= \frac{\overline{n}^c}{\overline{n}^o}\left[\begin{array}{c} Z_s^{cd} cov_{I_s^{cd}}\left(\widetilde{n}_z^{c,cd},\Delta\widetilde{p}_z^{cd}\right) + \sum_{z\in U_s^{cd}}\widetilde{n}_z^{c,cd}\frac{1}{Z_s^{cd}}\sum_{z\in I_s^{cd}}\Delta\widetilde{p}_z^{cd} \\ +Z_s cov_{I_s}\left(\widetilde{\widetilde{n}}_z^{cd},\Delta\widetilde{p}_z^{cd}\right) + \sum_{z\in I_s}\overline{n}_z\Delta\widetilde{p}_z^{cd} - \sum_{z\in U_s^{cd}}\widehat{n}_z^{cd}\Delta\widetilde{p}_z^{cd}\end{array}\right] \\
&\geq \frac{\overline{n}^c}{\overline{n}^o}\left[-\sum_{z\in U_s^{cd}}\widetilde{n}_z^{c,cd}\frac{1}{Z_s^{cd}}\sum_{z\in I_s^{cd}}\Delta\widetilde{p}_z^{cd} - \sum_{z\in U_s^{cd}}\widehat{n}_z^{cd}\Delta\widetilde{p}_z^{cd}\right]
\end{aligned}$$

The first equality uses $n_z^c = \widetilde{n}_z^{c,cd} + \widehat{n}_z^{cd}$ and the fact that $I_s^{cd} = I_s - U_s^{cd}$. The second equality uses $\widehat{n}_z^{cd} = \widetilde{\widetilde{n}}_z^{c} + \overline{n}_z$ to decompose the second term, and also uses the fact that $\sum_{z\in I^j} x_z y_z = Z_j cov_{I^j}(x_z, y_z) + \frac{1}{Z^j}\sum_{z\in I^j} x_z \sum_{z\in I^j} y_z$. The inequality uses assumptions 4 and 5, and also the definition of $\widetilde{P}_s^c$ in (10), which implies that $\sum_{z\in I_s}\overline{n}_z\Delta\widetilde{p}_z^{cd} = 0$.

Decomposing $\Delta\widetilde{p}_z^{cd}$ according to its definition in (36) and using assumption 6, after some simple algebra manipulation we obtain

$$\frac{\sum_{z\in I_s^{cd}} n_z^c\left(\frac{\widetilde{p}_z^c}{\widetilde{P}_s^c}\right)^{1-\sigma_s}}{\sum_{z\in I_s^{cd}} n_z^c\left(\frac{\widetilde{p}_z^d}{\widetilde{P}_s^d}\right)^{1-\sigma_s}} \geq 1 \tag{56}$$

which implies that

$$\ln\phi_s^c = \ln\left(\frac{\sum_{z\in I_s^{cd}} n_z^c\left(\frac{\widetilde{p}_z^c}{\widetilde{P}_s^c}\right)^{1-\sigma_s}}{\sum_{z\in I_s^{cd}} n_z^c\left(\frac{\widetilde{p}_z^d}{\widetilde{P}_s^d}\right)^{1-\sigma_s}}\right)^{\frac{1}{1-\sigma_s}} \leq 0 \tag{57}$$

Substituting this result into $\ln H_s^{cd} \leq \ln P_s^{cd} + \ln\phi_s^c$ in equation (29), we obtain

$$\ln H_s^{cd} \leq \ln P_s^{cd} \tag{58}$$

Finally, since with an analogous proof it can be shown that $\ln P_s^{cd} \leq \ln L_s^{cd}$, we can establish that the Paasche and Laspeyres indexes bound the Impure Price Index:

$$\ln H_s^{cd} \leq \ln P_s^{cd} \leq \ln L_s^{cd}.$$

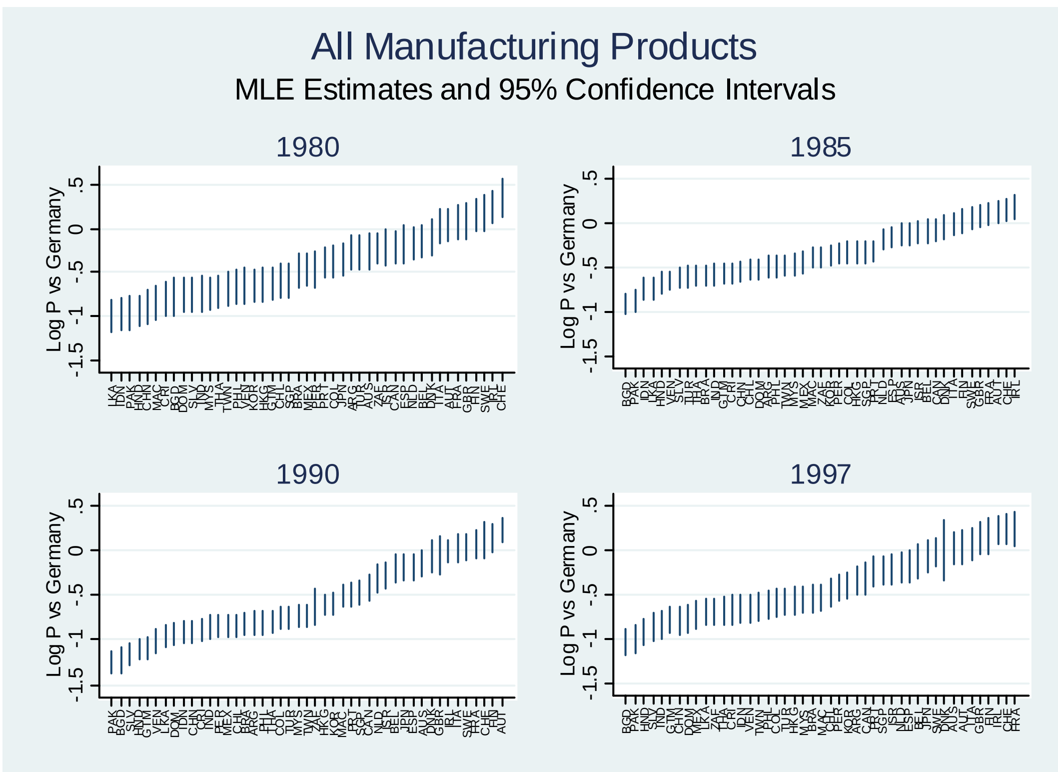


Figure 1: First Stage Maximum Likelihood Estimates of the Impure Price Index $\left(\ln \widehat{P}_{st}^{co}\right)$ for Four U.S. Trading Partners, 1980 to 1997

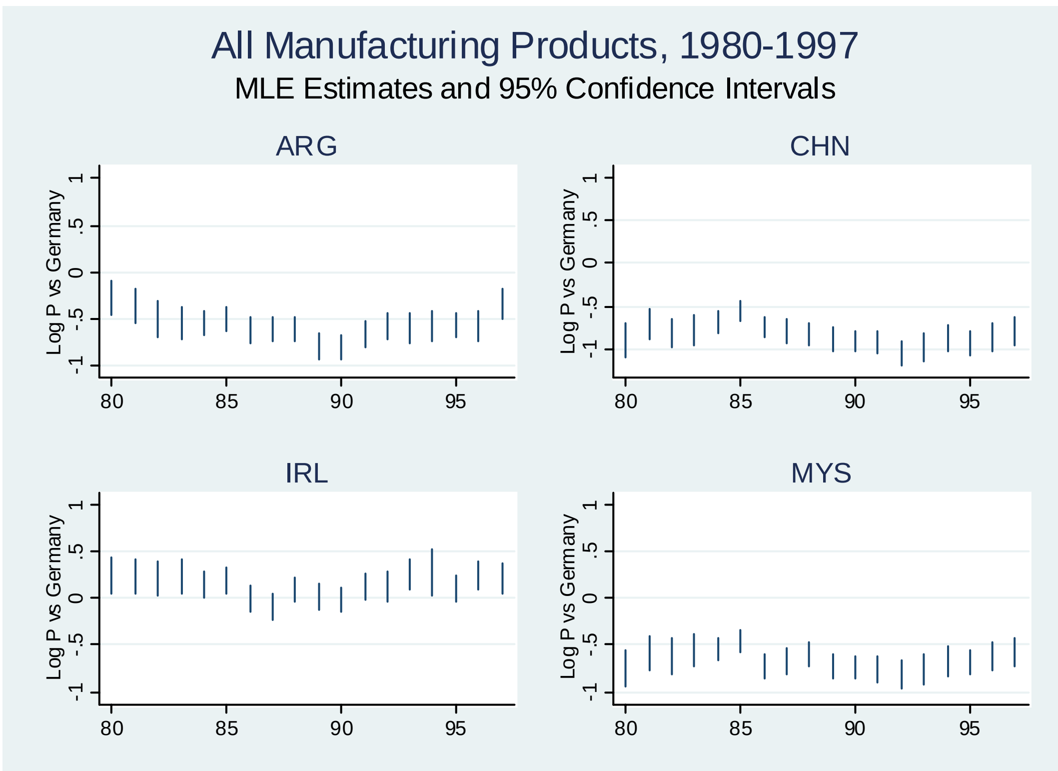


Figure 2: First Stage Maximum Likelihood Estimates of the Impure Price Index $\left(\ln \widehat{P}_{st}^{co}\right)$ for Four U.S. Trading Partners, 1980 to 1997

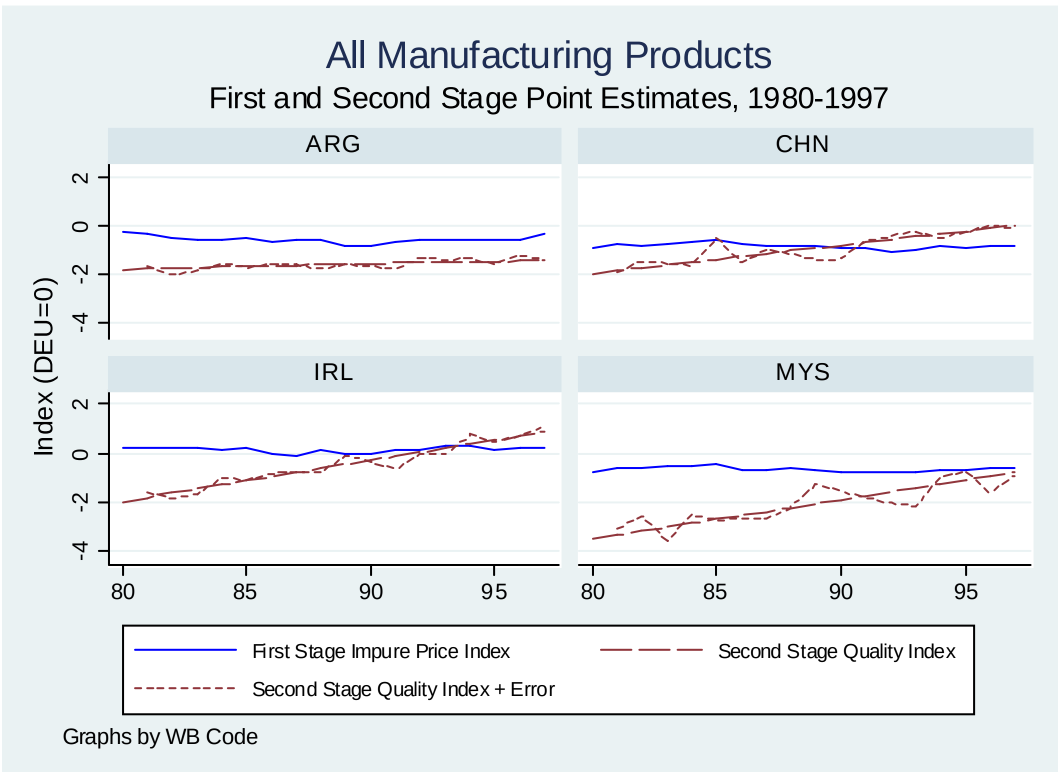


Figure 3: 2SLS Estimates of Linear Quality Trends for Four U.S. Trading Partners, 1980 to 1997

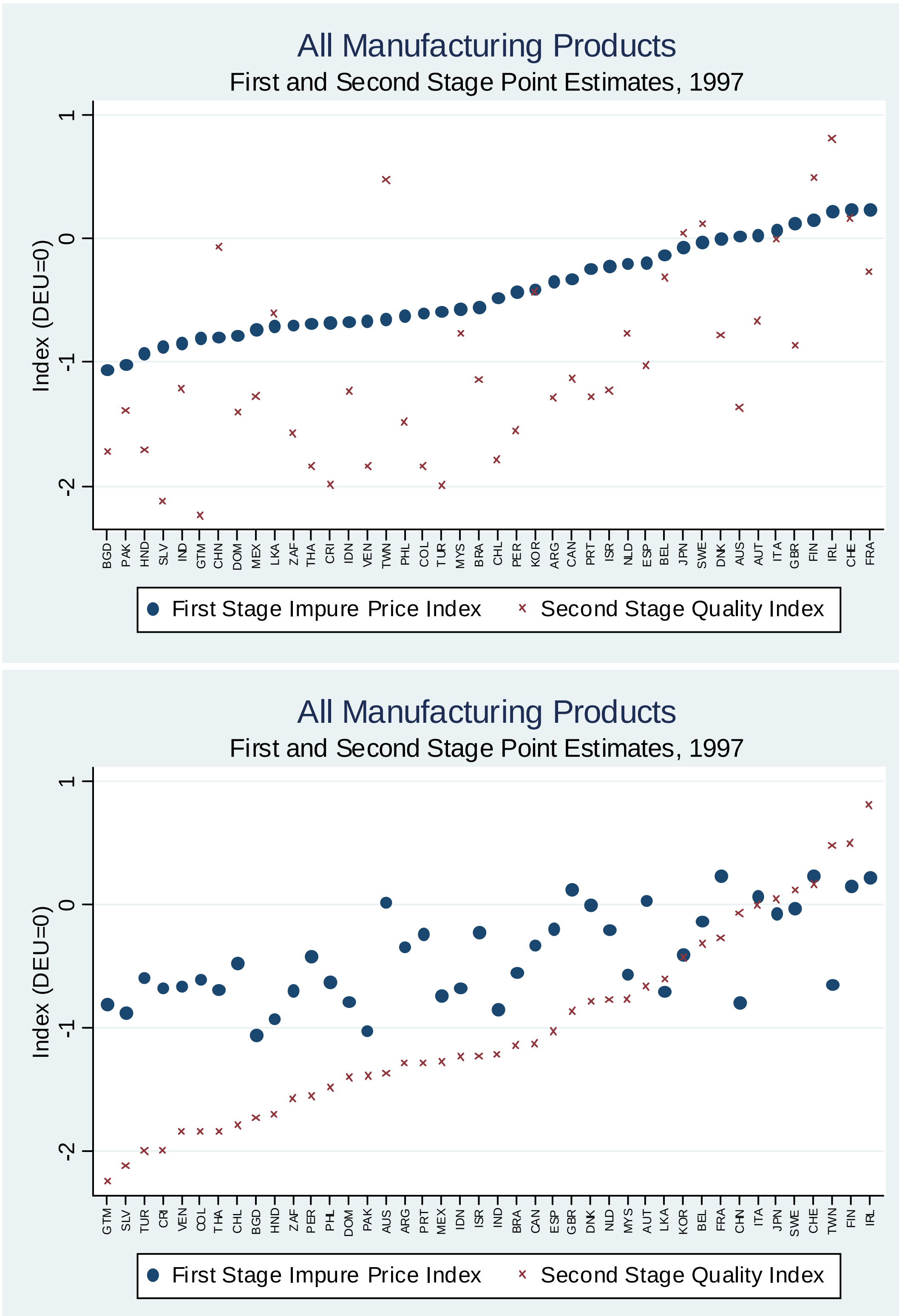


Figure 4: Comparison of First- and Second-Stage Estimates, 1997

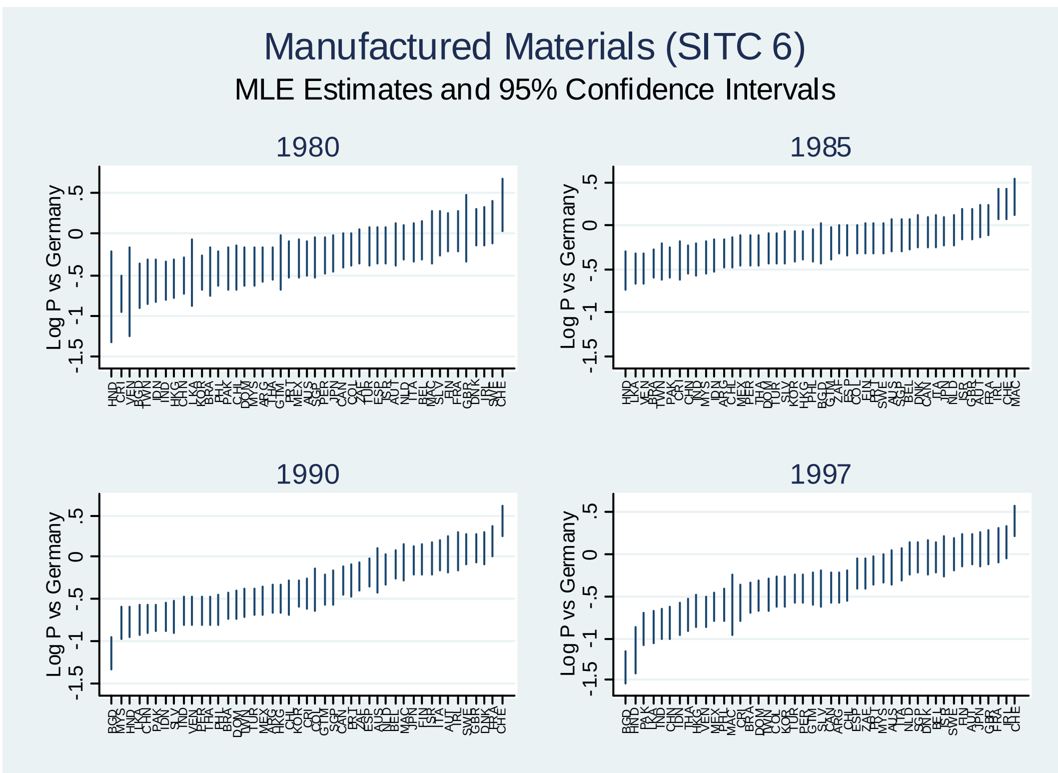


Figure 5: First Stage Maximum Likelihood Estimates of the Impure Price Index $\left(\ln \widehat{P}_{st}^{co}\right)$ for Four U.S. Trading Partners, 1980 to 1997

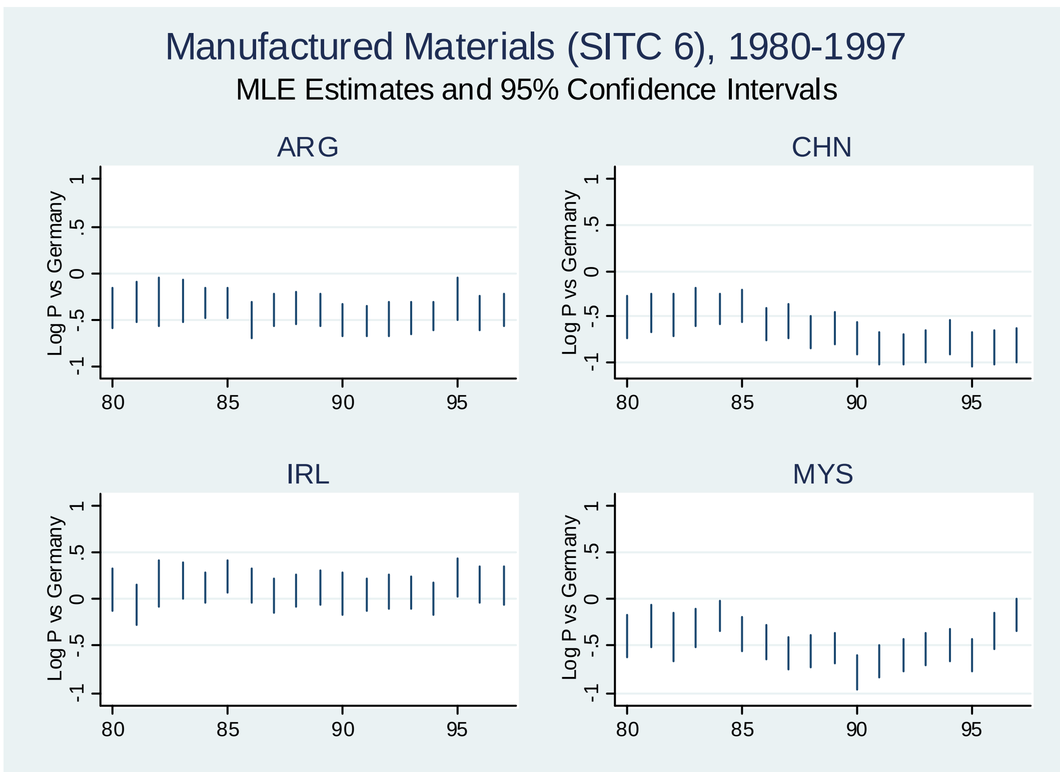


Figure 6: First Stage Maximum Likelihood Estimates of the Impure Price Index $\left(\ln \widehat{P}_{st}^{co}\right)$ for Four U.S. Trading Partners, 1980 to 1997

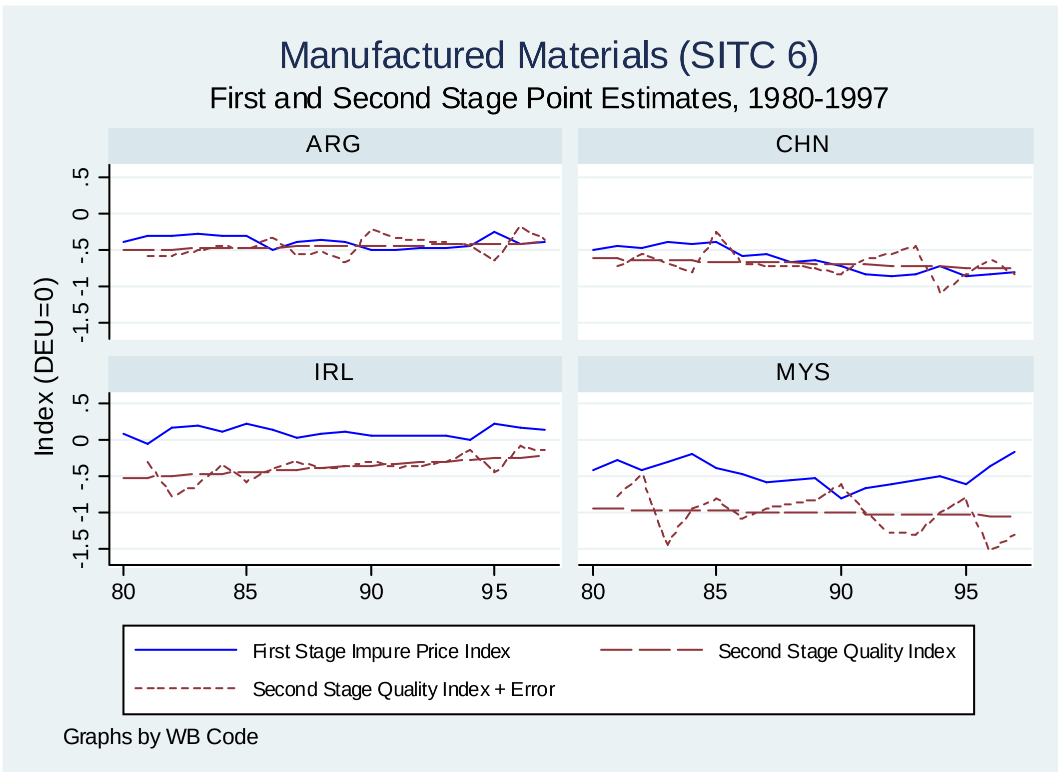
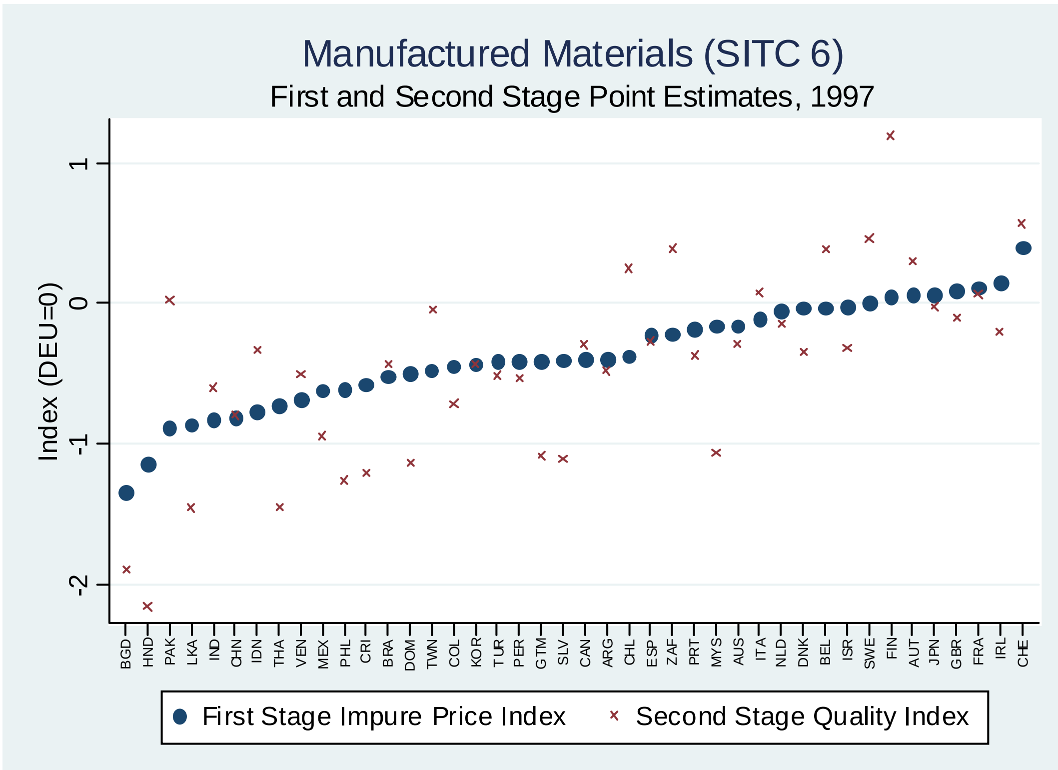


Figure 7: 2SLS Estimates of Linear Quality Trends for Four U.S. Trading Partners, 1980 to 1997



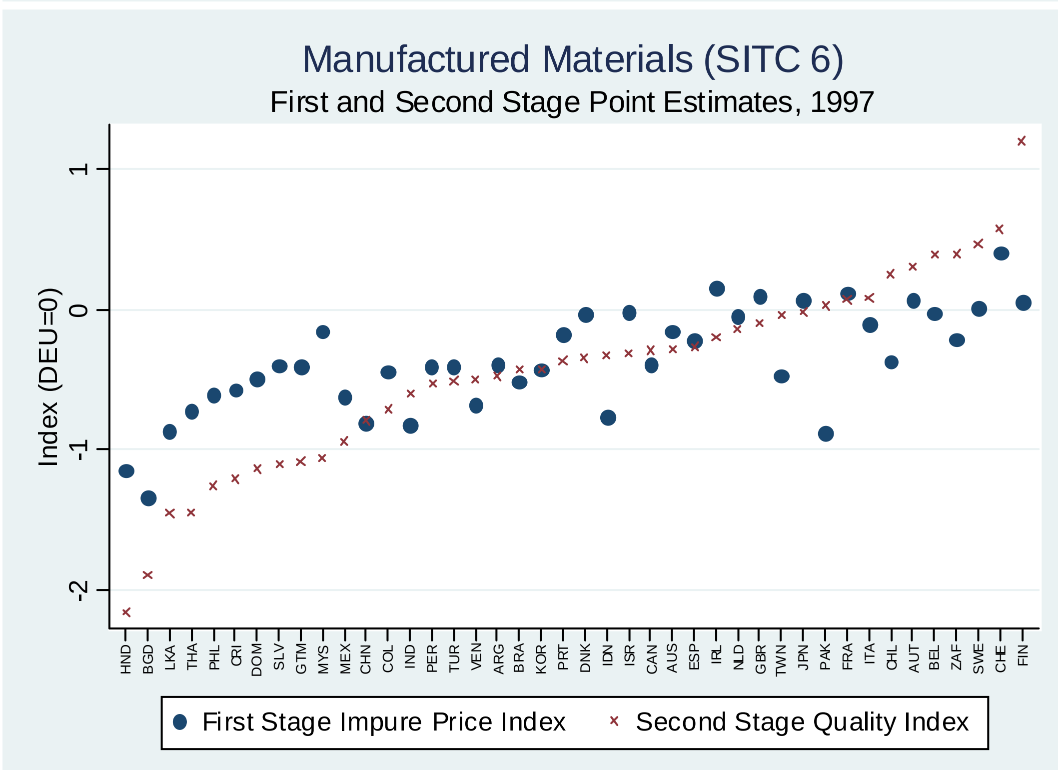


Figure 8: Comparison of First- and Second-Stage Estimates, 1997

Year	Log Likelihood	Mean Commonly Exported Products	Median Commonly Exported Products	Psi
1980	-906	332	220	-0.89
1981	-853	363	234	-0.72
1982	-886	387	246	-0.82
1983	-931	435	287	-0.78
1984	-816	570	389	-0.69
1985	-710	697	505	-0.55
1986	-682	717	529	-0.75
1987	-796	717	524	-0.79
1988	-695	722	535	-0.84
1989	-556	896	649	-0.88
1990	-511	882	648	-0.91
1991	-537	856	616	-0.92
1992	-612	876	641	-1.05
1993	-622	923	682	-0.97
1994	-567	1009	755	-0.88
1995	-426	1113	841	-0.93
1996	-388	1167	879	-0.86
1997	-396	1260	975	-0.79

Notes: Table summarizes results from first stage maximum likelihood estimation on sample of 45 U.S. trading partners relative to base country Germany. Trade data encompasses all manufacucturing products. The number of observations ranges from 1878 in 1980 to 1980 in 1997. First column reports log likelihood. Second two columns reports mean and median number of manufacturing products produced in common across bilateral pairs. Final column reports estimate of psi.

Table 1: First-Stage Estimates, All Manufacturing

	OLS	2SLS	2SLS
Impure Price Index	-0.019	-0.088 ***	-0.096 ***
	0.011	0.024	0.023
Outbound Transport Cost			-1.292
			0.971
Inbound Transport Cost			1.177 *
			0.743
Fixed Effects	Country	Country	Country
Observations	713	713	713
R^2	0.05	.	.
Transport Cost Chi^2 Pvalue	.	.	0.13
First-Stage Fstat	.	102	79

Notes: Results of 2SLS estimation of equation (47) for the years 1980 to 1997. The instrument for the Impure Price Index is the real exchange rate. The transport cost p-value tests the null hypothesis that the coefficient for the inbound transport cost less the coefficient for the outbound cost equals zero.

Table 2: Second Stage IV Estimation, All Manufacturing

Year	Log Likelihood	Mean Commonly Exported Products	Median Commonly Exported Products	Psi
1980	-879	83	45	-0.51
1981	-916	91	50	-0.46
1982	-955	97	52	-0.48
1983	-988	109	61	-0.39
1984	-962	143	84	-0.42
1985	-992	174	102	-0.39
1986	-999	179	114	-0.58
1987	-1011	179	115	-0.55
1988	-978	181	117	-0.67
1989	-760	224	147	-0.63
1990	-730	221	147	-0.74
1991	-680	214	138	-0.84
1992	-722	219	147	-0.86
1993	-686	231	155	-0.83
1994	-665	252	174	-0.72
1995	-706	278	198	-0.86
1996	-689	292	208	-0.84
1997	-661	315	225	-0.82

Notes: Table summarizes results from first stage maximum likelihood estimation on sample of 45 U.S. trading partners relative to base country Germany. Trade data encompasses all manufactured material products (SITC 6). The number of observations ranges from 1564 in 1980 to 1938 in 1997. First column reports log likelihood. Second two columns reports mean and median number of manufacturing products produced in common across bilateral pairs. Final column reports estimate of psi.

Table 3: First-Stage Estimates, Manufactured Materials (SITC 6)

	OLS	2SLS	2SLS
Impure Price Index	0.0019	-0.065 ***	-0.065 ***
	0.0040	0.016	0.016
Outbound Transport Cost			-0.175
			0.400
Inbound Transport Cost			0.386
			0.284
Fixed Effects	Country	Country	Country
Observations	713	713	713
R^2	0.03	.	.
Transport Cost Chi^2 Pvalue	.	.	0.39
First-Stage Fstat	.	31	36

Notes: Results of 2SLS estimation of equation (47) for the years 1980 to 1997. The instrument for the Impure Price Index is the real exchange rate. The transport cost p-value tests the null hypothesis that the coefficient for the inbound transport cost less the coefficient for the outbound cost equals zero.

Table 4: Second Stage IV Estimation, Manufactured Materials (SITC 6)