4

Specific Trade Costs, Quality, and Import Prices

Benjamin Bridgman
Bureau of Economic Analysis

Recently, quality differences among internationally traded goods have received significant attention. Differing import quality across markets, characterized by lower-income countries producing lower-quality products, is a robust empirical finding. Johnson (2011) shows that quality differences account for most firm heterogeneity in trade. Baldwin and Harrigan (2011) argue that, in order to match the data, trade models must account for such differences. However, international price indices frequently cannot make quality adjustments. Correctly accounting for quality differences is important to the measurement of real trade, since mismeasurement of trade filters into other indicators such as real GDP and productivity. (See Feenstra et al. [2013], Houseman [2007], and Houseman et al. [2011].)

While quality measurement is an issue for all price indices, it is a particular challenge for international prices. There has been a significant increase in the number of goods that are traded. A large number of goods that are traded are only traded intermittently. The “new goods problem”—determining the quality of new goods relative to previously traded ones—is a frequent issue in international prices. A lack of quantifiable characteristics or agency resources often prevents explicit adjustments for quality, such as hedonics.

Statistical agencies have developed techniques to deal with environments with shifting sets of goods. A common way of accounting for the quality of newly measured goods is matched modeling. If an explicit adjustment for quality cannot be done, a good may be matched to a similar good. The price difference is attributed to quality differences.

To avoid having frequent replacement of goods in the sample, sampling techniques intentionally focus on consistently traded goods. Price
changes of consistently traded goods within a category stand in for price changes of all goods in that category.

I examine these techniques in light of recent advances in trade theory. I use a version of the model in Baldwin and Harrigan (2011) to show theoretically that both methods are vulnerable to mismeasurement for goods with quality differences that pay specific (per-unit) trade costs. I then analyze the quantitative impact of these forces using U.S. import data.

I show theoretically that matched modeling will tend to overstate quality differences between goods. Specific trade costs weaken the link between price and quality. Prices are set as markup over production and trade costs. Lower-quality goods cost less to produce, and all goods pay the same specific cost regardless of quality. Therefore, a bigger share of a low-quality good’s price is due to trade costs. The price difference between goods will be smaller than their quality differences. Using matched modeling will tend to overstate real imports of new goods. Since matched modeling overstates the quality of new goods, it underestimates the (quality-adjusted) price.

Dropping intermittently traded goods will tend to underestimate price changes. Specific trade costs systematically make goods that enter and exit different from continuing goods. Lower-quality goods are the least profitable, so they are the most sensitive to cost changes. Relatively small cost changes can make a previously profitable market unprofitable, and vice versa. Low-quality goods are more likely to be traded intermittently, and the prices of these intermittently traded low-quality goods are likewise more sensitive to cost shocks.

I show that the quantitative impact of this mismeasurement can be significant: in some cases, applying matched modeling leads to significant overstatement of the quality of new goods. For leather footwear, a major import category, matched modeling understates the quality gap between the highest- and lowest-quality goods by over 30 percent. However, the average impact has fallen over the period of 1974–2004, since transportation costs, which tend to be specific, have fallen.

The impact of dropping intermittently traded goods from the sample has likely increased. The size of the effect is proportional to the price gap between continuing and newly traded goods, a gap that has widened. By 2004, the model predicts that newly traded goods’ prices were twice as sensitive to cost shocks as previously traded ones.
This chapter is part of a literature that examines mismeasurement of international prices. Feenstra and Romalis (2012) also examine international prices with specific trade costs. However, their focus is on macro-level data, while I analyze the micro-level data and the techniques used by statistical agencies. A number of papers have examined difficulties in matched modeling. Reinsdorf and Yusavage (2011) examine country substitution bias, which arises when imports are sourced from new countries with different price levels. Gagnon, Mandel, and Vigfusson (2012) and Nakamura and Steinsson (2012) look at whether the tendency to introduce price changes at product introduction biases import price indices. This chapter is complementary to those papers, as it looks at a different mechanism. Berman, Martin, and Mayer (2012) examine whether entry and exit in response to exchange rates dampen the pass-through of exchange rate fluctuations. The mechanism is similar, though they do not examine its impact on statistical agency methods.

A theoretical literature examines how to accommodate new goods in international price measurement. Feenstra (1994) derives a method of calculating the ideal price index with new goods. This chapter focuses on statistical agency practice and does not deal with welfare.

MODEL

The model is adapted from that found in Bridgman (2013). This model is based on the Quality Heterogeneous Firm Trade (QHFT) model developed by Baldwin and Harrigan (2011) and is similar to that of Gervais (2008).

Households

There are \( J \) number of countries. The preferences of the representative household in each country is given by the following equation:

\[
U = \left( \sum_{i \in \Omega_j} (c_j(i)q(i))^{\frac{1}{\sigma}} \right)^{-\frac{1}{\gamma-1}},
\]

\( (4.1) \)
where $c_j(i)$ is units consumed of variety $i$ in country $j$, and $\Omega_i$ is the set of available varieties. The preference parameters $q(i)$ are the quality of the variety and $\sigma > 1$. The household is endowed with $L$ units of labor.

**Production**

Consumption goods are produced using labor. The wage in country $j$ is $w_j$. There is a constant set of firms, each endowed with a technology to produce a variety. Output of a variety is

$$y(i) = \frac{L(i)}{a(i)}.$$  

Higher-cost firms produce higher-quality goods. A firm with unit cost $a$ produces a good of quality $q$ according to the following equation:

$$(4.2) \quad q(i) = a(i)^{\theta},$$

where $\theta > 0$. The assumption that $\theta > 0$ implies that the consumer’s valuation of quality increases faster than marginal cost, so profit increases in marginal cost. Baldwin and Harrigan (2011) argue that the data support this assumption. Following Eaton, Kortum, and Sotelo (2012), profits are spent outside the economy.

**Trade**

There are three costs to export a variety. There is a market-entry fixed cost of $F^i_{od}$ units of labor to export variety $i$ from origin country $o$ to destination country $d$. There is a specific (per-unit) cost with unit labor requirement $F^i_{od}(i)$. Finally, there is an ad valorem charge $\tau_{od}(i)$. Given a mill price $p_{od}(i)$, consumers pay delivered price

$$p_d(i) = p_{od}(i)(1 + \tau_{od}(i)) + w_d F^i_{od}(i).$$

**Solution**

This section characterizes the solution but does not fully solve it. A full solution to the model would require specifying a distribution of unit costs. Since the results do not require a distribution, I do not fully close the model.
Each representative household chooses \( c_j(i) \) for \( i \in \Omega_j \) to maximize Equation (4.1) subject to

\[
\sum_{i \in \Omega_j} p_j(i) c_j(i) \leq L w_j .
\]

For varieties that are available in a market, expenditure in destination country \( d \) is given by

\[
(4.3) \quad p_d(i) c_d(i) = \left[ \frac{q(i)}{p_d(i)} \right]^{\sigma-1} B_d ,
\]

where \( B_d = \frac{w_d L_d}{p_d^{1-\sigma}} \) and

\[
(4.4) \quad P_d = \left[ \sum_i \left( \frac{p_d(i)}{q(i)} \right)^{1-\sigma} \right]^{1/(1-\sigma)}
\]

is the quality-adjusted price index of destination country \( d \). The demand function in terms of the mill price \( p_{od}(i) \) in origin country \( o \) for a good exported to destination country \( d \) is

\[
c_d \left( p_{od}(i) \right) = \left[ q(i) \right]^{\sigma-1} \left[ p_{od}(i) \left( 1 + \tau_{od}(i) \right) + w_o F_{od}^s(i) \right]^{-\sigma} B_d .
\]

Firms are monopolistic competitors that set prices to maximize profits. They can set different prices for each market. As a simplifying assumption, the firm takes the price index \( P \) as given. The optimal mill price \( p_{od}(i) \) is the solution to

\[
(4.5) \quad \max_{p_{od}(i)} p_{od}(i) c_d \left( p_{od}(i) \right) - w_o a(i) c_d \left( p_{od}(i) \right) - F_{od}^s(i) w_o .
\]

The mill price solution is

\[
p_{od}(i) = \frac{w_o}{\sigma-1} \left[ a(i) \sigma + F_{od}^s(i) \frac{1}{1+\tau_{od}(i)} \right] ,
\]

which generates the delivered price

\[
p_d(i) = \frac{w_o \sigma}{\sigma-1} \left[ a(i) \left( 1 + \tau_{od}(i) \right) + F_{od}^s(i) \right] .
\]
The firm will only export if profits are nonnegative. The goods that are available are determined by whether it is profitable to sell to the market. A variety $i$ will be exported from origin country $o$ to destination country $d$ if

$$
\left\{ \frac{a(i)^{1+\vartheta}}{a(i) + F_{od}(i)} \right\}^{\sigma^{-1}} \left( \frac{\sigma - 1}{W_o} \right)^{\sigma^{-1}} \frac{B_d}{1 + \tau_{od}(i)} \sigma^\sigma \geq F_{od}^{0}(i) W_o.
$$

**SAMPLING**

In this chapter, I attempt to match the model to how international prices are actually collected. Statistical agencies cannot collect price data for all goods that are traded. They must use a sample to stand in for nonsampled goods. In this section, I describe the sampling process the Bureau of Labor Statistics (BLS) uses for its International Price Program (IPP). The BLS’s sampling process is the most germane, since the empirical work examines U.S. trade. The sampling techniques and constraints faced are similar at other statistical agencies, so much of the description applies to other countries.

**Selecting Quotes**

Based on trade data, the BLS sets a sample to determine the number of price quotes needed for each item. The BLS then selects a set of companies to ask for quotes and determines which quotes to ask of each company. A field economist then approaches the company to determine the particular products that will be priced.

The BLS sets a number of goals for its price program and faces a number of constraints when setting its sample. Therefore, the sampling is not a pure proportional probability sample, but a compromise that attempts to achieve its goals within the constraints.

The sample is designed to get prices covering total trade as well as a number of subaggregate price indices. Therefore, it will oversample some products to maintain sufficient coverage of those subindices.

The survey is voluntary and requires the ongoing cooperation of importers or exporters. Resource constraints restrict the number of
new prices that can be gathered and how often the sample is reset. It is more difficult to obtain prices from intermittently traded products, since items that trade too infrequently do not yield usable price changes, so field economists focus on items that are regularly traded. Firms that are involved in trade intermittently tend to cooperate with data collection less frequently; therefore, the survey design intentionally downweights such products and companies.

Not all intermittent trade is due to the effects identified in this chapter. For import prices, the BLS does not have jurisdiction to ask overseas exporters for price data. A foreign company’s goods may be imported consistently, just not by the same importer. Since the BLS can only track the importer’s side of the relationship, the goods from that foreign company will be intermittently traded in the sample. Some goods, like machinery installed in a new factory, are only demanded irregularly.

**Quote Replacement**

Quotes will drop out and need to be replaced periodically. There is both planned and forced substitution.

Planned substitution is replacement built into the sampling design. The sample is reset periodically to reflect changes in the set of products that are traded. Old items are cycled out and replaced by items in the new sample. Forced substitution is due to a product being discontinued or a firm ceasing business. In such cases, the field economist will attempt to get a replacement quote from the trading firm if possible.

Nakamura and Steinsson (2012) report that about half of the quotes that drop out do so because of forced substitution and that a quarter drop out because of planned substitution. The remaining quarter are cases where the firm ceases to provide quotes and gives no reason for stopping. Depending on factors such as how much longer the item was to be included in the sample, the item may either be replaced by a new quote from a different firm or discontinued.

If there is a forced substitution and the new item is substantially different, the reporter is asked for the value of change so it can be subtracted from the new item’s price, a process called “linking.” Gagnon, Mandel, and Vigfusson (2012) and Nakamura and Steinsson (2012) argue that this method is used relatively infrequently. If a new item is added (as in a planned substitution), there is no item with which to link.
When the import prices are put together by the BEA to deflate trade, quality adjustments are made to a few items, largely durable goods, where established hedonic methods are available (BEA 2011).

Nakamura and Steinsson (2012) argue that since explicit quality adjustment is done infrequently, import/export prices are approximately matched-model indices. That is, level differences between items within an index are attributed to quality differences and omitted. Of course, the data collection does not explicitly use matched modeling. However, quotes are often added to a cell without quality adjustment, and level differences between items do not get included. From the standpoint of the theory, this method is equivalent to matched modeling.

RESULTS

This section examines the theoretical difficulties in adjusting for quality. Specifically, I examine matched modeling and the problems posed by sampling intermittently traded goods less frequently. I show that specific trade costs interfere with the assumptions that support the use of these methods.

In the subsections that follow, I will focus on how statistical agencies measure international price change. The BLS uses a Laspeyres index for its import price indices (BLS 1997). The expression for the index measuring a price change from period 0 to period $t$ is

\[ P_t = \sum_i \frac{\omega_d(i) \cdot p_{od,t}(i)}{\sum_i \omega_0(i) \cdot p_{od,0}(i)}, \]

where $\omega_d(i) = p_{od,0}(i) \cdot c_{d,0}(i)$.

This measure is distinct from the theoretical price index that measures the welfare effects of price change. The BLS (1997) states explicitly that the purpose of the international price indices is not to measure welfare.

To isolate the differential impact of costs on price across goods of different qualities, I assume throughout this section that trade costs are the same for all varieties.
Matched Modeling

Matched modeling works off the assumption that if two similar goods are available in the market at different prices, the price gap reflects differences in quality. We can recover the quality gap between an existing and a new good by examining the price gap. In this section, I show that specific costs weaken the link between price and quality.

Without specific costs \((F^s = 0)\), prices closely reflect quality. The relationship between unit cost \(a(i)\) and quality \(q(i)\) can be written as

\[
a(i) = q(i)^{1/\sigma}.
\]

The relative price of two goods \(i\) and \(i'\) that only differ in quality is

\[
\frac{p_{od}(i)}{p_{od}(i')} = \frac{q(i)^{1/\sigma} \sigma w_o}{q(i')^{1/\sigma} \sigma w_o} = \left[\frac{q(i)}{q(i')}\right]^{1/\sigma}.
\]

In this case, matched modeling works well. As long as wages paid by the producers of the two products are the same, the price difference reflects only quality differences. If a comparison good from a producer with similar input costs can be found (for example, from the same country), matched modeling is a practical method for dealing with the new-goods problem.\(^4\)

This clear relationship between price and quality breaks down with specific costs. The relative price is now

\[
\frac{p_{od}(i)}{p_{od}(i')} = \frac{q(i)^{1/\sigma} \sigma + F^s_{od}}{q(i')^{1/\sigma} \sigma + F^s_{od}}
\]

As the specific cost term increases, prices are determined more by trade costs than by quality. Breaking the relationship between price and quality makes matched modeling more difficult. In matched modeling, the price gap between an old and a new good is attributed to quality. As Proposition 1 shows, this method underestimates the quality gap.
Proposition 1: Suppose \( a_k < a_H \). Then
\[
\frac{p_o(L)}{p_o(H)} > \frac{qL}{qH}.
\]

Proof: From the solution to the model, \( \frac{p_o(L)}{p_o(H)} > \frac{qL}{qH} \) if
\[
\frac{w_o}{\sigma-1} \left[ \frac{a(L)\sigma + F_{od}^s}{1 + \tau_{od}} \right]^\theta > \left[ \frac{a(H)}{a(L)} \right]^\theta.
\]
This condition holds if
\[
a(H)\sigma \left[ \left( \frac{a(H)}{a(L)} \right)^\theta - 1 \right] + \frac{F_{od}^s}{1 + \tau_{od}} \left[ \left( \frac{a(H)}{a(L)} \right)^\theta - 1 \right] > 0.
\]
This condition is always true, since \( \frac{a(H)}{a(L)} > 1 \) and \( \theta > 0 \) by assumption.

The specific cost \( F_{od}^s \) has more influence on the price of low-quality goods. Therefore, the price difference will be smaller than quality differences. New goods are of lower quality than prices indicate. This force will tend to overstate the real value of new goods imports.

Sampling

As long as the nonsampled prices move in the same way as the sampled goods, this method gives accurate price measures. However, specific trade costs can introduce differences. Newly and intermittently traded goods are likely to have systematically lower quality than continuing goods. These lower-quality goods react to trade cost changes differently, so deflating these goods by prices of high-quality goods can lead to mismeasurement.
Quality of new goods

Newly traded goods tend to be of lower quality than continuing goods. Since lower-quality goods are the least profitable, they are the most sensitive to cost changes. High-quality-goods exporters will serve even high-trade-cost markets, since they have high margins. Low-margin exporters of low-quality goods are much closer to the zero-profit cutoff. Relatively small cost increases can make a market unprofitable, so these exporters are the most likely to exit.

In the paragraphs that follow, I will vary a cost and hold all other quantities constant. That is, if an exercise changes a specific trade cost so that

\[ F_{od,t+1}^s \neq F_{od,t}^s, \]

all other trade costs and wages are held constant:

\[ F_{od,t+1}^f = F_{od,t}^f, \tau_{od,t+1} = \tau_{od,t}, \text{and } w_{o,t+1} = w_{o,t}. \]

I define cutoff quality \( \overline{q}_{od} \) as the quality level that sets Equation (4.6) at equality; however, changes in trade costs or input prices will change this cutoff. Lemma 1 shows that falling wages and trade costs (holding the other quantities constant) will lead to entry of low-quality goods.

**Lemma 1**: Holding all other quantities constant, if any of the following four conditions hold:

1) \( F_{od,t+1}^s < F_{od,t}^s \),

2) \( F_{od,t+1}^f < F_{od,t}^f \),

3) \( \tau_{od,t+1} < \tau_{od,t} \), or

4) \( w_{o,t+1} < w_{o,t} \),

then \( \overline{q}_{od,t+1} < \overline{q}_{od,t} \).

**Proof**: For proofs of the first three conditions, see Lemmas 2, 3, and 4 in Bridgman (2013). For the proof of the final condition, rearranging the cutoff condition (Equation [4.6]) gives us
If $w_{o,t+1} < w_{o,t}$, the right-hand side of the condition falls. This decline is equivalent to the fixed cost $F_{o,d}^{o}$ falling. Following the proof of Lemma 3 in Bridgman (2013), this implies that $\bar{q}_{o,d,t+1} < \bar{q}_{o,d,t}$.

**Quality and price changes**

The fact that new and intermittently traded goods are of lower quality would not be a problem for sampling if the price changes of low- and high-quality goods were the same. However, low-quality goods react more to cost changes than do high-quality goods. Since more of the price of low-quality goods reflects trade costs, these goods are more sensitive to changes in these costs. The prices of low-quality goods fall (rise) more when specific trade costs fall (rise) than do the prices of higher-quality goods. I show this formally in Proposition 2.

**Proposition 2:** If $a(H) > a(L)$ and either

1) $F_{o,d,t+1}^{o} \not= F_{o,d,t}^{o}$ or

2) $\tau_{o,d,t+1} \not= \tau_{o,d,t}$,

then

$$\frac{p_{t+1}(L) - p_{t}(L)}{p_{t}(L)} > \frac{p_{t+1}(H) - p_{t}(H)}{p_{t}(H)}.$$  

**Proof:** Define $\Delta p(i)$ by $p_{t+1}(i) = p_{t}(i) + \Delta p(i)$. For the condition

$$\frac{p_{t+1}(L) - p_{t}(L)}{p_{t}(L)} > \frac{p_{t+1}(H) - p_{t}(H)}{p_{t}(H)}$$  to hold, we have

$$\frac{\Delta p(L)}{p_{t}(L)} > \frac{\Delta p(H)}{p_{t}(H)}.$$
If either trade cost \( F_{od,t}^i \) or \( \tau_{od,t} \) changes, \( \Delta p(L) = \Delta p(H) \). Formally, if 
\[ F_{od,t+1}^i \neq F_{od,t}^i, \]
then
\[ \Delta p(L) = \Delta p(H) = \frac{F_{od,t+1}^i - F_{od,t}^i}{1 + \tau}, \]
and if \( \tau_{od,t+1} \neq \tau_{od,t} \), then
\[ \Delta p(L) = \Delta p(H) = \frac{F_{od,t}^i}{1 + \tau_{od,t+1}} - \frac{F_{od,t}^i}{1 + \tau_{od,t}}. \]
The condition holds if 
\[ \frac{1}{p_t(L)} > \frac{1}{p_t(H)}. \]
Since \( a(H) > a(L) \), \( p(H) > p(L) \) and the condition holds.

Since they show more price volatility, dropping low-quality goods will tend to underestimate price changes. To see this more concretely, consider the case where both a high- and a low-quality good—\( c(H) \) and \( c(L) \), respectively—are traded in a category, but only the high-quality good is included in the sample. Suppose the specific cost falls \( F_{od,t}^i < F_{od,t+1}^i \). The measured price change for the category is

\[ P^M_t = \frac{[\omega_t(H) + \omega_t(L)] p_t(H)}{\sum_i \omega_t(i) p_{t-1}(H)}. \]
The price change should be

\[ P_t = \frac{\omega_t(H) p_t(H)}{\sum_i \omega_t(i) p_{t-1}(H)} + \frac{\omega_t(L) p_t(L)}{\sum_i \omega_t(i) p_{t-1}(L)}. \]

By Proposition 2, 
\[ \frac{p_{t+1}(L)}{p_t(L)} > \frac{p_{t+1}(H)}{p_t(H)}. \]

That implies that

\[ (4.12) \quad \frac{\omega_t(H) p_t(H)}{\sum_i \omega_t(i) p_{t-1}(H)} + \frac{\omega_t(L) p_t(L)}{\sum_i \omega_t(i) p_{t-1}(L)} > \frac{\omega_t(H) p_t(H)}{\sum_i \omega_t(i) p_{t-1}(H)} + \frac{\omega_t(L) p_t(L)}{\sum_i \omega_t(i) p_{t-1}(L)}. \]

Therefore, \( P^M_t > P_t \), so the measured price change underestimates the price fall.
EMPIRICAL EVIDENCE

The previous section showed theoretically that specific trade costs can lead to mismeasurement. In this section, I examine how important this mismeasurement is empirically.

This section only performs an initial assessment of the empirical scope of the theoretical mechanisms. It does not “fix” the import price index. While I find that these mechanisms appear to have a quantitative impact in some cases, doing a full adjustment of the data will require additional work.

Data

The basic data I use in the data analysis are U.S. goods import data from the Census Bureau, as collected by Hummels (2007). These data give trade value on a customs value (FOB, or free on board), tariffs, freight charges, and weight of shipments from 1974 to 2004. A “good” is defined as an SITC Revision 2 item-and-country-of-origin pairing.

There are a couple of caveats to using these data. First, they are not the data that are used by the BLS to calculate import price indices. The price concept I use is unit value (value per weight) rather than price per product. A product is much more aggregated compared to the prices used by statistical agencies, so it will likely underestimate the real impact of specific costs. Furthermore, the analysis does not cover all trade. Weight data only cover shipments brought in by water or air. Therefore, the portion of trade with Mexico and Canada shipped by rail or truck is excluded. Additionally, not all goods report a weight.

Despite the limitations of the data, they do have advantages that lead me to use them. Most importantly, they are publicly available, unlike the microdata. Quality variation across exporters and locations is a robust finding. (For example, see Bastos and Silva [2010]; Choi, Hummels, and Xiang [2009]; and Hummels and Klenow [2005].) Therefore, country-level variation generates sufficient quality differences to get a first-pass impact of quality difference on price measurement.

In the subsection that follows, I assume that tariffs are all ad valorem charges and that freight rates are all specific costs. That is, $\tau_d(i)$ is the tariff rate and $w_d F_{ad}(i)$ is freight charge per kilogram. Price $p_d(i)$ is unit value. Hummels and Skiba (2004), among many others, find that freight
rates are charged on a specific basis. Tariffs in the post–World War II era are typically charged on an ad valorem basis.

**Matched Modeling**

As documented in Proposition 1, specific trade costs change the relationship between quality and price compared to the case without such costs. To get an empirical measure of this impact, I compare the model’s estimates of the cost parameter \( a(i) \) with and without specific costs. Since we know that specific costs are present, I will assume that the specific trade-cost model is the “true” model. I will use the ratio of the “true” \( a(i) \) and the estimate without these costs, as is usually assumed, as my indicator of quality mismeasurement.

The mill price is given by

\[
p_{odi}(i) = \frac{w_o}{\sigma-1} \left[ a(i)\sigma + \frac{F_{odi}^s(i)}{1+\tau_{odi}(i)} \right].
\]

We can rewrite this equation as follows:

\[
(4.13) \quad a(i)w_o = \frac{1}{\sigma} \left[ p_{odi}(i)(\sigma-1) - \frac{w_oF_{odi}^s(i)}{1+\tau_{odi}(i)} \right].
\]

Neglecting the impact of specific costs (setting \( F^s = 0 \)) will give an estimate of \( \hat{a}(i) \):

\[
(4.14) \quad \hat{a}(i)w \approx \frac{p_{odi}(i)}{p_{odi}(i)}.
\]

Taking the ratio gives us a measure of the overstatement of quality differences from assuming only ad valorem costs:

\[
(4.15) \quad \frac{a(i)}{\hat{a}(i)} = 1 - \frac{w_oF_{odi}^s(i)}{p_{odi}(i)(1+\tau_{odi}(i))(\sigma-1)}.
\]

Specific trade costs are more likely to be an issue when one or more of the following characteristics are present:

1) High specific-cost goods (high \( F^s \))
2) Low-quality goods (low \( a(i) \))
3) Inelastically demanded goods (low \( \sigma \))
Microanalysis

I begin the empirical analysis by examining one good, leather footwear, in detail. I selected this good for a number of reasons. It is one of the 10 largest import categories in the period examined. A wide variety of countries export this good to the United States, with the potential for significant quality differences. In addition, there are few observable attributes that can be used for hedonic quality adjustment. Therefore, there may be room for alternative methods such as the one proposed in this chapter.

I need a value of $\sigma$ to estimate the mismeasurement. I use a value of 2.02, taken from Broda and Weinstein (2006). Table 4.1 reports the estimated $a(i)$ ratio for Switzerland and Sri Lanka at the beginning and end of the sample period. I use these two countries since they represent the high and low ends of unit value, with Swiss exporters charging more than five times the price of their Sri Lankan counterparts in 1974. This spread reflects the fact that the richer countries tend to export higher-quality goods (Fajgelbaum, Grossman, and Helpman 2011).

In 1974, the impact of specific costs on mismeasurement is much stronger for Sri Lanka than it is for Switzerland. Price overstates quality by nearly 40 percent for Sri Lanka, whereas it overstates quality by only 4 percent for Switzerland. FOB prices are selected as a markup over production cost, which is correlated with quality, and over specific cost, which is not. For Switzerland, trade costs are low relative to price. Therefore, most of the price reflects production cost, which reflects quality. Specific costs relative to unit value are much higher for Sri Lanka, so more of the charged price is a markup over trade costs. In 2004, Sri Lanka’s mismeasurement falls significantly. Specific costs

Table 4.1  Leather Footwear $a(i)$ Ratios

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_O$</td>
<td>15.41</td>
<td>2.97</td>
<td>30.97</td>
<td>13.01</td>
</tr>
<tr>
<td>$F^{S_{wo}}$</td>
<td>0.70</td>
<td>0.93</td>
<td>1.49</td>
<td>0.80</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.08</td>
<td>0.13</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>$\frac{\delta_i}{\delta	heta}$</td>
<td>1.04</td>
<td>1.37</td>
<td>1.05</td>
<td>1.06</td>
</tr>
</tbody>
</table>

relative to unit value are much lower. Switzerland and Sri Lanka are much more similar in cost structure, so prices are more reflective of quality.

If each good was mismeasured by the same amount, there would be no impact on matched modeling. As shown above, specific costs affect low-quality goods more, so we would not expect the impact to be the same. To measure the impact on measurement, we need to compare goods across producers. An issue with the trade data is that the producers are different countries, so input costs are unlikely to be the same. The price levels of wealthier countries tend to be higher, as a result of the “Penn effect.” (See Marquez, Thomas, and Land [2012] for a recent empirical confirmation of this effect.) Certainly, wages in Switzerland and Sri Lanka are different.

We can use the model to eliminate the wages from our estimates. If good $k$ is produced by countries $i$ and $j$, the price ratio without specific costs is

$$\frac{p_i(k)}{p_j(k)} = \frac{w_i \hat{a}_i(k)}{w_j \hat{a}_j(k)}.$$  

The equivalent ratio with the true $a(i)$ ratio is

$$(4.16) \quad \frac{w_i a_i(k)}{w_j a_j(k)} = \frac{w_i \hat{a}_i(k)}{w_j \hat{a}_j(k)} \times \frac{a_i(k)}{\hat{a}_i(k)} \times \frac{\hat{a}_j(k)}{a_j(k)}.$$  

The degree to which matched modeling underestimates quality gaps is

$$(4.17) \quad \frac{w_i a_i(k)}{w_j a_j(k)} \times \frac{w_i \hat{a}_i(k)}{w_j \hat{a}_j(k)} = \frac{a_i(k)}{\hat{a}_i(k)} \times \frac{\hat{a}_j(k)}{a_j(k)}.$$  

In 1974, the unadjusted price ratio overstates the quality difference by 32 percent. The unadjusted price ratio

$$\frac{w_i \hat{a}_i(k)}{w_j \hat{a}_j(k)}$$  

is 5.19, while the adjusted ratio $\left( \frac{w_i a_i(k)}{w_j a_j(k)} \right)$ is 6.83.

In 2004, the overstatement falls to 1 percent. Since Swiss and Sri Lankan costs are more similar, so is the degree of mismeasurement. Therefore, the data better reflect the assumptions of matched modeling.
Overstating the quality of new goods will overstate imports. The effect is strongest for low-quality goods. Therefore, this effect will tend to overstate the U.S. trade deficit. American producers tend to produce higher-quality goods, since the United States is a high-income country. U.S. imports have begun to shift to lower-income countries, for whom the effect is stronger. Therefore, imports are more likely to be overstated than exports.

The size of the mismeasurement is sensitive to the elasticity used. For example, the $a(i)$ ratio for Sri Lanka in 1974 drops from 1.37 to 1.10 if $\sigma$ is increased from 2.02 to 4.00. On the other hand, the ratio jumps up to 2.24 if $\sigma$ falls to 1.50. The elasticity governs the degree to which price is marked up over cost. For low values of $\sigma$ (inelastic goods), there are high markups that magnify the impact of specific costs.

**Aggregate analysis**

I now turn to the aggregate effect on quality measurement. I use $\sigma = 4$ for all goods. This is the value Simonovska and Waugh (2011) settle on as a consensus value using U.S. data; the value is within the usual range used in the literature. This will tend to underestimate the impact, since more differentiated goods tend to have a lower value of $\sigma$.

The impact of specific costs is heterogeneous. The range is large, from a ratio of 1 (no distortion) to 3 (200 percent overstatement).

The average $\hat{a}(i)$ ratio over the sample is 1.039, with a standard deviation of 0.067. The goods with the largest ratios are those shipped by air. The mismeasurement is larger for goods with high specific-trade costs. Since air charges are much larger than charges for goods shipped by water, goods shipped mostly or exclusively by air are more subject to this distortion.

So far, I have treated each good equally. To get a sense of the overall impact, Figure 4.1 plots the $a(i)$ ratio against its share in total imports within the sample for 2004. The most distorted goods tend to be a smaller share of imports. However, there are a number of goods that are relatively important that show significant distortion.

As a measure of the aggregate impact, I calculate a trade-weighted ratio of all goods:
Specific Trade Costs, Quality, and Import Prices

Figure 4.1 Estimated Mismeasurement vs. Log Share of Total Imports for 2004 ($\sigma = 4.00$)


\[(4.18) \quad \sum_i s_i \frac{w_i \hat{a}(i)}{w_i a(i)}, \]

where \( s_i = \frac{p_o(i)c(i)}{\sum p_o(i)c(i)} \).

Figure 4.2 shows the weighted ratio, which declines from 1.029 to 1.015. This decline follows the fall in freight rates documented in Hummels (2007). Of course, what matters for matched modeling is the relative mismeasurement within a category. As shown above with Sri Lankan shoes, the decline in specific costs will reduce the scope of this source of mismeasurement. Since the typical good’s price reflects its quality more over time, the typical relative mismeasurement will likely decline as well.
The matched modeling issue may be important for at least some goods. There is reason to believe that this calculation underestimates the degree of mismeasurement. The data may understate actual specific costs. They do not include any other specific costs that accrue because of internal transportation and wholesale and retail trade. Rousslang and To (1993) find that internal trade barriers are significant. Internal transportation costs are 37 percent of international rates. If any of these costs are specific, these estimates will be too low. Using Norwegian data, Irarrazabal, Moxnes, and Opromolla (2011) estimate that the median specific trade cost is 34 percent of a good’s value.

Using the same $\sigma$ for all goods understates the impact on some differentiated goods whose demands are less elastic than $\sigma = 4$. If we set $\sigma$ at 2, the magnitude of the average mismeasurement increases to 9.9 percent in 1974 and 5.4 percent in 2004.

Even if the impact for the average good is small, there are some goods for which it is likely to matter. Lower-income countries, which

Figure 4.2 U.S. Trade–Weighted Estimated Mismeasurement ($\sigma = 4.00$)

tend to produce lower-quality goods, have become more important in U.S. imports. The shift to air transportation, which has much higher freight rates, has increased the specific cost for some goods.

**Sampling**

The sampling method is less likely to collect quotes for intermittently traded goods, whereas the model predicts that new and intermittently traded goods are of lower quality than continuing goods. I begin the analysis by examining whether these goods are of lower quality in the data. Though quality cannot be observed directly, there is evidence that such goods are of lower quality than continuing goods.

Goods that were not traded in the previous year have lower unit values. This set of goods includes both completely new goods and intermittently traded goods that are imported again. Bridgman (2013) shows that newly traded goods enter at a lower unit value, while trade costs are similar across the two sets of goods.

Beginning with Besedes and Prusa (2006a,b), a growing literature has examined the duration of trading relationships. (See Besedes and Prusa [2010] for a survey.) This literature finds that most trade relationships are very short, with the median product being traded for only a year. Lower exporter income, as measured by GDP per capita, is associated with shorter trading relationships. As discussed above, lower-income countries tend to export lower-quality goods.

There is direct evidence that entering and exiting items are of lower quality than continuing goods. Mandel (2010) finds that U.S. goods that cease to be exported are of lower quality. In a later work, Mandel (2013) finds that Chinese exporters to the United States entered at low quality.

**Impact of sampling**

These data do not allow us to assess the quantitative impact of sampling, since we cannot identify which goods are excluded from the sample. However, we can do a back-of-the-envelope calculation to get a sense of quantitative impact. I examine the impact of trade cost changes for low- and high-quality goods. Specifically, I compare what the theory predicts the new prices would be if $F'$ changed to $F''$. To parameterize the exercise, I use new and old goods in 2004 as reported in Bridgman (2013). I identify old and new goods as high- and low-
quality goods, respectively. Their prices are $p_i(H)$ and $p_i(L)$. ($H$ and $L$ stand for high and low.)

Equation (4.19) gives us $a(i)\omega$. Using the price equation, we can calculate $p'_i(i)$ for $i \in \{H, L\}$, the predicted price when $F^v$ changes to $F^v$ and all other quantities are held constant.

If trade costs $F^v$ and $\tau$ are the same for high- and low-quality goods, which is the case for new and old goods in 2004, the relative growth rate of prices is

$$\left(\frac{p'_i(H)}{p_o(H)}\right) \left(\frac{p'_i(L)}{p_o(L)} - 1\right) = \frac{p_o(L)}{p_o(H)}.$$ 

In 2004, we have $\frac{p_o(L)}{p_o(H)} = \frac{0.59}{1.20} = 0.49$.

Therefore, high-quality goods are half as responsive to a change in specific trade costs.

While this example is quite stylized, it indicates that there can be significant differences in price responsiveness among goods of different quality. Using only high-quality goods will tend to underestimate price changes.

There are forces mitigating this effect. Most trade value results from trade relationships that are long lasting. If a trade relationship survives the first few years, the chances that it will end fall significantly (Besedes and Prusa 2010).

Trade relationships in differentiated goods tend to be longer. Besedes and Prusa (2006b) compare trade duration for goods in organized markets with differentiated goods using the classification reported in Rauch (1999). Trade relationships for commodities traded in organized markets tend to be shorter, since such markets lower the cost of switching. These are the goods for which the measurement issues resulting from quality differences are less important.

The impact on aggregate trade measurement is probably small. Most trade value is not impacted by this effect. However, it may have an impact on subindices. The price gap between new and old goods has been increasing, suggesting that the scope for mismeasurement is increasing.
This increasing scope of mismeasurement could have an impact on some of the other uses of trade prices, aside from deflating trade. For instance, it may have a role in explaining the low responsiveness of trade prices to exchange rates. Nakamura and Steinsson (2012) note that trade quotes change very little over time. The items that tend to be included in the price sample are those that are the least affected by cost shocks.

**CONCLUSION**

This chapter shows theoretically that two frequently used techniques in international price measurement, matched modeling and dropping intermittently traded goods from the sample, will mismeasure prices when there are quality-differentiated goods and specific trade costs. Specific costs weaken the link between a good’s quality and its price. This effect causes matched modeling to overstate the quality of low-quality goods. Intermittently traded goods are typically low-quality goods, those whose prices are the most sensitive to shocks. Removing them from the sample will understate price movements. These effects may lead us to overstate the amount of trade from new, low-income exporters, since they tend to produce lower-quality goods. Determining the extent of this overstatement will require additional work using more granular data. However, initial data work indicates that these effects may be quantitatively important for some types of goods.

**Notes**

I thank Jeffrey Blaha, John Greenlees, Larry Lang, and Dave Mead for comments. The views expressed in this chapter are solely those of the author and not necessarily those of the U.S. Bureau of Economic Analysis or the U.S. Department of Commerce.

1. For example, see Hallak (2006); Hallak and Schott (2011); Henn, Papageoriou, and Spatafora (2013); Hummels and Skiba (2004); Irarrazabal, Moxnes, and Opremolla (2011); Manova and Zhang (2012); Martin (2012); and Spearot (2011).
2. This assumption provides closed-form solutions for prices. As shown in Bridgman (2013), the impact of this assumption is small as long as there are a large number of varieties sold.
3. I thank the BLS’s Jeffery Blaha, Larry Lang, and Dave Mead for extensive assistance in explaining the sampling process.

4. There are other issues with match modeling. If there are menu costs, firms may use the introduction of new models as an opportunity to change prices (Nakamura and Steinsson 2012). That concern does not arise in this model, since prices are fully flexible and there are no strategic or informational reasons for not adjusting prices. Therefore, that literature is complementary to this paper.

5. Amiti and Davis (2009) use unit values and argue that they are a reasonable proxy for broad price movements.

6. Quality is actually a function of this cost \( q(i) = a(i)^{1 + \theta} \). By only examining the ratio of the \( a(i) \), we do not have to assign a value for \( \theta \). This ratio shows the impact of specific trade costs on quality measurement, but we would need a value of \( \theta \) to assess the impact on welfare measurement.

7. This value is the 1974–1988 value for SITC Revision 2 Code 85102, taken from the working-paper version (Broda and Weinstein 2004). The published version reports elasticities for the more aggregated three-digit SITC level, while the working paper reports at the five-digit level.

8. I log both variables to make the figure easier to see. I use 2004, the final year of the sample, since it has the most observations.

References


Nakamura, Emi, and Jón Steinsson. 2012. “Lost in Transit: Product Replace-