Product Differentiation and Duration of US Import Trade*

Tibor Besedeš† Louisiana State University

Thomas J. Prusa‡ Rutgers University and NBER

December 2005

Abstract

We examine the extent to which product differentiation affects duration of US import trade relationships. The results are consistent with a matching model of trade formation. Using highly disaggregated product level data we estimate the hazard rate is at least 23 percent higher for homogeneous goods than for differentiated products. The results are not only highly robust but are often strengthened under alternative specifications. As the smallest relationships are dropped, differences across product types increase. Controlling for potential measurement errors also result in larger differences across product types.

Key words: duration of trade, product differentiation, survival analysis
JEL classification: F10, F14

*We would like to thank seminar participants at the NBER ITI group, the 2003 EIIT conference, the Fall 2002 Midwest International Economics Meetings, and Syracuse University for their many useful suggestions. We would also like to thank two anonymous referees for their comments. This paper is a substantially revised version of our 2004 NBER Working Paper, No. 10319.

† *Corresponding Author:* Tibor Besedeš, Department of Economics, E.J. Ourso College of Business, Louisiana State University, Baton Rouge, LA 70803, USA, besedes@lsu.edu (email), 225.578.3803 (telephone), 225.578.3807 (fax)

‡ Thomas J. Prusa, Department of Economics, New Jersey Hall, Rutgers University, New Brunswick, NJ 08901, prusa@econ.rutgers.edu

1 Introduction

The influence of the pioneering work of Krugman (1979) and Helpman and Krugman (1985) has been so profound nearly all discussion of prominent trade issues such as inter- and intra-industry trade, the home market effect, and imperfect competition involve the implicit assumption that trade involving homogeneous and differentiated goods is different. Since Rauch (1999) developed a classification of products into homogeneous, differentiated, and an intermediate category, researchers have explored how trade in homogeneous and differentiated products differs. A set of trade facts has emerged. Differentiated products as compared to homogeneous goods are characterized by: (1) a greater importance of proximity, common language, and colonial ties as well as higher search barriers to trade (Rauch 1999); (2) the home market effect (Feenstra, Markusen, and Rose 2001); (3) lower price elasticities (Erkel-Rousse and Mirza 2002, Broda and Weinstein 2004); (4) higher markups (Feenstra and Hanson 2004); (5) a greater impact of communication costs (Fink, Mattoo, and Neagu 2002); (6) less frequent use of the US dollar as the main currency for trade (Goldberg and Tille 2004); and (7) a lower border effect (Evans 2003).

In this paper we extend this set of facts by using Rauch and Watson's (2003) matching model to generate three testable hypotheses: (1) differentiated products involve smaller initial purchases, (2) differentiated products are traded longer than homogeneous goods, and (3) for each product type duration increases with the initial purchase size. We derive the duration of trade for hundreds of thousands of trade relationships in the US import market and use Rauch's (1999) classification scheme to divide products into three types.

The results not only provide strong support for the notion that product type matters but also indicate there is much to be learned from product level analysis. With respect to the first hypothesis, we find trade relationships involving homogeneous goods consistently start with considerably larger transactions than those involving differentiated goods. On average, relationships involving homogeneous goods start with 40–350 percent larger values as those involving differentiated goods. This result is quite robust as it holds whether we look at more recent or more distant time periods, whether we use highly disaggregated or more aggregated data, and whether we look at the mean or other quantiles in the distribution. Simply put, transactions involving differentiated goods tend to involve smaller values than homogeneous goods.

With respect to the second hypothesis, we find differentiated products have the longest duration, followed by reference priced products, and then homogeneous goods. As compared with differentiated products, the median survival time for homogeneous products is about half as long and the hazard rate for homogeneous goods is about 23 percent higher. As was the case with the first hypothesis, the results are quite robust. Under a variety of alternative assumptions we find the hazard rate for homogeneous goods is consistently larger than that for differentiated goods.

The third hypothesis also finds strong support. Initial transaction size has a long term, persistent effect on duration. Limiting the analysis to relationships with larger initial transactions increases the k-year survival rate for all time horizons and for each product type. Larger initial transactions shift up the survival function throughout the entire horizon. Moreover, the impact on survival is *greater* the longer is the relationship, which suggests a persistent effect on duration.

Interestingly, we find differences in duration between product types increase with initial transaction size. Limiting the analysis to relationships with trade in the first year greater than $100,000 ($1,000,000) the hazard rate for homogeneous goods 71 percent (182 percent) higher than for differentiated goods. Size does matter and it serves to intensify, not diminish, the importance of product type.

Our results are intuitive and consistent with many theoretical models. One might expect relationships involving homogeneous goods to be quite fragile. For products such as corn, wheat, and oil one can imagine a world market where all foreign suppliers ship and all buyers purchase their products and the lowest price rules the day. Trade relationships might be very short; relationship-specific factors may not matter and source country may be irrelevant.

There are a number of explanations why differentiated products may exhibit long duration. The seminal work of Krugman (1979) and Helpman and Krugman (1985) asserts each variety of a differentiated product is desired by consumers. Hence, trade relationships for differentiated goods should be very long-lived and should begin with smaller purchases. If variety is associated with source country, then their models are consistent with our results. Sunk costs and relationship-specific investments can also explain differences in duration. If differentiated goods require larger initial investments, models such as Melitz (2003) suggest once relationships are established they will tend to be robust. In a series of papers Rauch uses network and search theory to explain why trade in differentiated products and homogeneous products is different (Rauch, 1999, 2001; Rauch and Casella, 2003; Rauch and Trindade, 2002).

2 Motivating the Empirical Approach

The finding that homogeneous and differentiated goods display distinctly different trade patterns is consistent with a variety of models ranging from the early work of Krugman (1979) to the more recent work of Melitz (2003) and Bernard, Redding and Schott (2004). These papers, however, do not directly address the issue of duration of trade relationships. Rauch and Watson (2003) do explore the duration issue and for motivational clarity we sketch their model.

The RW model proceeds in three stages: search, investment, and rematch. In the first stage a domestic buyer searches over a large pool of foreign suppliers who differ in their per-period production costs. After paying a search cost and being matched with a foreign supplier the buyer immediately observes the supplier's per period cost. What is unknown, however, is whether the foreign supplier will be successful in fulfilling a large order. Large orders generate greater surplus for the buyer. In the second stage the buyer must decide whether to make a lump-sum investment with the supplier. If the supplier turns out to be reliable, the lump-sum investment means a large surplus will be earned immediately. If the supplier turns out to be unreliable, the lump-sum investment is lost and the buyer must search again. As an alternative to the buyer making a lump-sum investment, the supplier's reliability can be learned over time via small orders which yield zero surplus. If the supplier proves to be reliable, the buyer makes the investment necessary for a large order and places a large order. In the third stage, assuming a good match has been realized, the buyer will periodically be given the option of switching to a new supplier whose reliability is unknown.

RW show there is a unique solution involving three possible actions for the buyer who has just been matched with a foreign supplier: start big (invest immediately), start small (learn), or reject the supplier. RW also characterize how the equilibrium changes as parameters change. All else equal (1) relationships starting with large orders will have longer duration; (2) a decrease in investment costs increases the probability a relationship starts large; and (3) a decrease in search costs increases the likelihood the buyer will opt to switch to a new supplier.

Rauch (1999) argues homogeneous goods are sold on organized markets, which minimize the search cost the buyer is required to pay in order to find an appropriate supplier. Differentiated goods are not sold on organized markets and search costs will be considerably higher as the buyer has to go out and find an appropriate supplier. Similarly, it is reasonable to expect the training cost required before the supplier can deliver a large order is smaller for homogeneous goods. These goods are standardized products which do not differ significantly across suppliers. Differentiated goods, with their multitudes of differences across many dimensions, will require larger and potentially sizeable training costs on the part of the buyer.[1] If we assume differentiated goods have higher search costs and require lower supplier-specific investments than homogeneous goods, RW's model generates the following testable hypotheses.

- The initial purchase size for relationships involving homogeneous goods should be larger than those involving differentiated products (RW, Proposition 3). The relatively small search costs associated with homogeneous goods will lead a buyer to not learn about a supplier with a small order because the buyer can easily find another less risky supplier with whom it can start with a big order.

- Holding initial purchase size constant, duration of relationships involving differentiated goods should be longer than those involving homogeneous goods (RW, Proposition 6). Lower investment costs associated with homogeneous goods will make buyers more likely to switch to an alternative supplier (which lowers duration).

- For each product type, duration of relationships starting with large orders should be longer than those starting with small orders (RW, Proposition 7). All else equal, initial large orders tend to be associated with suppliers who have low marginal costs, and low costs tend to increase duration.

[1] Homogeneous goods may also require substantial training costs. But we believe that training costs will be lower on average for homogeneous goods. For example, differentiated goods suppliers may have to learn the specifications required by a buyer.

We note that search and investment costs have similar implications for the hypotheses. For instance, the smaller either the search costs or the supplier-specific investment, the more likely will a buyer immediately place a large order. Low investment cost means the buyer has less to lose if the supplier proves to be unreliable making it more likely the buyer will start with a large order. Since we expect investment costs to be lower for homogeneous goods we have an additional reason why we should expect initial purchases to be larger for homogeneous goods.

3 Data

The analysis is based on US import statistics as compiled by Feenstra (1996) and extended by Feenstra, Romalis, and Schott (2002).[2] We provide a brief description of data and refer the interested reader to Besedeš and Prusa (2006) for a more detailed description, as well as a discussion of all relevant issues when applying survival analysis methods. The data record annual US imports of products between 1972 and 2001 from virtually every trading partner and include value of trade, quantity, customs collected, and other relevant information.[3] The data set can be divided in two periods based on the nature of the product classification scheme used by US Customs.

From 1972 to 1988 products were classified according to the 7-digit Tariff Schedule of the United State (TS). There is a total of some 23,000 different products imported from about 160 countries. On average, each year some 10,000 products are imported. From 1989 on products are classified according to the 10-digit Harmonized System (HS). Some 23,000 different products are observed since 1989 with an average of 15,000 imported every single

[2] Details on the sources of our data are included in the appendix.

[3] The coverage of some variables is not complete. For example, some observations have missing values for quantity and duty collected.

year from some 180 countries.

Annual data for each country-product pair are used to create spells of service indicating continuous service to the US market. If the US imports product i from country c continuously from 1980 to 1985 then this represents a spell of six years. Such transformation of data results in 495,763 observed trade relationships between the US and a source country for the 1972–1988 period and 620,177 for the latter period. A number of trade relationships experience a dormant phase at some point — the relationship is alive and the US imports a product for a number of years, stops for at least a year, and then starts importing the same product from the same country again. For both periods under study the number of spells of service exceeds the number of trade relationships. There are 693,963 spells of service between 1972 and 1988 and 918,236 in the latter period.[4]

One of the most subtle issues in survival analysis is the treatment of censored observations. Censoring appears in three flavors. Relationships observed in 1972 and 1989 have an uncertain starting date as those are the first years under observation — they may have commenced in 1972 (1989) or before. Similarly, relationships observed in 1988 and 2001 may have truly ended in those years or at a later unobserved time. Both types of censoring are common in duration analysis.

The third flavor is unique to data used. US Customs revises product codes on an annual basis. Some codes deemed to represent products incorrectly, because they are either obsolete or too broad, are discontinued and new codes are introduced. Reclassification may split a single code into several new ones or several codes could be merged into a single new code. We are not aware of any US Customs documentation on reclassification and given

[4] About one-quarter (one-third) of trade relationships experience multiple spells of service at the TS (HS) level and about two-thirds of those experience just two spells. Less than ten percent of trade relationships have more than three spells.

the multitude of changes we are unable to create our own reclassification mapping.[5] As a result, trade relationships in reclassified codes, either discontinued or new ones, are all treated as censored.[6] All censored spells in data have the interpretation of being observed for at least x years.

The benchmark analysis is based on the 7-digit TS product level data. We use the 10-digit HS data to verify whether results are specific to the 1972–1988 period or extend to the 1989–2001 period as well. Industry level data serve as another robustness check. Product level data may be too disaggregated and lead us to observe excessive entry and exit, and result in overly short spells of service. Trade relationships might be better measured using industry level data. We report results using relationships defined at the 5- and 4-digit SITC industry level. An additional benefit of using industry level data is the lack of reclassification of industry codes on an annual basis.[7] The only remaining censoring stems from the start and end of the observation period. Industry level data serve as a check for the treatment of reclassified products as censored.

The next major task involves characterizing the extent of product differentiation for each product. We follow Rauch (1999) and classify commodities into three categories: homogeneous, reference priced, and differentiated. Rauch classified products traded on an organized exchange as homogeneous goods. Products not sold on exchanges but whose benchmark prices exist were classified as reference priced; all other products were deemed differentiated.

Although coarser than one would like, Rauch's classification scheme has several virtues.

[5]Frequent merging of several TS codes into a single HS and splitting of a single TS code into several HS codes make it impossible to establish a concordance between TS and HS codes.

[6]In Besedeš and Prusa (2006) we discuss in great detail the impact of censoring and examine alternative treatments of reclassified product codes. There is no evidence that any of our results significantly depend on how censoring is handled.

[7]While revision 3 of the SITC industry classification was introduced in 1987, we use revision 2 in both periods to identify industries.

First and foremost, it is the only classification scheme we know of that exists at a highly disaggregated level. Rauch classifies products at the 4-digit SITC level and we mapped his codes to the product level codes using the concordance found in Feenstra (1996).

Second, Rauch's scheme makes intuitive sense since it broadly captures what economists mean by product substitutability. Products sold on organized exchanges (e.g., corn, oil, wheat, etc.) are exactly those products typically cited as being homogeneous. Consumers may neither know nor care about the source of the product they are purchasing. Reference priced products are products whose prices are listed in industry guides and trade journals, but are not traded on organized exchanges. These are products that likely have some unique attributes (e.g., quality may vary by source country), but are essentially substitutable. Consumers will know the source country, but it may only have a small impact on their purchasing decision. In the final category are differentiated products. These are products with many characteristics that vary across suppliers and may even be specifically tailored to the end-user's needs. Automobiles are perhaps the most often cited example of such a good; in fact, most consumer goods (e.g., toys, apparel, cookware) are classified as differentiated.

Third, Rauch's classification scheme is quite comprehensive. Some 97 percent of all observed spells of service are covered by the classification in both time periods analyzed. Rauch's scheme is broad as seven of ten 1-digit SITC industries are represented by products from each product type.

We note that the data we use require us to assume goods are differentiated by country of origin. A more realistic assumption might be that goods are actually differentiated by firms within and across countries. If this is the case, the firm-level breaks within countries will not be observed and our duration results might be biased upwards. About 70 percent of spells are observed for two or fewer years indicating the bias is not large when using annual data even if firm-product level data were available.

4 Empirical Findings — Benchmark Product level Data

4.1 Initial Purchase Size

We begin our analysis by looking at whether product type affects initial purchase. In the RW model, all else equal, goods with lower search costs will tend to start with a larger purchase. To examine this we gathered information on the value of trade in the first year of the spell. We then estimated a linear regression of the form

$$v_i = \alpha + \beta D_i + \gamma \mathrm{AVP}_i + \delta \mathrm{NS}_i + \boldsymbol{x}'_{it}\boldsymbol{\phi}_{it} + \epsilon_i, \tag{1}$$

where v_i denotes initial transaction size for the ith spell, D_i is a dummy indicating product type, AVP_i is the average value traded for each product in the year when the spell commences, NS_i is the total number of observed suppliers for a given product, and $\boldsymbol{x}_{it}$ is a vector denoting industry, country, and time effects. All non-dummy variables are in logs.[8]

Three comments are in order. First, there are some spells that are not observed from their actual starting point — we do not know if the first observed year of some spells is in fact the first year of the relationship. For instance, it is not clear whether all relationships observed in 1972 begin in 1972 or started in some year before 1972. This is a problem since the model's prediction is about the *initial* transaction size. The most straightforward way to address the issue here is to exclude all such spells from the analysis in this section.[9]

Second, we measure product types relative to differentiated products, implying $\hat{\beta} > 0$ supports the hypothesis. To streamline the presentation we focus only on the difference

[8] We use the CPI to convert nominal trade values into real 1987 dollars.

[9] We also estimated the regression including all spells. The results are very similar to those in Table 1 and are available upon request from the authors.

between homogeneous and differentiated products.[10]

Third, we include a variety of additional variables controlling for other exogenous reasons for differences in initial size. It is possible, for instance, that product codes classified as homogeneous have on average larger trade values. In addition, suppose initial trading relationships within a product category are randomly distributed across all potential partners. In this case the size of initial flows would be larger for homogenous goods with few potential partners, even if the RW hypothesis was wrong.[11] To mitigate this concern, we control for the size of trade in the category and the number of potential trading partners. To compute the former, for each product code we calculate the average observed traded value in the year when the spell starts.[12] To calculate the latter we use the total number of observed suppliers in each product category over the observed time period.

In addition, we include country, year, and industry dummies. Transaction size may vary by country rather than product type. As mentioned, the RW search cost model was motivated by trade with developing countries. In our empirical application, however, we use trade with developed and developing countries. To control for country characteristics we include country fixed effects.[13] Transaction size might vary by industry rather than product type. We report specifications with 2-digit SITC industry dummies. It is also possible that differences in starting size might reflect different starting years. This would be the case if relationships involving homogeneous goods tended to start in certain years and those involving differentiated goods tended to start in other years. In addition, it may

[10] We estimated the regression including reference price products as well. The results are very similar to those in Table 1 and are available upon request.

[11] We thank a referee for pointing this out.

[12] For example, for a spell starting in 1978, the average reflects the average value of all relationships in 1978 regardless of when they started.

[13] We also used GDP per capita and the World Bank Income Classification of countries instead of country fixed effects. The results are very similar to those presented ($\hat{\beta}$ is positive and statistically significant in every specification) and are available upon request.

also be the case that differences might decline or increase over time. To investigate we include calendar year dummies for the year the spell started.

In Table 1 we report the results. We do not report standard errors in any table because virtually all coefficients are statistically significant at the 1% level. We use a superscript * to denote coefficients that are *not* statistically significant at the 1% level.

Overall the results show clear and consistent support for the hypothesis: $\hat{\beta}$ is positive and statistically significant. Homogeneous goods start with a larger initial transaction than differentiated products. As we include additional fixed effects and control for the average trade value and the number of suppliers $\hat{\beta}$ decreases in size, but it is positive and statistically significant in every specification. Moreover, the estimates are economically relevant. The initial trade value for homogeneous goods is 40–350 percent larger than that for differentiated goods.

To verify the robustness of our results we estimate least absolute deviation regressions for several points in the distribution (25th, 50th, and 75th percentiles). As with the OLS regression, $\hat{\beta}$ is greater than zero and highly significant in every specification. The LAD estimates produce a greater range for the initial size difference, ranging from as low as 10 percent for the 25th percentile to over 700 percent for the 75th percentile.

HS and industry level data further confirm the robustness of the results. Looking at HS data, we note that the point estimate for $\hat{\beta}$ is slightly smaller than with TS data. Nevertheless, $\hat{\beta}$ is positive and statistically significant in every specification. The results using industry level data are remarkably similar to those found with product level data.

Overall, the results are a striking confirmation of the first hypothesis: product type matters for initial purchase. In comparison with homogeneous products, relationships involving differentiated products consistently start smaller.

4.2 Duration Across Product Types

Survival Functions

We now turn to the question whether duration of relationships involving differentiated goods should be longer than those involving homogeneous goods. We begin by examining non-parametric Kaplan-Meier estimates of survival functions across product types. Estimates for the 7-digit TS data are graphed in the upper-left panel of Figure 1. The median survival times are extraordinarily short: five years for differentiated products and two years for reference priced and homogeneous goods. One half of trade relationships involving reference priced and homogeneous goods fail during the first two years.

While survival functions across product types are similar in their general shape, there are notable differences in support for the second hypothesis. As reported in Table 2 and seen in Figure 1, differentiated products dominate the other product types in their survival rates at any stage of a relationship. In year one, 69 percent of relationships involving differentiated goods survive to year two, while only 55 and 59 percent of relationships involving homogeneous and reference priced goods do. By year four, these rates decline to 52 percent for differentiated and 33 percent for homogeneous goods. Between years four and twelve survival rates are stable declining by just 7 percentage points for each product type. The differences in survival across product types are statistically significant.[14]

For comparison we calculate survival rates for the 1989–2001 period using the 10-digit HS data and graph them in the upper-right panel of Figure 2 and present them in Table 2. While differences between product types are not as large as in the earlier period, they are still present and statistically significant. Differentiated goods exhibit similar survival rates in both periods: in the first year during the 1972–1988 period it is 69 percent while

[14]The standard errors are in the range of 0.001 to 0.02.

it is 66 percent for the later period. The biggest difference stems from the higher survival experience of reference priced and homogeneous goods during the 1989–2001 period.

As discussed in Besedeš and Prusa (2006), trade relationships face a high hazard in initial years. As indicated by the declining slopes of the estimated survival functions, hazard rates decline very rapidly for each product type. In the first year they are quite high, 31 percent for differentiated and 45 percent for homogeneous goods in the earlier period. Over the next three years combines hazard rates range from 25 percent for differentiated and 40 percent for homogeneous goods. Beyond the fourth year there is little additional hazard, primarily because most exits are classified as censored rather than failures, resulting in lower hazard rates.

Cox Hazard Estimates

In order to control for a myriad of factors that might be influencing duration, we estimate the Cox proportional hazard model of the form

$$h(t, \boldsymbol{x}, \boldsymbol{\theta}) = h_0(t) \exp(\boldsymbol{x}'\boldsymbol{\theta}),$$

where $\boldsymbol{x}$ denotes a vector of explanatory variables and $\boldsymbol{\theta}$ is to be estimated. The baseline hazard, $h_0(t)$, characterizes how the hazard function changes as a function of time. A particular advantage of the Cox model is that the baseline hazard is left unspecified and is not estimated.

The basic estimation model includes regressors designed to control for country and product characteristics that might influence duration. We include GDP since it is well known that larger countries tend to trade more with each other, which might affect duration. In addition, we include country fixed effects to control for unobserved differences across countries.[15] Suppliers capable of large deliveries may be less prone to disruption and if such capability is correlated with country characteristics we need to include country fixed effects.

We include an ad-valorem measure of transportation costs which we compute as the cif/fob ratio for US imports as reported in Feenstra (1994) and Feenstra et al. (2002). When calculated at the product level this ratio is a reasonable measure for transportation costs (Hummels and Lugovskyy, 2003). We drop a handful of observations with unrealistic transportation costs (e.g., those with zero transportation costs and those with transportation costs being more than double the value of the traded product). Given country fixed effects, the ad-valorem measure captures product specific costs. An alternative interpretation is that distance captures the time dimension of trade — the cost imposed by the length of time it takes to deliver a product to the market (Evans and Harrigan, 2005).

We include several variables that capture relative cost and competitiveness issues. The industry level tariff rate, calculated at the 4-digit SITC level, controls for the ease with which foreign firms can enter the market. The exact effect of tariffs on hazard depends largely on whether time series or cross-section variation dominates. For a given product, an increase in the tariff should lead some foreign firms to exit since higher tariffs raise the cost of servicing the US market. Time series variation in tariffs should lead us to find higher tariffs raise the hazard. Looking across industries higher tariffs mean less competition for incumbent firms. Both domestic and foreign firms servicing the US market face less risk and a lower hazard. Cross section variation in tariffs should lead to higher tariffs lowering the hazard. Given the relative absence of time series variation we expect the cross section effect to dominate, which means higher tariffs should lower the hazard.

The change in the relative real exchange rate should capture the impact of cost changes

[15] In a working paper version we included other regressors found in the gravity literature such as contiguity, distance, and language. Given that we have chosen to include country fixed effects these additional regressors are no longer needed.

on the hazard. To construct it we began by defining each country's exchange rate so that an increase corresponds to a real depreciation (i.e., foreign currency per US dollar). In each year we normalize a country's exchange rate against the US dollar by the average real exchange rate of all countries relative to the US dollar and then calculate annual percentage changes. It gives us a measure of how each country's exchange rate changed relative to its competitors'. An increase in the measure reflects a country's currency has weakened relatively more than its competitors. If one country's currency depreciates relative to other countries' currencies, its firms should become more competitive vis-à-vis other foreign and domestic firms and less likely to exit.

Schott (2001) argues even with data as disaggregated as the product level import data, some products are more broadly defined than one would like. The coefficient of variation of unit values for each TS (or HS) product in each year controls for diversity of products within each product code. We expect the smaller the coefficient of variation the more standardized the product and the greater the hazard.

There is also the issue of multiple spells — trade relationships with multiple periods of service separated by a period with no service. Some trade relationships are observed for a period of consecutive years (spell 1), followed by a period of no trade, and then again observed for another service spell (spell 2). We believe the first failure makes a second failure more likely (higher hazard). On the other hand, it is possible the return of the foreign supplier to the market is a positive sign making a second failure less likely. In either case the hazard rate will depend on whether we are observing a second spell and should be controlled for. We treat multiple spells as independent and use a dummy to control for any impact of higher order spells.[16]

[16]We considered alternative methods for addressing multiple spells and found no significant changes in the results. Results are available upon request.

Finally, since agricultural products are generally classified as homogeneous products and since agricultural products are more likely to be subject to weather or disease disruption we include an agricultural dummy. Agricultural goods are defined as those belonging to the 4-digit SIC industries starting with a 0 (e.g., 0181, 0272, 0912).

In the first column of Table 3 we report the benchmark estimates based on the 1972–1988 period. Throughout we present results in terms of hazard ratios. An estimated hazard ratio less (greater) than 1 is interpreted as implying the variable lowers (raises) the hazard rate. A ratio equal to 1 implies no impact on the hazard rate.

Larger countries, as measured by GDP, face a lower hazard. Given the variance in GDP across countries, the size of the effect depends significantly on country size. Thinking either in terms of differences across time or countries a $100 billion increase in GDP lowers the hazard rate by about 5 percent; a $1 trillion increase in GDP lowers the hazard rate by about 50 percent.

Transportation costs have a sizeable impact. A 10 percent increase in transportation costs raises hazard by 7 percent. Industries with higher tariffs face a lower hazard. A 1 percentage point higher tariff lowers hazard by approximately 2 percent. Depreciation of a country's currency lowers its hazard. A 10 percent depreciation in the real exchange rate (relative to other suppliers) lowers hazard by about 10 percent.

Products with higher variation in unit values face a significantly lower hazard rate which suggests high and low prices for the same product increase duration. The result confirms Schott's (2001) contention that products within a single product code may not be identical. For instance, the "cotton T-shirt" product might include commodity grade products from China and Bangladesh and fashion designer products from Italy. It would be surprising if the effect were common for all three product types. As we will show below, the effect differs across product types.

We are primarily interested in the product type estimates. Letting differentiated products be the benchmark, reference priced products have a 17 percent higher hazard and homogeneous goods a 23 percent higher hazard. The estimates strongly support what Figure 1 suggested: namely, product type matters.

We ran a similar regression for the 1989–2001 HS data as a robustness check. The estimated coefficients, along with those from other robustness checks, are presented in Table 5. In the interest of brevity and since we are primarily interested in the effect of product type we only report estimates for product type dummies.[17] As was the case with the TS data, product type matters, albeit somewhat less. Reference price products face a 6 percent higher hazard and homogeneous goods face an 8 percent higher hazard.

In Table 4 we report the results when we consider a more flexible specification,

$$h(t, \boldsymbol{x}, \boldsymbol{\theta}) = h_0(t) \exp\left(\sum_{i=h,r,d} D_i \boldsymbol{x}' \boldsymbol{\theta} \right),$$

where D_i denotes the ith dummy corresponding to product type. We allow each variable, including country characteristics, to have a product type specific effect. Given that the RW model was motivated by developed country trade with developing countries, the more flexible specification accounts for the significant differences among suppliers.

The estimates confirm what we learned from the basic specification. Namely, larger countries, high tariff products, and weaker currencies, all have lower hazard rates for each product type. Just as standard economic theory would predict, higher variation in unit values lowers hazard for differentiated products, but increases it for homogeneous goods. In a homogeneous good market one would be surprised if high priced suppliers would have long lived spells of service; in a differentiated product market, high priced suppliers could

[17]All other regression coefficients are qualitatively similar to those in the benchmark case. Complete regression results are available upon request.

easily have a long service spell if their high prices reflect quality differences.

Even after allowing for systematic differences across product types, reference priced products and homogeneous goods face a significantly higher hazard than differentiated products, on the order of 32 and 24 percent. Overall, the nonparametric and semiparametric results provide strong support for the hypothesis that relationships involving differentiated goods should have longer duration than those involving homogeneous goods.

4.3 Initial Size and Duration Across Product Types

The RW model implies spells with large values of trade to be longer lived. All else equal, one might expect a relationship with $1 million of trade in the first year of the spell to survive longer than one starting with $100,000. We now explore whether small valued spells are at greatest risk. This exercise will also shed light on whether the observed differences in duration across product type are driven by the small-value relationships.

In order to investigate whether small valued spells are at greatest risk we filtered out small dollar-value observations; that is, we eliminated spells with trade in the first year below some minimum level. We then estimate survival functions for each product type after dropping the small-valued observations (Figure 1). In Table 2 we report survival rates based on dropping all observations where the value of trade in the first year of the spell was less than $100,000 ($1,000,000).

Two important insights are gained. First, survival functions shift up as we progressively drop observations. Spells that begin with small trade value are at greatest risk. This is true for all product types. For example, for differentiated goods the one-year survival rate increases from 0.69 to 0.92 to 0.99 as observations with progressively larger initial transactions are used. This pattern holds for each product type in both time periods studied. As implied by the model, the larger the initial purchase, the longer the duration for each product type.

Second, the estimates provide no evidence differences among product types are driven by small observations. Differences among product types grow as we eliminate the smaller trade observations. When we restrict the sample to only those spells with initial transactions exceeding $1,000,000 the one-year survival rate is 99 percent for differentiated and 75 percent for homogeneous goods which compare with 69 and 55 percent for the benchmark.

To get a more precise quantification of these effects we re-estimate the Cox proportional hazard model filtering out spells that start small. The estimates are reported in columns two and three of Tables 3 and 5. Comparing these estimates with the benchmark (column one), we find the impact of GDP, tariffs, and multiple spells increase. Most importantly, product type dummies indicate our results are not driven by small value spells. Compared to differentiated products, homogeneous goods face a 71 percent higher hazard at the $100,000 cutoff level, and a 182 percent higher hazard at the $1,000,000 cutoff level. Reference priced products face a 59–155 percent higher hazard.

We use the same size cutoffs for the 1989–2001 HS data and report these results in Table 5. As was the case with TS data, filtering out small observations results in an increased impact of product type. Reference price products face a 26–68 percent higher hazard and homogeneous goods face a 41-124 percent higher hazard. These all exceed the benchmark HS estimates.

The bottom-line is size does matter. Rather than weakening the results controlling for size strengthens them: product type is even more important than the benchmark results suggest. We emphasize that the importance of initial purchase is product type specific. That is, within each product type larger initial purchase size increases duration. If we do not control for product type the impact of size is much murkier. As we saw in section 4.1 relationships involving homogeneous goods tend to start much larger than those involving differentiated goods, but homogeneous goods also tend to have significantly higher hazard.

5 Robustness

We perform several exercises to investigate whether differences between product types are robust. There are three concerns we explore. First, are the results affected by a potential measurement error regarding the end of the spell? Second, are the same patterns found in industry level data? Third, are differences across product types driven by the distribution of product types across industries?

5.1 Measurement Error

We examine how the results are affected if we incorrectly infer the end of a trade relationship. The concern involves trade relationships with multiple spells. If the time between spells is short, it is possible the gap is mis-measured and interpreting the initial spell as "failing" is inappropriate. The two spells might be better thought of as one longer spell. If the measurement error is related to product type in that homogeneous goods are more likely to have short gaps, then the results will misrepresent the role of product type.

We address this issue by assuming a one-year gap between spells is an error, merge the individual spells, and adjust duration accordingly. Gaps of two or more years are assumed to be accurate and no merging is done. As an example, suppose the US imports a product from country c in 1973–74, 1976–77, and 1979–1980. Without any adjustment this trade pattern is interpreted as three distinct spells, each of length two years. Assuming all gaps of one year are errors, there is just one distinct spell with a length of eight years.

Survival functions for the product level data using the gap-adjustment modification are shown in the lower right panel of Figure 1 and in Table 2. The probability of survival for each product type increases at each point in time, yet differences across product types remain. The same can be inferred from the estimation results reported in Table 5. The impact of product type increases slightly as compared with the benchmark specification. As compared with differentiated good, reference priced goods have a 21 percent higher hazard while homogeneous goods have a 26 percent higher hazard in the 1972–1988 period, and 9 and 11 percent higher hazard in the 1989–2001 period.

5.2 Aggregation

We can define trade relationships using industry level data. In addition to mitigating the censoring problem, industry level analysis allows us to explore whether differences across product types are due to the highly disaggregated nature of our data. The concern is that the TS (HS) classification system is too fine and leads us to observe excessive entry and exit. Trade relationships might be better measured at the industry level. Although the results presented in section 4.2 and as argued by Schott (2001) suggest the product level classification is not too disaggregated, we do want to investigate this possibility.

We calculated spells of service using the 5- and 4-digit SITC industry data. We plot the Kaplan-Meier survival function for each product type in the bottom-half of Figure 2 using the 5-digit SITC data.[18] As with benchmark data, in both periods the survival function for differentiated products is above the survival function for reference priced products, which in turn is above the survival function for homogeneous goods. The figure suggests differences among product types are somewhat attenuated as compared to benchmark data. The median survival time for all three product types is just two years. Nevertheless, differences persist and are statistically significant.

[18]The Kaplan-Meier plots are similar for the 4-digit data and available upon request.

In Table 5 we report the Cox proportional hazard estimates for both periods.[19] Interestingly, differences among product types are about the same using industry level and product level data for the 1972–1988 period and are somewhat larger for the latter period. The industry level analysis confirms the main finding: differentiated products face a significantly lower hazard than reference price goods which in turn have a significantly lower hazard than homogeneous goods.

5.3 Industry by Industry Analysis

A vast majority of products are classified as differentiated, while fairly few are classified as homogeneous. Furthermore, two industries, "Machinery" (SITC=7) and "Miscellaneous Manufactures" (SITC=8), are composed entirely of differentiated products.[20] Are product type differences driven by the distribution of product types across industries?

Using the product level data we re-estimate the benchmark specification for each 1-digit industry separately. We chose this approach rather than simply including industry dummies because the majority of products are in two industries (SITC=7 and 8) which contain only differentiated products. We report the results in Table 6. We again only report estimates for product type dummies. In the top panel we report estimates for the benchmark 7-digit data. The results largely confirm earlier findings — for both periods differentiated products have lower hazard rates and involve longer-lived spells than homogeneous goods for six of seven industries.

There are two important caveats. First, in the early period for four industries (SITC=1,

[19] Not all variables are available at the SITC level. In particular, the regressions do not include the coefficient of variation of unit values because units vary across products within the same industry. Also, Hummels and Lugovskyy (2003) argue that the ad valorem transportation cost is unreliable at the industry level.

[20] A third industry, "Other" (SITC=9), is composed of only differentiated and homogeneous goods and is not presented.

2, 3, 4) the coefficient on reference priced goods is not statistically significant. In the later period the reference price good coefficient is statistically significant for six of seven industries. Second, differentiated products have a higher hazard rate than reference priced or homogeneous goods for "Mineral Fuels" (SITC=3). We are not troubled by this result because (i) we do not know what it means for a mineral fuel to be differentiated and (ii) in this industry being differentiated may be undesirable.

All in all, these findings confirm the main result — differences across product types — is not driven by the distribution of product types across industries.

6 Conclusion

This paper offers additional empirical evidence that trade in differentiated and homogeneous products is different. We do so by focusing on duration and initial value of trade relationships. Using the Rauch and Watson (2003) search cost model we show relationships involving differentiated goods should begin with smaller initial purchases and should last longer, and for each product type larger initial purchase should result in a longer relationship. Homogeneous and differentiated products differ by the extent of search and investment costs that a buyer must spend before a supplier can deliver a large order. Our working assumption is that differentiated goods involve larger search and investment costs.

Differentiated goods start with considerably smaller initial purchases. On average, the initial purchase for homogeneous goods is 40–350 percent larger than for differentiated goods. Differentiated products have a median duration more than twice as long as either reference priced or homogeneous goods. The hazard rate for differentiated products is 18 percent lower than for reference priced products and 25 percent lower than for homogeneous goods. Furthermore, for each product type, as the value of initial purchase increases, duration increases as do differences across product types. We perform a number of robustness exercises which show differences in duration and initial purchase across product types are systematic. The results indicate lower transportation costs, higher GDP, higher tariffs, and depreciation of the source country's currency all lead to longer durations. Higher variation in unit values lowers hazard for differentiated products and raises it for homogeneous goods.

There are several avenues that can be explored by trade economists beyond this paper. The analysis indicates survival in US import markets will be longer if a differentiated good rather than homogeneous product is traded. An open question to be answered is whether exporters should focus on differentiated products. Ffuture work should study whether a country's development experience is related to its shift from homogeneous to differentiated products.

The results suggest search costs and networks are more important for differentiated goods than for homogeneous goods, as has been suggested by Rauch (1999). Higher search costs for differentiated goods result in lower initial purchases but also in longer lasting relationships. Low search costs (or their absence) for homogeneous goods result in higher initial purchases, but less stable relationships. Future work should focus more on incorporating the search cost/network view in explaining trade in differentiated goods.

Our results indicate there is a great deal of turnover in international trade. The majority of trade relationships last for only a few years, which could potentially indicate there is a lot of uncertainty in the international market. As the Rauch and Watson (2003) model states, relationships start with small purchases in order to avoid potential costs resulting from uncertainty present when looking for a trade partner. Uncovering the sources of uncertainty and ways to mitigate it should be a part of future research with implications for both firms and policy makers.

While traditional models of trade fit aggregate data reasonably well, they glance over many interesting economic phenomena. Aggregate data preclude the possibility of investigating duration of trade relationships, the value of initial purchase, uncertainty present in different markets, search costs needed to find appropriate partners, the extent to which importers switch among foreign suppliers of a product, the extent to which exporters switch across different markets, and the difference in prices of similar goods from different suppliers among other issues. In order to investigate these issues future research should focus on using product- and firm-level data in the vein of work done by Roberts and Tybout (1997), Bernard and Jensen (2004), and Schott (2001) among others. Future research should work on developing firm-product-level data which would track products particular firms produce and trade. Such data would allow us to examine duration at the firm-product-level and uncover within country and within product firm heterogeneity, which is currently obscured by the use of country-product level data.

References

[1] Bernard, A.B., Redding, S., Schott, P.K., 2004. Comparative Advantage and Heterogeneous Firms. NBER Working Paper No. 10668.

[2] Bernard, A.B., Jensen, J.B., 2004. Why Some Firms Export. Review of Economics and Statistics 86, 561-569.

[3] Besedeš T., Prusa, T.J., 2006. Ins, Outs, and the Duration Trade. Canadian Journal of Economics 39, forthcoming.

[4] Broda, C., Weistein, D.E., 2004. Globalization and the Gains from Variety. NBER Working Paper No. 10314.

[5] Erkel-Rousse, H., Mirza, D., 2002. Import Price Elasticities: Reconsidering the Evidence. Canadian Journal of Economics 35, 282–306.

[6] Evans, C.L., 2003. The Economic Significance of National Border Effects. American Economic Review 93, 1291-1312.

[7] Evans, C.L., Harrigan, J., 2005. Distance, Time, and Specialization: Lean Retailing in General Equilibrium. American Economic Review 95, 292–313.

[8] Feenstra, R.C., 1996. US Imports, 1972–1994: Data and Concordances. NBER Working Paper No. 5515.

[9] Feenstra, R.C., Hanson, G.H., 2004. Intermediaries in Entrepot Trade: Hong Kong Re-Exports of Chinese Goods. Journal of Economics and Management Strategy 13, 3-35.

[10] Feenstra, R.C., Markusen, J.R., Rose, A.K., 2001. Using the Gravity Equation to Differentiate Among Alternative Theories of Trade. Canadian Journal of Economics 34, 430–447.

[11] Feenstra, R.C., Romalis, J.E., Schott, P.K., 2002. US Imports, Exports, and Tariff Data, 1989–2001. NBER Working Paper No. 9387.

[12] Fink, C., Mattoo, A., Neagu, I.C., 2002. Assessing the Impact of Communication Costs on International Trade. World Bank Policy Research Working Paper 2929.

[13] Goldberg, L.S., Tille, C., 2004. Vehicle Currency Use in International Trade. Federal Reserve Bank of New York Staff Report No. 200.

[14] Helpman, E., Krugman, P.R., 1985. Market Structure and Foreign Trade. MIT Press, Cambridge.

[15] Hummels, D., Lugovskyy, V., 2006. Usable data? Matched partner trade statistics as a measure of international transportation costs. Review of International Economics 14, forthcoming.

[16] Krugman, P.R., 1979. Increasing returns, monopolistic competition, and international trade. Journal of International Economics 9, 469–79.

[17] Marazzi, M., Sheets, N., Vigfusson, R.J., Faust, J., Gagnon, J., Marquez, J., Martin, R.F., Reeve, T., Rogers, J., 2005. Exchange Rate Pass-through to U.S. Import Prices: Some New Evidence. Federal Reserve Board International Finance Discussion Papers No. 833.

[18] Melitz, M., 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71, 1695–1725.

[19] Rauch, J.E., 1999. Networks versus Markets in International Trade. Journal of International Economics 48, 7–35.

[20] Rauch, J.E., 2001. Business and Social Networks in International Trade. Journal of Economic Literature 39, 1177–1203.

[21] Rauch, J.E., Casella, A., 2003. Overcoming Informational Barriers to International Resource Allocation: Prices and Ties. Economic Journal 113, 21–42.

[22] Rauch, J.E., Watson, J., 2003. Starting Small in an Unfamiliar Environment. International Journal of Industrial Organization 21, 1021–1042.

[23] Rauch, J.E., Trindade, V., 2002. Ethnic Chinese Networks in International Trade. Review of Economics and Statistics 84, 116–130.

[24] Roberts, M.J., Tybout, J., 1997. The Decision to Export in Columbia: An Empirical Model of Entry with Sunk Costs. American Economic Review 87, 545-564.

[25] Schott, P.K., 2001. Do Countries Specialize?. NBER Working Paper No. 8492.

[26] Swenson, D.L., 2005. Overseas Assembly and Country Sourcing Choices. Journal of International Economics 66, 107–130.

A Data Appendix

Data used in this paper are available from public sources.

Variable	**Source**
7-digit TS and 10-digit HS Import Data	Robert Feenstra's online data collection at http://data.econ.ucdavis.edu/international/
5- and 4-digit SITC Import Data	http://data.econ.ucdavis.edu/international/
Rauch Product Type Classification	http://www.haveman.org
GDP and GDP per capita	World Development Indicators (World Bank Statistics)
Consumer Price Index	Bureau of Labor Statistics at http://www.bls.gov/cpi/
US Industry Level Tariffs	Calculated from Robert Feenstra's data collection at the 4-digit SITC level
Real Exchange Rates	US Department of Agriculture's Economic Research Service at http://www.ers.usda.gov/Data/exchangerates/
Unit Values and Ad Valorem Transportation Costs	Calculated from the product level import data from http://data.econ.ucdavis.edu/international/

Figure 1 - Survival Functions for Rauch's Product Classification

7-digit TSUSA

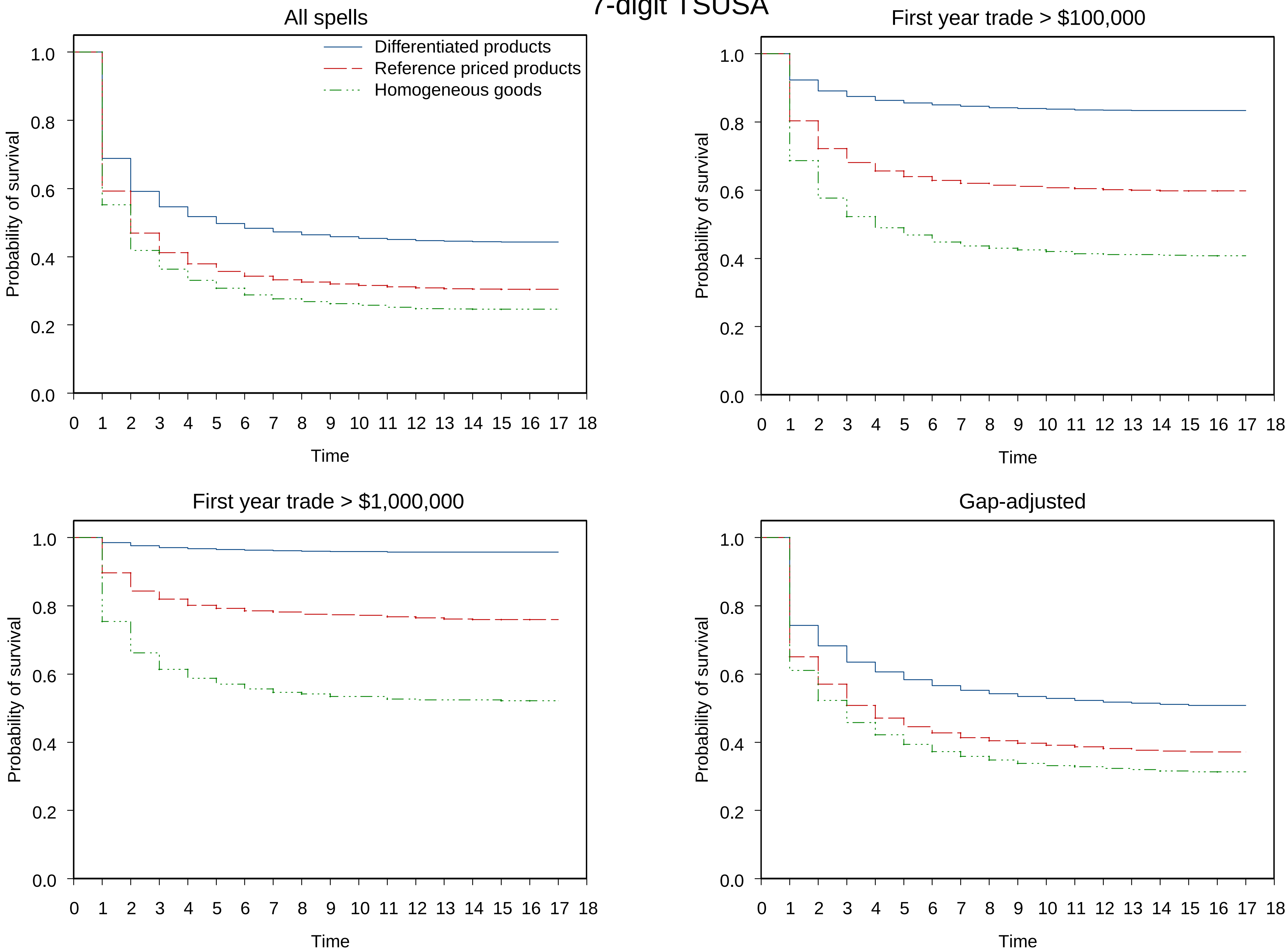

Figure 2 - Robustness Check Survival Functions

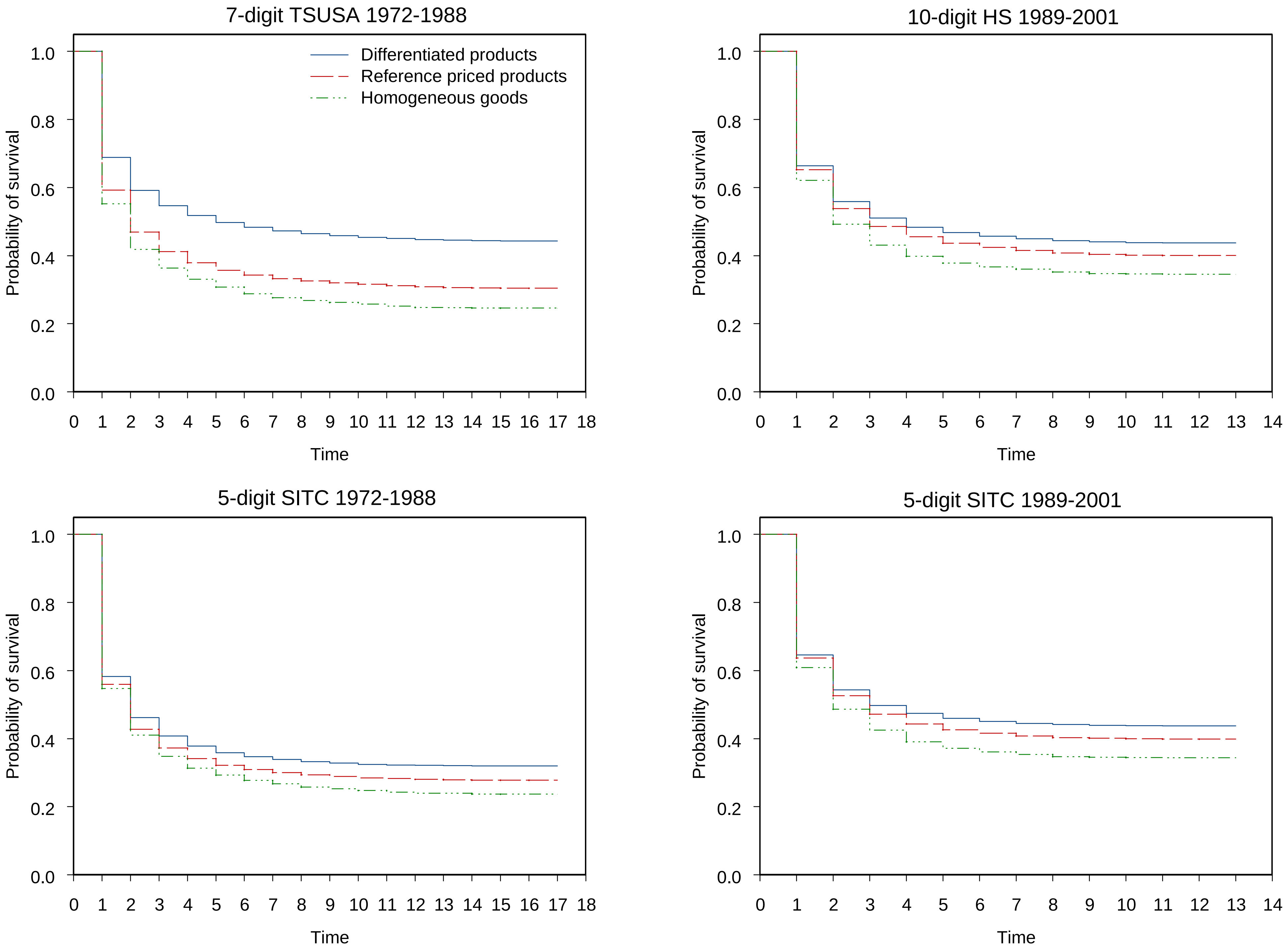

Table 1 - Analysis of Starting Size of Relationships (Known Starts)

		Product Level (TS & HS)						Industry Level (SITC 5 digit)					
		1972-1988			1989-2001			1972-1988			1989-2001		
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
OLS	α	-5.32	-2.68	-2.45	-5.27	-2.14	-2.38	-5.09	-2.40	-2.76	-5.08	-0.97	-1.83
	β	1.55	1.05	0.33	1.44	0.93	0.24	1.64	1.14	0.35	1.16	0.80	0.25
	γ		0.34	0.30		0.33	0.23		0.35	0.29		0.33	0.23
	δ		-0.54	-0.46		-0.63	-0.43		-0.62	-0.51		-0.82	-0.56
25ptile LAD	α	-6.72	-4.96	-4.37	-6.54	-3.79	-4.19	-6.58	-5.12	-4.70	-6.34	-3.56	-4.11
	β	0.65	0.50	0.10	0.77	0.55	0.11	0.71	0.52	0.07*	0.51	0.31	0.10*
	γ		0.20	0.18		0.27	0.13		0.19	0.15		0.21	0.12
	δ		-0.32	-0.28		-0.56	-0.24		-0.29	-0.19		-0.56	-0.26
50ptile LAD	α	-5.58	-2.86	-2.52	-5.47	-2.37	-2.56	-5.37	-2.87	-3.26	-5.32	-1.66	-2.39
	β	1.42	1.03	0.27	1.19	0.80	0.20	1.58	1.18	0.34	0.98	0.68	0.28
	γ		0.33	0.30		0.31	0.22		0.32	0.27		0.28	0.21
	δ		-0.55	-0.47		-0.63	-0.42		-0.58	-0.43		-0.70	-0.44
75ptile LAD	α	-4.17	-1.07	-0.84	-4.11	-0.69	-0.96	-3.88	-0.30*	-1.02	-4.02	0.75	-0.37*
	β	2.12	1.39	0.41	1.60	1.10	0.28	2.41	1.61	0.43	1.41	1.09	0.31
	γ		0.44	0.39		0.37	0.31		0.44	0.37		0.38	0.29
	δ		-0.67	-0.56		-0.69	-0.53		-0.87	-0.68		-0.94	-0.67
	Country FE	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
	Year FE	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
	Industry FE	N	N	Y	N	N	Y	N	N	Y	N	N	Y
Number of observations:			295,373			376,438			88,740			64,510	

Note: * denotes estimates not significant at the 1% level

Table 2 - Kaplan-Meier Survival Rates

	Data		Differentiated Products			Reference Priced Products			Homogeneous Goods		
			Year 1	Year 4	Year 12	Year 1	Year 4	Year 12	Year 1	Year 4	Year 12
1972-1988	7-digit TSUSA	Benchmark	0.69	0.52	0.45	0.59	0.38	0.31	0.55	0.33	0.25
		Obs>$100,000	0.92	0.86	0.83	0.80	0.66	0.60	0.69	0.49	0.41
		Obs>$1,000,000	0.99	0.97	0.96	0.90	0.80	0.76	0.75	0.59	0.52
		Gap-adjusted	0.74	0.61	0.52	0.65	0.47	0.38	0.61	0.42	0.32
	5-digit	SITC	0.58	0.38	0.32	0.56	0.34	0.28	0.55	0.31	0.24
	4-digit	SITC	0.59	0.39	0.34	0.58	0.36	0.31	0.56	0.34	0.27
1989-2001	10-digit HS	Benchmark	0.66	0.48	0.44	0.65	0.46	0.40	0.62	0.40	0.35
		Obs>$100,000	0.92	0.85	0.83	0.86	0.75	0.71	0.76	0.59	0.55
		Obs>$1,000,000	0.98	0.96	0.96	0.93	0.86	0.85	0.79	0.65	0.63
		Gap-adjusted	0.73	0.58	0.52	0.71	0.55	0.48	0.68	0.49	0.41
	5-digit	SITC	0.65	0.47	0.44	0.64	0.44	0.40	0.61	0.39	0.34
	4-digit	SITC	0.66	0.50	0.46	0.66	0.48	0.44	0.62	0.42	0.38

Note: The survival functions across the product types within each dataset are statistically significant at the 1% level using the logrank test

Table 3 - Cox Proportional Hazard Estimates for 1972-1988 7-digit TSUSA Data

		Benchmark	Obs>$100,000	Obs>$1,000,000	Gap-adjusted
Ad-valorem transportation cost	(unit = 10%)	1.068	1.039	1.048	1.071
GDP	(unit = $100bil)	0.946	0.940	0.906	0.986
Tariff rate, 4-digit SITC	(unit = 1%)	0.979	0.945	0.877	0.976
%Δ relative real exchange rate	(unit = 10%)	0.906	0.897	0.941	0.878
Coefficient of variation of unit values		0.927	0.864	0.922	0.915
Multiple spell dummy		1.495	2.254	2.366	1.487
Agricultural goods		1.040	0.949*	0.734	1.033
Reference priced products		1.173	1.594	2.551	1.206
Homogeneous goods		1.226	1.712	2.819	1.257
Observations		1,140,896	356,141	128,871	1,140,945
No. Subjects		444,378	85,629	26,236	386,191

Country fixed effects inlcuded but not reported
Note: * denotes estimates not significant at the 1% level

Table 4 - Product-type Specific Cox Proportional Hazard Estimates for 1972-1988 7-digit TSUSA Data

	Differentiated Products	Reference Priced Products	Homogeneous Goods
Ad-valorem transportation cost	1.081	1.033	1.038
GDP	0.947	0.957	0.969
Tariff rate, 4-digit SITC	0.980	0.974	0.990
%Δ relative real exchange rate	0.898	0.922	0.929
Coefficient of variation of unit values	0.900	0.994*	1.049
Multiple spell dummy	1.662	1.155	1.016*
Agricultural goods	1.134	1.028	1.148
Reference priced products	1.317		
Homogeneous goods	1.238		
Observations	1,140,896		
No. Subjects	444,378		

Country fixed effects inlcuded but not reported
Note: * denotes estimates not significant at the 1% level

Table 5 - Product Type Dummy Estimates for Both Periods and Various Specifications

	Product Level				Industry Level	
7-digit TS Data 1972-1988	Benchmark	Obs> $100,000	Obs> $1,000,000	Gap-adjusted	5-digit SITC	4-digit SITC
Reference priced products	1.173	1.594	2.551	1.206	1.226	1.192
Homogeneous goods	1.226	1.712	2.819	1.257	1.273	1.324
10-digit HS Data 1989-2001						
Reference priced products	1.063	1.260	1.679	1.088	1.296	1.229
Homogeneous goods	1.082	1.408	2.241	1.105	1.503	1.435

Country fixed effects inlcuded but not reported
Note: * denotes estimates not significant at the 1% level

Table 6 - Product Type Dummy Estimates for 1-digit SITC Industries for Product Level Data in Both Periods

7-digit TS Data 1972-1988	SITC=0	SITC=1	SITC=2	SITC=3	SITC=4	SITC=5	SITC=6
Reference priced good	1.113	1.151*	1.023*	0.929*	1.095*	1.128	1.050
Homogenous good	1.185	1.583	1.197	0.869*	1.213	1.314	1.194
Observations	93,072	13,511	48,429	6,541	3,929	95,622	356,572
No. Subjects	33,128	4,649	18,580	2,384	1,613	31,680	137,676
10-digit HS Data 1989-2001							
Reference priced good	1.137	1.196	1.121	0.957*	1.210*	1.087	1.088
Homogenous good	1.167	1.229*	1.268	0.815	1.250	1.204	1.261
Observations	163,848	16,058	70,479	13,496	6,076	192,476	570,288
No. Subjects	58,123	5,362	24,690	4,739	2,123	59,094	182,793

Country fixed effects inlcuded but not reported
Note: * denotes estimates not significant at the 1% level