

# The Importance of Being Different: Product Differentiation and the Duration of Trade

Tibor Besedes\*  
Rutgers University

Thomas J. Prusa†  
Rutgers University and NBER

December 2002

## Abstract

This paper studies how product differentiation affects the dynamics of US import trade. Using nonparametric and semiparametric techniques we find that trade in homogenous and reference priced products is of much shorter duration than trade in differentiated products. Differentiated products have a median survival time of five years; reference priced goods and homogenous products have a median survival time of only two years. As compared with differentiated products the hazard rate for reference priced goods is 16 percent higher and for homogenous goods 20–24 percent higher. Importantly, we find that these results are not driven by small trade relationships. To the contrary, we find that the differences in survival rates among the product types increase when we drop the smallest trade relationships. If we limit our analysis to those observations whose trade value in the first year is greater than \$100,000 we find that the hazard rate for reference priced goods is 50 percent higher and for homogenous goods is 60–85 percent higher. These results provide compelling evidence that trade in differentiated goods is different than trade in homogenous goods.

---

\*Department of Economics, New Jersey Hall, Rutgers University, New Brunswick, NJ 08901, besedes@economics.rutgers.edu

†Department of Economics, New Jersey Hall, Rutgers University, New Brunswick, NJ 08901, prusa@economics.rutgers.edu

## 1. Introduction

One of the benefits of increased international economic integration is increased competition. Greater market competition can lead to more rapid innovation, lower prices, and efficiencies due to specialization. Typically, however, it is other consequences of international competitive pressures—closed factories and lost jobs—that are featured in news stories and are the source of the unhappiness and consternation surrounding liberal trade policies.

Michael Porter (1985, 1990) argues that countries and businesses can best harness the positive aspects of international competition while minimizing its potential negative consequences by creating what he refers to as “competitive advantage.” In simplest terms, Porter’s view is that operational efficiency by itself is insufficient for a firm (and by extension, a country) to successfully compete in the international marketplace. Rather, it is necessary for a firm to carve out a distinct place in the market.

In this paper we offer a systematic study of competition in the US import market and we are able to identify how being “different” affects a firm’s (country’s) ability to compete for sales to the United States.<sup>1</sup> While our empirical tests are not designed to directly test Porter’s competitive advantage thesis, we nevertheless find striking confirmation that being different (i.e., selling a differentiated product) is valuable.<sup>2</sup>

In order to examine competition in import markets we focus on “trade relationships,” by which we define as a given country selling a given product to the US. When a country begins selling the product to the US, the relationship starts; when the country stops selling the product to the US, the relationship ends. We find this an extremely useful way to look at the data as the length of a trade relationship is a measure of the exporting country’s ability to successfully compete in the international marketplace.

Our findings in this paper extend those in Besedeš and Prusa (2002) where we found that regardless of aggregation or source country most US trade relationships are typically of very short duration—over half of US import trade observations are for a single year.

---

<sup>1</sup>To date, the empirical support for Porter’s view has been based on case studies.

<sup>2</sup>Of course, Porter is neither alone nor the first to argue that trade and competition in differentiated products is different from that involving homogenous products. In a series of seminal papers Paul Krugman (1979, 1980, 1981) was the first to rigorously model the economics behind trade in differentiated products. Subsequently, Helpman and Krugman (1985) offered a unified framework for understanding increasing returns and imperfect competition.

In this paper we document exactly how short depends significantly on the type of product being examined. Intuitively, one would expect that trade relationships involving homogenous products could be quite fragile. For products such as corn, wheat, and oil one can imagine a “world market” where all foreign suppliers ship their product and where all buyers purchase. In this case, trade relationships might be very short as the identity of the source country may be irrelevant. On the other hand, for differentiated products where relationship-specific investments matter and where the source country may be important the trade relationships would be expected to be far longer lived.

To motivate our empirical investigation we present two standard models of trade with differentiated goods (section 2). Both models imply that differentiated goods will be less sensitive to changes in the economic environment (e.g., cost or tariff changes) and have longer duration.<sup>3</sup>

With the economic motivation established, we then turn to US import statistics to examine the duration of trade relationships (sections 3–5). Our analysis is conducted at the most disaggregated level, the 7-digit line-item tariff level. On average, this means that in each year we observe import trade for about 10,000 products sourced from about 150 countries. Using Rauch (1999) we are able to classify the highly disaggregated import data into three product types: differentiated, reference priced, and homogenous.

The results provide support for the predictions of the theoretical models: differentiated products tend to have the longest survival, followed by reference priced products and then homogenous products. Differentiated products have a median survival time of five years; reference priced goods and homogenous products have a median survival time of only two years. This is an extraordinarily short life expectancy. We also find that compared with differentiated products the hazard rate for reference priced goods is 16 percent higher and for homogenous goods 20–24 percent higher. Put another way, trade relationships involving homogenous products are much more likely to fail than those involving differentiated products. In this sense, our results provide stark empirical support for Porter’s thesis.

In section 6 we investigate the robustness of our results. We find that these results are not driven by small trade relationships. To the contrary, we find that the differences

---

<sup>3</sup>There is a large literature on trade and growth, but the papers in this literature have not focused on the dynamics of homogenous and differentiated products trade. See Grossman and Helpman (1991) for a summary.

in survival rates among the product types increase when we drop the smallest trade relationships. If we limit our analysis to those observations whose trade value in the first year is greater than \$100,000 we find that the hazard rate for reference priced goods is 50 percent higher and for homogenous goods is 60–85 percent higher.

We also investigate measurement error in defining a trade relationship, by which we mean that we incorrectly infer that a trade relationship has ended. We adjust duration length by not inferring that short gaps in service imply the end of a relationship. By allowing a trade relationship to survive with short gaps in service, we convert two (or more) short spells of service into one longer spell of service. We nevertheless continue to find a significant difference between product types—reference priced goods have an 18 percent higher hazard rate while homogeneous goods have a 27 percent higher hazard rate than differentiated goods.

## **2. The Importance of Being Different: Theory**

In this section we use standard models of product differentiation to motivate our empirical exercise. We begin with oligopoly models that take firms’ geographical locations as exogenous and emphasize the strategic interaction among these firms. In these models transportation costs (distance) play a key role in differentiating the products. In addition, the strategic interaction among firms can result in firms exiting the market. These models imply the greater is the differentiation between products, the less sensitive are the supplying firms to changes in costs and tariffs.

We also use a standard model of trade with increasing returns and monopolistic competition. This model utilizes Dixit-Stiglitz preferences thereby abstracting from strategic issues but nevertheless allowing us to identify the equilibrium number of firms in the market and to show that changes in costs have a greater impact on homogenous product industries than in differentiated product industries.

### **2.1. Spatial Competition**

Begin by considering a version of Hotelling’s linear city model applied to trade. In our application we assume that domestic consumers of mass  $N$  are uniformly distributed along the  $z \in [0, 1]$  interval. A foreign firm is located at each endpoint with firm  $A$  located at the left endpoint. Given the nature of our dataset, it is convenient to think

of each firm being located in a different foreign country.<sup>4</sup> Hence, domestic consumers can import the product from either country  $A$  or country  $B$ .

In models of spatial competition, transportation costs serve to differentiate the products. For simplicity we will assume that transportation costs are linear in distance shipped.<sup>5</sup> Letting  $d$  denote the marginal cost of shipping the product, it follows that the products are less substitutable the larger is  $d$ .

The products are produced with constant marginal cost,  $c_j > 0$ ,  $j = A, B$ . Every consumer wants at most one unit and derives a gross benefit of  $v$  from its consumption. For a consumer located a distance  $x$  from firm  $j$  the total cost of buying is  $t_j p_j + dx$ , where  $t_j$  is one plus the ad valorem tariff. Without loss of generality let firm  $A$  have lower effective costs, i.e.,  $t_{ACA} \leq t_{BCB}$ .

Assuming that all consumers obtain strictly positive surplus by purchasing the good from one of the two firms, one can show that the location of the consumer indifferent between the two firms is

$$z = \frac{3d + t_{BCB} - t_{ACA}}{6d}. \quad (1)$$

When the firms are symmetric ( $t_{BCB} = t_{ACA}$ )  $z = 1/2$  and each firm supplies half of the market. For a given value of  $d$ , firm  $B$ 's sales decrease as its tariff or its costs increase. In other words, firm  $B$  loses market share as it becomes less competitive. Similarly, for a given value of  $d$ , firm  $B$ 's sales decrease as firm  $A$ 's tariff or costs decrease. In other words, firm  $B$  loses market share as its rival becomes more competitive. In addition, note that the larger is  $d$  the smaller is the impact of a given change in costs or tariffs. Put another way, tariff and cost changes have a smaller impact on each firms's sales the greater is the product differentiation.

Now, suppose that each firm must incur fixed costs of production,  $f_j \geq 0$ , to produce the good. As a result firm  $B$  will not service the domestic market if

$$t_{BCB} > t_{ACA} + 3d - (18f_B d)^{1/2}.$$

Clearly, if  $c_B$  (or  $t_B$ ) become sufficiently large, firm  $B$  will exit the market. Em-

---

<sup>4</sup>While the trade data is extremely disaggregated at the product level, we only observe the source country of imports. The identity of the specific firm selling the product is not reported.

<sup>5</sup>It is easy to demonstrate that a model with quadratic transportation costs yields similar results.

pirically, this means that we might observe firm  $B$  active in the market for a period of time and then, after a cost change, exit the market. Note that the greater the product differentiation, the greater is the increase in cost before we will observe firm  $B$  exiting the market. Taken together these comparative static results suggest that we should find that firms competing in differentiated product industries are less sensitive to tariff and cost changes and hence will experience less exit.

These insights can be extended to the case when there are more than two firms using the model of Salop (1979). In this case it is assumed that consumers are located uniformly on a circle with perimeter 1 and  $N$  firms could potentially service the market. Once again we think of each firm as a foreign firm/country supplying the domestic market, where different consumers prefer to source their imports from different countries.

In the Salop model one can show that the greater is the product differentiation the greater is the profit margin. As was the case in the two-firm linear model, this suggests that we should observe less exit in product markets with greater product differentiation.

## 2.2. Monopolistic Competition

A standard model of trade with increasing returns to scale and monopolistic competition can also be used to gain insight about how the length of trading relationships varies with the elasticity of substitution.

Let there be  $N$  countries and  $M$  sectors, where each sector has a large number of product varieties,  $n_m < N$ . For simplicity (and for consistency with our data), we will assume each country produces only one variety and so we will associate countries with varieties. Domestic consumer preferences are represented by a Cobb-Douglas utility function

$$U = \prod_{i=1}^M Q_m^{\gamma_m},$$

where  $\gamma_m$  is the share of expenditures on sector  $m$  and  $Q_m$  is a composite of symmetric imported product varieties in sector  $m$  given by

$$Q_m = \left( \sum_{i=1}^{n_m} x_{mi}^{\frac{\sigma_m-1}{\sigma_m}} \right)^{\frac{\sigma_m}{\sigma_m-1}}, \quad \sigma_m > 1,$$

where we have assumed the subutility function has the CES form. The parameter  $\sigma$  is the elasticity of substitution between varieties. The smaller is  $\sigma$  the greater is the extent of product differentiation (Dixit and Stiglitz, 1977).

We can write the budget constraint for a representative domestic consumer as

$$\sum_{i=1}^M P_i Q_i \leq I.$$

We allow for iceberg transport costs in shipping goods between countries and for import tariffs, so the consumer's price is  $P_i = p_i \tau_i d_i$ , where  $p_i$  is the price the foreign firm earns,  $\tau_i$  is one plus the ad valorem tariff and  $d_i$  represents transportation costs.

Each producer (country)  $i$  is identified by good  $i$ . The production of a differentiated good  $i$  involves a fixed cost  $f_m$  and a constant marginal cost  $c_m$ . Because of the fixed costs the number of differentiated goods actually being consumed is far less than the number of potential differentiated goods,  $n_m < N$ . As is usually done, we will assume that  $n_m$  is large and that free entry guarantees zero profits.

If we assume symmetry and solve three conditions simultaneously (profit maximization, zero profits, and utility maximization) we can solve for the equilibrium solutions,

$$p_m = c_m \sigma_m / (\sigma_m - 1), \tag{2}$$

$$x_m = (\sigma_m - 1) f_m / c_m, \tag{3}$$

$$n_m = \gamma_m I / (\tau_m d_m f_m \sigma_m) \tag{4}$$

As has been noted elsewhere (Neary, 2000), the price and quantity are functions of costs and the elasticity of substitution. Changes in other parameters (e.g., tariffs and transportation costs) lead to adjustments in the number of foreign firms (countries) only.

Now consider two industries where products in industry  $j$  are more differentiated than those in industry  $i$ ,  $\sigma_j < \sigma_i$ . From (4) it follows that industry  $i$  is more sensitive to changes in tariffs, transportation costs, and fixed costs. Therefore, we expect changes in these variables to lead to more exit in homogenous good industries than in differentiated

product industries.<sup>6</sup>

### 3. Data

Our analysis is based on US import statistics as reported in the NBER Trade Database (NBERTD) compiled by Feenstra (1996).<sup>7</sup> From 1972 through 1988 import products are classified according to the seven digit Tariff Schedule of the United States (TS). Beginning in 1989 imports have been classified according to the ten digit Harmonized System (HS). Given that all products were recoded in 1989, we limit our analysis to the period 1972-1988 in order to avoid the problematic concordance issue.<sup>8</sup> In simplest terms, for each product and year we can identify all of the countries from whom the United States purchased the product.

Unless explicitly noted, all analysis is done at the most disaggregated level, the TS 7-digit level. On average, this means that in each year we observe import trade for about 10,000 products sourced from about 150 countries. We are interested in studying the length of time until a country no longer exports a product to the US, an event that we will refer to as a “failure.” As a result calendar time is not as important as analysis time, which measures the length or duration that a country exports a product to the US. Hence, for each product and country we use the annual data to create *spell* data. By this we mean that if we observe trade in product  $i$  from country  $c$  from 1976–1980, we would say that the  $ci^{\text{th}}$  trade relationship has a spell length of five. Thinking of imports in terms of spells our dataset has 693,963 observations. The observations have a median (mean) spell length of 4 (2.7) years.

For duration analysis we believe that highly disaggregated data is a necessity. First of all, we need the data to be sufficiently disaggregated to allow us to identify the extent of product differentiation. The more aggregated is the data, the more we will be identifying industries rather than products and many different types of products can be

---

<sup>6</sup>The substitution parameter also plays a key role in Melitz’s (2002) model with heterogenous firms. While Melitz’s primary interest is how domestic firms respond to trade, his model implies that changes in tariffs, transportation costs, and fixed costs have greater impact in homogenous good industries than in differentiated product industries.

<sup>7</sup>We only study imports because there is no concordance between the disaggregated import and export codes. In addition, the export data is self-reported making it more likely that exports are misreported.

<sup>8</sup>Many TS products are mapped into multiple HS codes and vice-versa. This makes it almost virtually impossible to discern actual exits from a concordance exit.

sold by a single industry. Hence, industry level analysis will make it more difficult to identify the role of product types.

In addition, if the products are too broadly defined, we cannot expect to see any source countries exit the data. For instance, at the extreme if we aggregate all imports from each country we will never observe any exit since the United States purchases some product from each source country every year.

Once we began to think of trade data in terms of spells it became apparent that we needed to account for censoring in our analysis. By this we mean that we often do not know whether a trade relationship ended because of failure or for some other reason. In practice this means that we do not know for certain the beginning or ending date for some trade relationships.<sup>9</sup>

As it turns out, censoring is pretty common in US import data—about half of the trade observations are censored. The censoring problem comes in several flavors. First, we have no information on trade relationships for the years before the beginning and after the end of our sample. Suppose, for example, that we observe that the US imported corn from the Philippines in 1972 and that this relationship was observed for  $x$  years. The Philippine trade relationship may have begun in 1972 (i.e., a trade relationship of exactly length  $x$ ) or it may have begun in some prior year (i.e., a trade relationship of length greater than  $x$ ). Given the nature of the data we cannot determine the true length of the Philippine spell. As such, we feel the most appropriate interpretation is that this relationship had a duration of at least  $x$  years. Similarly, suppose we observe that the US imported corn from India in 1988 and also for the  $y - 1$  preceding years. Unfortunately, the TS data does not continue beyond 1988 and hence we cannot be sure how long that spell ultimately lasted. Once again, we feel the most appropriate interpretation is that this relationship had a duration of at least  $y$  years.<sup>10</sup>

Another common cause of censoring is due to reclassification. US Customs is constantly revising and adjusting the product definitions for the TS and HS codes. As an example, in the early 1970s, a single product code was sufficient to identify the trade

---

<sup>9</sup>There are a number of important issues (including the aggregation issue) involved in applying duration analysis to import data. Most of the issues are not central to the primary focus in this paper (the impact of product differentiation). Besedeš and Prusa (2002) contains an extensive discussion of the import data and its applicability to duration analysis.

<sup>10</sup>Note that both examples are said to be right censored even though in the Philippine example it is the left side of the spell that is not observed. However, in duration analysis left censoring is defined as knowing that the subject failed before it came under observation.

in semiconductors. Over time too many types of semiconductors were being accounted for by a single code and the need arose for more codes (e.g., 4K and 16K DRAMs). Once new codes are introduced, the original one is no longer being used and trade in that product is no longer observed. But, it is not clear whether this is due to the end of the relationship or because of the redefinition. Such exits are classified as censored. The interpretation of a censored spell is then that the relationship lasts at least  $\alpha$  years. Trade relationships in newly defined products are also censored because it is not known whether this is truly a new relationship or just a redefinition.

Given how US Customs collects the data, censoring cannot be ignored. As explained above, our approach has been to classify a trading relationship as censored whenever there is uncertainty about the length of the trading relationship. As a result, our results may understate the actual number of exits and hence overstate the true duration of the typical trading relationship. The alternative would be to classify these trading relationships as “exits,” which we believe would more seriously distort what is actually occurring in the data.

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

Although coarser than one would like, Rauch’s classification scheme has several virtues. First and foremost, it is the only classification scheme that we know of that exists at a disaggregated level. In a first best world, we would know each product’s elasticity of substitution. Unfortunately, such elasticities are generally only available for 2-digit industries and a handful of 3-digit industries. Given the requirement that we use highly disaggregated data in our analysis, Rauch’s scheme is preferred. Second, Rauch’s scheme makes intuitive sense since it broadly captures what economists mean by product substitutability. Products like corn, oil, wheat, etc., that are sold on organized exchanges are exactly those products typically cited as being homogenous. Consumers may neither know nor care about the source of the product they are purchasing. On the other hand, products like many types of steel and chemicals whose prices are listed in industry guides and trade journals will likely have some unique attributes (e.g., quality

may vary by source country) but are essentially substitutable. In this case, consumers will know the source country, but this may only have a small impact on their purchasing decision. In the final category are differentiated products. These are products that not only have many characteristics that vary across suppliers but may even be specifically tailored to the end-user’s needs. Automobiles are perhaps the often cited example of this type of good; in fact, most consumer goods (e.g., toys, apparel, cookware) are classified as differentiated. Third, Rauch’s classification scheme is comprehensive, covering about 98 percent of all US import products.<sup>11</sup> Thus, our analysis is quite comprehensive. There is no selection bias as might be the case if we limited our analysis to only those products where substitution elasticities exist.

#### 4. Modelling Duration

In this paper we are interested in modelling the duration of exports to the US, or to be more precise the likelihood that a country will stop exporting a product in the next short interval of time given that it has been exporting that product until time  $t$ . One approach to this could be to assume that conditioned on a set of covariates  $x$  time  $t$  has a normal distribution and then proceed to characterize the probability distribution of observed times. There are several reasons why such a classical estimation approach is not desirable. For one, the duration time analyzed is positive by construction and a variable following the normal distribution can take negative values. Additionally, as stated above, we are not interested in the probability distribution of observed durations, but rather a conditional probability of failure, which is given by the hazard rate. Hence, modeling of duration requires us to model the hazard rate as the main object of interest.

Let  $T$  be a nonnegative random variable denoting time to failure, or what is referred to as survival time. Instead of focusing on cumulative density and probability density functions of  $T$ ,  $F(t)$  and  $f(t)$ , respectively, the duration analysis focuses on survival and hazard functions,  $S(t)$  and  $h(t)$ , respectively. The survival function is defined as the reverse cumulative density

$$S(t) = 1 - F(t).$$

---

<sup>11</sup>After dropping those TS codes that are not classified by Rauch, we have 673,885 trade relationships.

This is the probability of surviving longer than  $t$  periods. The hazard function,  $h(t)$ , of survival time  $T$  gives the conditional failure rate. It is defined as the probability that a failure occurs in a small interval of time, conditional on the subject having survived to the beginning of the interval,

$$h(t) = \lim_{\Delta t \rightarrow 0} \Pr(t + \Delta t > T > t \mid T > t) \Delta t = \frac{f(t)}{S(t)}.$$

Since time in our analysis is discrete, the relevant expressions are as follows. Let  $T$  be a discrete random variable taking on values  $t_i$ ,  $i = 1, 2, \dots, n$  with probability density function  $p(t_i) = \Pr(T = t_i)$ ,  $i = 1, 2, \dots, n$  where  $t_1 < t_2 < \dots < t_n$ . The survival function for a random variable  $T$  is given by

$$S(t) = \Pr(T > t) = \sum_{t_i > t} p(t_i).$$

The hazard function is

$$h(t_i) = \Pr(T = t_i \mid T \geq t_i) = \frac{p(t_i)}{S(t_{i-1})}, \quad i = 1, 2, \dots, n$$

where  $S(t_0) = 1$ . The survival and hazard functions are related through the following expression

$$S(t) = \prod_{t_i < t} [1 - h(t_i)].$$

The main objective of the duration analysis is then to estimate the survival function to characterize duration of an event and the hazard function to evaluate the impact of different factors (covariates) on the hazard (and hence survival) function.<sup>12</sup> We will estimate both the survival and hazard functions first nonparametrically to characterize the duration of countries' exports to the US. We will follow then with semiparametric estimation of the hazard function to examine what impact different factors have on duration.

---

<sup>12</sup>Keifer (1988) provides a good survey of duration models, while Hosmer and Lemeshow (1998), Lancaster (1990), and Lee (1992) provide good introductions to survival analysis.

#### 4.1. Nonparametric Estimation

The survivor function  $S(t)$  is usually estimated nonparametrically using the Kaplan-Meier product limit estimator. Derivation is as follows. Assume a sample contains  $n$  independent observations denoted  $(t_i; c_i)$ ,  $i = 1, 2, \dots, n$ , where  $t_i$  is the survival time, while  $c_i$  is the censoring indicator variable  $C$  (taking on a value of 1 if failure occurred, 0 otherwise) of observation  $i$ . Assume there are  $m < n$  recorded times of failure. Denote the rank-ordered survival times as  $t(1) < t(2) < \dots < t(m)$ . Let  $n_j$  denote the number of subjects at risk of failing at  $t(j)$  and let  $d_j$  denote the number of observed failures. The Kaplan-Meier estimator of the survival function is then

$$\hat{S}(t) = \prod_{t^{(i)} < t} \frac{n_j - d_j}{n_j},$$

with the convention that  $\hat{S}(t) = 1$  if  $t < t(1)$ . Given that many of our observations are censored, we note that the Kaplan-Meier estimator is robust to censoring and uses information from both censored and non-censored observations.

In the discrete case, the hazard function can be estimated by

$$\hat{h}(t_j) = \frac{d_j}{n_j}.$$

#### 4.2. Semiparametric Estimation

The Cox proportional hazards regression model (Cox 1972) asserts that the hazard rate is

$$h(t, x, \beta) = h_0(t) \exp(x\beta),$$

where  $x$  denotes a vector of explanatory variables and the  $\beta$ s are to be estimated. The baseline hazard,  $h_0(t)$ , characterizes how the hazard function changes as a function of time. It captures individual heterogeneity that is not explained by the covariates.

In the Cox proportional hazards model one subject's hazard is a multiplicative replica of another's; for instance, the ratio of subject  $j$  to subject  $m$ 's hazard is

$$HR(t, x_m, x_j) = h(t, x_j, \beta) / h(t, x_m, \beta) = \frac{h_0(t) \exp(x_j \beta)}{h_0(t) \exp(x_m \beta)} = \frac{\exp(x_j \beta)}{\exp(x_m \beta)}.$$

One particular application of this property is to compare one subject with covariates  $\{x_1, x_2, \dots, x_k\}$  with another with the same covariates, except that  $x_2$  is incremented by 1. Taking the ratio of the hazards of these two subjects, we have

$$h(t, x_1, x_2 + 1, \dots, x_k, \beta) / h(t, x_1, x_2, \dots, x_k, \beta) = \exp(\beta_2).$$

Note that this ratio is not a function of  $t$ . Hence, exponentiated individual coefficients have the interpretation of the ratio of the hazards for a one-unit change in the corresponding covariate.

A particular advantage of the Cox model is that the baseline hazard is left unspecified and is not estimated. This allows us to account for unobserved heterogeneity in a more tractable approach than the Kaplan-Meier estimator<sup>13</sup> and a less restrictive approach than a fully parametric model.<sup>14</sup> To see that the baseline hazard is not needed for estimation note the following derivation of the partial likelihood function to be maximized. Let the risk set  $R(t_i)$  be defined as the set of individuals who are alive at time  $t_i$ , or alternatively the individuals who can die at time  $t_i$ . The probability that one individual dies at time  $t_i$  with covariates  $x_i$  given that one of the individuals in the risk set  $R(t_i)$  dies, is given by

$$\begin{aligned} & \Pr[\text{individual dies at } t_i \mid \text{one death at } t_i] \\ &= \frac{\Pr[\text{individual dies at } t_i \mid \text{survival to } t_i]}{\Pr[\text{one death at } t_i \mid \text{survival to } t_i]} \\ &= \frac{h(t_i | x_i)}{\sum_{j \in R(t_i)} h(t_i | x_j)} = \frac{h_0(t_i) \exp(x_i \beta)}{\sum_{j \in R(t_i)} h_0(t_i) \exp(x_j \beta)} \\ &= \frac{\exp(x_i \beta)}{\sum_{j \in R(t_i)} \exp(x_j \beta)} \end{aligned}$$

---

<sup>13</sup>The Kaplan-Meier estimator is largely immune to the heterogeneity problem, but is limited in the amount of information that can be extracted from it.

<sup>14</sup>Parametric estimation of proportional hazards models involves making a distributional assumption about the baseline hazard function (exponential and Weibull distributions are among the most popular ones). Modelling heterogeneity in such a setting involves making an assumption about the distribution of unobserved heterogeneity (the Gamma distribution is frequently used), imposing more structure on the model.

The vector of coefficients  $\beta$  is estimated by maximizing the partial likelihood function (or the log-likelihood)

$$L(\beta) = \prod_{i=1}^n \frac{\exp(x_i\beta)}{\sum_{j \in R(t_i)} \exp(x_j\beta)}. \quad (5)$$

Notice that the contribution of an individual to the likelihood function does not involve the baseline hazard.

In the derivation of the partial likelihood in (5) it is assumed that each subject dies at a different time, or in other words that there are no tied failures. Our data set, due to its size, contains a multitude of tied failures.<sup>15</sup> In order to deal with tied failures we need an alternate partial likelihood. There are a variety of approaches suggested by different authors and we use Breslow's method primarily because it is computationally fast. Breslow's method is as follows. Let  $t_1 < t_2 < \dots < t_D$  be the  $D$  distinct and ordered duration times. Let  $d_i$  be the number of deaths at  $t_i$  and  $\mathbb{D}_i$  the set of all individuals who die at time  $t_i$ . Let  $s_i = \sum_{j \in \mathbb{D}_i} x_j$  and  $R_i$  the set of all individuals at risk just prior to  $t_i$ . The adjusted partial likelihood suggested by Breslow is

$$L_1(\beta) = \prod_{i=1}^D \frac{\exp(s_i\beta)}{\left[ \sum_{j \in R_i} \exp(x_j\beta) \right]^{d_i}}. \quad (6)$$

This likelihood considers each of the  $d_i$  events at a given time as distinct, constructs their contribution to the likelihood function, and obtains the contribution to the likelihood by multiplying over all events at time  $t_i$ .

To allow for time varying covariates the basic model becomes

$$h(t, x(t), \beta) = h_0(t) \exp(x(t)\beta),$$

while (6) becomes

$$L_1(\beta) = \prod_{i=1}^D \frac{\exp(s_i(t)\beta)}{\left[ \sum_{j \in R_i} \exp(x_j(t)\beta) \right]^{d_i}}.$$

---

<sup>15</sup>The Kaplan-Meier estimator handles tied failures and censoring without any difficulties.

Different censoring patterns do not affect the partial likelihood. A subject whose spell is censored between duration  $t_k$  and  $t_{k+1}$  appears in the summation in the denominator of the contribution to the partial likelihood of (ordered, uncensored) observations 1 through  $k$ , but not in any others.

## 5. Empirical Findings

As we have discussed, our primary interest is to better understand the dynamics of import supply. The theoretical models discussed above suggest that the extent of product differentiation will affect duration. Whether one believes differentiation is due to spatial considerations or “love of variety” preferences, economic theory predicts that changes in the economic environment will have greater impact in homogenous product industries.

### 5.1. Nonparametric Results

We begin by performing some nonparametric estimates. Besedeš and Prusa (2002) show that the estimated survival function US import trade relationships is downward sloping with a declining slope and that the estimated hazard function behaves in the same way. They also show that this general description of survival and hazard functions holds across regions, countries, and industries. The implication of this results is that US import trade relationships face a large probability of failure in their first few years. The risk declines quite markedly once trade relationships last for 4-5 years. We have no reason to expect that survival and hazard functions across product types will behave any differently.

The Kaplan-Meier survival function,  $\hat{S}(t)$ , is graphed in Figure 1 for each of the three product types. The estimated survival functions across product types have the expected shape. They are similar in that each is downward sloping with a decreasing slope. More striking than their similarity in shape is how the estimated survival functions differ across product types. At each time the survival function for differentiated products is above the survival function for reference priced products, which in turn is above the survival function for homogenous products. Moreover, given the large number of observations in our sample, the differences between each of estimated survival functions is statistically significant.<sup>16</sup>

---

<sup>16</sup>The standard errors are in the range 0.001 to 0.02.

These results provide support for the predictions of the theoretical models: differentiated products tend to have the longest survival, followed by reference priced products and then homogenous products. Differentiated products have a median survival time of five years; reference priced goods and homogenous products have a median survival time of only two years. Put another way, this means that **half** of the trade relationships involving homogenous goods fail during the first two years. This is an extraordinarily short life expectancy. Not surprisingly, the chance of observing a spell of more than 10 years is far greater for differentiated products than for either reference priced goods or for homogenous products. Given the extraordinary failure rates during the first few years, however, it is surprising that we nevertheless find that 45% of differentiated products, 32% of referenced priced products, and 25% of homogenous goods survive for more than 10 years. This suggests that the failure rate drops precipitously after the first few years.

To examine this issue we plot the hazard rates for each of the three product types in Figure 2. As expected, given the results presented in Besedeš and Prusa (2002) and given that survival functions across product types have decreasing slopes, the hazard rate functions are also decreasing. Also as expected, homogenous products have the highest hazard rates, followed in order by reference priced goods and differentiated products.

What is striking is how high the hazard rates are in the first few years and how quickly the hazard rates fall to very small levels. The conditional probability of failing in the first year is remarkably high, 45 percent for homogenous goods, 41 percent for reference priced goods, and 31 percent for differentiated goods. However, by year 5, the conditional probability of failing for each type has fallen to less than seven percent; after 10 years, the hazard rate for each type has fallen to less than two percent.

In Besedeš and Prusa (2002) we concluded these results indicate that there is negative duration dependence in trade, meaning that the conditional probability of failure decreases as duration increases. Also, this result suggests the presence of a threshold—countries that are able to survive the first few years exporting to the US are likely to continue exporting for a long period of time.

## 5.2. Semiparametric Results

Tables 1–4 contain the Cox proportional hazard estimates. Our basic estimation model includes the standard regressors that appear in gravity equation estimates (e.g., GDP, language, continuity, distance). While we do not offer a theoretical model explaining how these variables affect survival, we believe it is reasonable to include gravity equation variables in our empirical model. After all, the above discussion suggests that survival is related to country size. The gravity equation states that the bilateral trade between two countries is directly proportional to the product of the countries' GDPs and therefore larger countries will tend to trade more with each other. Given the gravity equation's remarkable job predicting bilateral trade, we view the gravity-motivated regressors as an exogenous control for the propensity for source countries to supply the US market.

In addition, we include several variables to capture relative cost and competitive issues. We use the industry level tariff rate to control for the ease in which foreign firms can enter the market. Whether higher tariffs increase or decrease the hazard depends largely on whether time series or cross-section variation dominates. For a given product, an increase in the tariff should lead to some foreign firms exiting because higher tariffs raise the costs of servicing the US market. It follows that time series variation in tariffs should lead us to find higher tariffs raise the hazard. On the other hand, looking across industries higher tariffs mean less competition for those firms currently in the market. As a result both domestic and foreign firms servicing the US market face less risk and hence a lower hazard. It follows that cross section variation in tariffs should lead us to find that higher tariffs lower the hazard. Given that there is relatively little time series variation in tariffs, we expect the cross section effect to dominate which means that higher tariffs should lower the hazard.

In order to capture the impact of cost changes on the hazard we included the change in the *relative* real exchange rate. To construct this measure of relative costs, we began by defining each country's exchange rate so that an increase corresponds to a real depreciation. Then, in each year we normalized the exchange rate by the average (percentage) change of all countries' exchange rates (relative to the US dollar). This gives us a measure of how each country's exchange rate changed relative to all other countries' exchange rates. An increase in this measure reflects that 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-a-vis other foreign and domestic firms and hence less likely to exit.

Schott (2001) argues that even with data as disaggregated as 7-digit TS import data, some products are more broadly defined than one would like. In order to control for diversity of the products Schott calculated the coefficient of variation of unit values for each TS product in each year. We follow his lead and include the coefficient of variation in our specifications. We expect that the smaller the coefficient of variation the more homogenous is the product and hence the greater the hazard.

Finally, there is the issue of multiple spells—trade relationships that have multiple periods of service separated by a period with no service. Specifically, some trade relationships are observed for a period of consecutive years (spell 1), then are followed by a period of no trade, and then again observed for another service spell (spell 2). We believe that the first failure makes a second failure more likely (higher hazard). On the other hand, it is also possible that the fact that the foreign supplier has returned to the market is a positive sign and hence make 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 in our estimation. In our base specifications we include spells after the first and control for multiple spells by using a dummy (=1). Alternatively, we can estimate the hazard rate using data only on first spells. We will consider alternative methods for addressing multiple spells in section 6.

In the first two columns of Table 1 we report our benchmark estimates based on the entire sample. Throughout we present results in terms of hazard ratios. Hence, an estimated hazard rate coefficient less than (greater than) (equal to) 1 is interpreted as implying that the variable lowers (raises) (has no impact on) the risk of failure. We report estimates with and without region dummies and find with few exceptions that the results do not vary significantly across specifications.

In many respects the estimates are consistent with those in the gravity literature. The estimated effects for the standard gravity variables, however, are generally small. In particular, our estimates imply

- Distance increases the hazard rate but by a fairly small magnitude, and in some specifications is statistically insignificant. With few exceptions (e.g. Australia) the differences in distance between the US and its trading partners would only

lead to 1–3 percent difference in hazard rates.

- Common language lowers the hazard, but again the effect is fairly small, about four percent. Countries that are contiguous with the US (Canada and Mexico) face significantly lower hazard rates—on the order of 15-25 percent (all observations).
- Larger countries (as measured by GDP) face lower hazard. Given the variance in country GDP the size of the effect depends significantly on country size. A \$100 billion increase in GDP lowers the hazard rate by about 2 percent; a \$1 trillion increase in GDP lowers the hazard rate by about 25 percent.
- Industries with higher tariffs have a lower hazard. A 10 percentage point increase in an industry’s tariff lowers the hazard by approximately 2 percent.

We do find that changes in the real exchange rate has a large impact on the hazard rate. We find that a 10 percent depreciation in the real exchange rate lowers the hazard by at least 10 percent.

We also find that products with higher variation in unit values face a significantly lower hazard rate. This supports Schott’s (2001) contention that TS product codes must viewed cautiously as a single TS code can capture differentiated products. For instance, the “cotton T-shirt” product might include commodity grade products from China and Bangladesh and also fashion designer products from Italy.

We are primarily interested in the product type estimates. Letting differentiated products be the benchmark, we find that reference priced goods have 16 percent higher hazard and homogenous goods 20–24 percent higher hazard. Thus, the estimates strongly support what Figures 1 and 2 suggested: namely, the value of being different.

The remaining columns of Table 1 contain estimates using different specifications for the exchange rate variable(s). In columns 3 and 4 we add to our benchmark specification the level of the relative real exchange rate. There are no significant changes in the estimated coefficients and we find that the higher level of the real exchange rate significantly lowers the hazard rate (about 21 percent). This result indicates that countries with weaker currencies face lower hazard rates when exporting their products to the US.

In columns 5 and 6 we separately estimate the impact of real exchange depreciations and appreciations. There are no significant changes in the estimate coefficients for the other covariates. A real depreciation lowers the hazard of exporting to the US by about 17 percent, while appreciation also lowers that hazard but by a much smaller amount, 2-4 percent. While we are not surprised by the impact of a real depreciation, we are puzzled by that of a real appreciation and it is one for which we do not have an explanation.

In Table 2 we restrict the sample to OECD countries and find that the results are fairly similar to those reported in Table 1. Most of the control variables are quite similar to those for the whole sample. The impact of distance increases threefold, while that of the coefficient of unit variation doubles. The contiguity effect is now much smaller due to the fact that OECD countries have survival experience similar to that of Canada (Japan's survival function lies completely above Canada's). One interesting difference involves the exchange rate: we find real depreciation has a much bigger effect for OECD trade. The impact of all of the exchange rate variables increases with the exception of the level of exchange rate. The main variables of interest, the product type dummies, are somewhat larger for the OECD sample: reference priced by 17 percent, and homogenous goods by 32 percent, continue to have significantly higher hazard rates as compared with differentiated products.

In Tables 3 (all countries) and 4 (OECD countries) we allow the coefficients on all variables to vary by product type. The parameter estimates confirm what we learned from the basic specification. Namely, Canada/Mexico, larger countries, high tariff products, and weaker currencies, all have lower hazard rates. Most importantly, even after allowing for systematic differences across product types, we continue to find that homogenous products face a significantly higher hazard than differentiated products.

Table 5 reports the results of tests for the equality of the product type specific estimates. Columns 1 and 4 correspond to the estimates in Tables 3 and 4. We find that except for the tariff rate, the product type specific coefficients are different from each other. Distance has a bigger impact for homogeneous goods than differentiated goods. Common language and common border also have a bigger impact on homogeneous products than on differentiated goods. Sharing a common language (border) with the US lowers transaction costs (transportation costs) and hence has a larger impact on trade in homogenous than differentiated goods, just as our models imply. The impact of

GDP is larger for differentiated goods than homogeneous goods, indicating that the size of the country is more important for trade in differentiated, rather than homogeneous goods.

Tariff rates have a slightly bigger impact on trade in differentiated goods than in homogeneous goods. The most surprising result in Tables 3 and 4 is that the impact of the exchange rate is largest for differentiated products. This is a bit surprising as the models imply that cost changes will have a bigger impact for homogenous products.

The coefficient of variation of unit values varies greatly across product types. The effects are similar in size (about 8 percent), but different in their sign. As we would expect, the estimates indicate that a higher variation of unit values for homogeneous goods will lead to a higher hazard rate (more failure) which suggest the importance of similar prices for homogenous goods. For differentiated products higher variation in unit values lowers the hazard rate.

Similarly to results in Tables 1 and 2, the different exchange rate specifications do not change the estimates significantly. The results for OECD are similar to those for all countries, except that the effects of changes in exchange rates are stronger for all products and the effect of contiguity for differentiated products is now positive. However, after including region dummies this coefficient becomes negative as we expect it to be.

Overall, our estimates show that there are significant benefits to exporting a differentiated product to the United States. Trade relationships involving reference priced products face a 17 percent higher hazard rate, while those involving homogeneous goods face a 20-24 percent higher hazard rate than those involving differentiated goods. Furthermore, after allowing for product type specific effects we not only find that homogeneous products face a higher hazard rate, but we also find that variables indicating costs of trading have a bigger impact on homogeneous goods than on differentiated goods. These results provide compelling evidence that trade in differentiated goods is different than in homogenous goods.

## 6. Robustness Results

We next perform several exercises to investigate the robustness of our results. There are three concerns we explore. First, are our results driven by differences in the value of

trade across product types? Second, are our results affected by potential measurement errors? Third, are our results distorted by our treatment of multiple spells? We now proceed to address each of these concerns in turn.

### 6.1. Small trade relationships

One concern is that our results are driven by differences in the value of trade across product types. It seems reasonable to expect the trade relationships with large values of trade to be longer lived. For instance, all else equal, one might expect an observation with \$1 million of trade (in year one) will survive longer than an observation with \$100,000 of trade (in year one). As a result, the findings reflected in Figures 1 and 2 might reflect that (i) small valued trade relationships are at greatest risk and hence (ii) given the ranking of product types, that homogenous goods tend to be small value trade relationships.

There is certainly support for the first conjecture, but less so for the second one. To show this we first look at the evidence that small valued trade relationships are at greatest risk. In Figure 3 we graph the distribution of spells lengths. On the  $x$ -axis we graph the duration of the spell, where we use the “+” to denote censored spells.

To provide a benchmark we begin plotting data on all trade observations. This benchmark distribution reflects all trade relationships and should be interpreted as an equal weighted distribution as each observation is treated symmetrically (plotted with a solid line). As shown 33 percent of the observations last only one year, and another 21 percent are censored at one year. More than 70 percent of all trade relationships are observed for less than three years. Overall, fewer than five percent of all trade relationships last more than 10 years.

One method to get a sense of whether value matters is to weigh each observation by the value of trade in year one of the spell. By doing so an observation with \$1 million of trade has 10 times the impact of an observation with \$100,000 of trade. This weighted distribution is plotted with a dashed line in Figure 3. As one can see this measure suggests that larger dollar relationships are longer lived. Weighed by value, 12.4 percent are observed for only one year; 29.9 percent are observed for less than three years—significantly less than the 70 percent attrition that we found in the unweighted sample. Looking at those relationships that last more than 10 years, long lived relationships are

about five times more likely when we weigh trade relationships by value than when we equally weigh the observations.

It strikes us, however, that weighing the observations by value overstates the role played by the largest trade observations. Given that we want to think about duration of trade relationships we believe that it is an overstatement to say that a \$1 billion crude oil purchase is 1000 times more important than a \$1 million auto parts purchase. We believe that weighing by trade value focuses too much on the biggest observations.

We feel a better way to get at the issue is to equally weigh the observations but filter out small dollar-value observations. In other words, eliminate spells whose trade in the first year is below some minimum level. This allows us to identify the role played by small value observations without overly focusing on the biggest observations. In Figure 3 we plot the distributions using this filtering scheme. We depict two alternative cut-off levels for year one trade, \$100,000 and \$1 million, but our results are not sensitive to the precise cut-off level used. As seen, imposing higher cut-offs results in higher distributions. This provides further evidence that the small valued trade relationships are at greatest risk.

This does not imply, however, that the differences in the survival functions depicted in Figure 1 only reflect size. If size were all that matters, then we should find that differentiated goods tend to involve larger trade values. Yet, the opposite is the case: differentiated goods tend to involve the smallest trade values. On average, in the first year homogenous products involve larger value transactions (\$4.5 million) than reference priced (\$730,000) or differentiated goods (\$700,000).<sup>17</sup>

Given this background in Figure 4 we plot the Kaplan-Meier survival function for each of the three product types after dropping the small-valued observations. As data in Figure 3 suggested, the survivor functions shift up as we progressively drop observations. This does indicate that small trade relationships tend to be shorter-lived. On the other hand, the estimates provide no evidence that differences in survivor functions among the product types are driven by small observations. In fact, the differences between product types appear to grow as we progressively eliminate the smaller trade observations.

We also re-estimate our basic specification using the dollar-cutoffs (Table 6). Comparing these estimates (all countries, columns 1 and 3) with those in Table 1, we find that the impact of common language, contiguity and tariff rates all increase. The impact

---

<sup>17</sup>If we compared the average value over the entire spell we continue to find the same ordering.

of distance increases at the \$100,000 cut-off level, but at the \$1,000,000 cut-off level the estimate is statistically insignificant. The impact of GDP and variation of unit values increases at both cut-offs, but the impact at the \$100,000 cut-off is greater than that at \$1,000,000 level. Interestingly, the impact of the change in exchange rate declines as we restrict ourselves to the bigger spells.

Most importantly, product type dummies indicate that the results are not driven by small trade relationships. Compared to differentiated goods, homogeneous goods face an 85 percent higher hazard rate at the \$100,000 cut-off level, and a 230 percent higher hazard rate at the \$1,000,000 cut-off level. Reference priced products face 54-160 percent higher hazard rates as we restrict our sample to larger trade value observations.

Estimates for OECD countries in Table 6 (columns 2 and 4) reveal that the impact of all variables except the exchange rate increases as we look at progressively larger valued trade relationships. The impact of changes in exchange rates first increases slightly and then decreases. Reference priced goods face 47-120 percent higher hazard rates than differentiated goods, while homogeneous goods face 88-180 percent higher hazard rates.

In Table 7 we report product-type specific estimates using the dollar-cutoff filter. Generally speaking, the coefficients increase for most variables across all product types. There is a decline in the impact of changes in exchange rate when we restrict the sample to the largest values of trade. Also, as we look at the highest valued trade relationships, variation of unit values is not significant for reference priced products, while the same is the case for GDP for homogeneous products. Looking at the tests for joint equality (Table 5) we can see that in the set of coefficients (all countries sample) for common language, shared border, tariff rate, and exchange rate changes are not statistically different across all product types. The other variables are. The unusual result that exchange rates have the biggest impact on differentiated products is no longer present. Product type dummies also increase. At the \$1,000,000 cut-off level, reference priced products face 86 percent higher hazard rates and homogeneous goods face 260 percent higher hazard rates than differentiated goods.

These same patterns hold when we look at the sample of OECD countries. Most coefficients increase in size and more of them become insignificant. For differentiated products contiguity is no longer significant at both cut-off levels; for reference priced goods distance, common language, and unit values are no longer significant at both cut-off levels; and for homogeneous goods common border is not significant at both levels,

while GDP and unit values are not significant at the \$1,000,000 level. Table 5 indicates that coefficients for tariff rates and exchange rate changes are not statistically different across product types, while the same is true for distance and common language at the \$1,000,000 cut-off level. Most importantly, product type dummies increase considerably. Reference priced products face a 220-500 percent higher hazard, while homogeneous goods face a 400-1040 percent higher hazard than differentiated goods.

We conclude that our results are not driven by small trade relationships. To the contrary, we find that the difference among the product types increases as we drop the small trade relationships.

## 6.2. Measurement errors

We are also concerned about potential measurement errors, by which we mean that we incorrectly infer that a trade relationship ended. What is the impact on our results of such an error?

To address this issue we focused on those trade relationships who have multiple spells. If the time between spells is short, it is possible that the gap is mis-measured. In this case, the two spells are really one longer spell. In what follows, we assume that a one-year gap between spells is an error, and hence merge the individuals spells. Gaps of two or more years are assumed to be accurate and no merging is done.

As an example, the data shows that the US imported “lace veils” from Germany in 1972, 1974, 1976, 1980, 1982, 1984-85, and 1987. Without any adjustment this trade pattern is interpreted as as seven distinct spells. Six spells are of length one year and one is of length two years. If we assume that all gaps of one year are errors, we would instead observe just two distinct spells, the first one from 1972 to 1976 (length five years) and the second one from 1980 to 1987 (length eight years).

Columns 5 and 6 of Tables 6 and 7 show the results using this gap-adjusted data. In the base specification (Table 6) we find that the impact of all covariates increases. Reference priced goods have an 18 percent higher hazard rate while homogeneous goods have a 27 percent higher hazard rate than differentiated goods. The results in Table 7 are quite similar.

These results indicate that potential measurement errors have little impact on our results and we thus conclude that our results are robust to these errors.

### 6.3. Multiple spells

Our last concern is that our treatment of multiple spells is not adequate. We have assumed that in the case of multiple spells for a single trade relationships all of the spells are independent and we treat it as a brand new relationship. In our estimated equations we only included a dummy variable indicating whether the spell is another spell for the same trading relationship. Across all specifications we found that multiple spells, or the return to the market (“reawakening” of the relationship) significantly increases the hazard rate. Our concern is that effects of the higher order spells is not entirely captured by the dummy variable.

As an alternative to the dummy variable specification, we estimated our models using the data for only the first spell. The last two columns of Tables 6 and 7 report the results. While the magnitudes of our estimates change somewhat, the signs do not change and all results are statistically significant. Limiting our analysis to the first spell only, produces similar results to including all spells in the analysis. In most cases our estimates indicate that in the case of first spells only the differences among product types are even bigger. These results lead us to conclude that our treatment of multiple spells as independent and including just a dummy variable indicating the re-appearance of a relationships is appropriate. We would now like to offer some interpretation for the estimated coefficients on this dummy.

In all of our regressions the spell dummy has a positive effect on the hazard rate. In other words, having already exported a product to the US, the return to the US market (after a gap) increases the hazard rate. The impact of this variable increases significantly as we limit the analysis to spells with larger values of trade, and it is always the biggest for differentiated goods and smaller for homogeneous goods. One explanation for these results is that once the country returns to the market, given the finite length of our sample, it has less time to create a long spell. Since multiple returns lead to progressively smaller potential lengths of spells, it is not surprising that we find that returning to the market raises the hazard rate. Given that differentiated goods are exported for longer periods of time, this can also explain the cross product type differences. It is entirely possible that in the return to the US market a country is able to maintain its presence longer having learnt something during its previous experience. However, our results indicate that the negative effect of the smaller time horizon is much

bigger.

## 7. Conclusion

In this paper we examined how the length of trade relationships differs for different product types, namely differentiated, reference priced, and homogeneous goods. We use two models, an oligopoly model and a standard model of trade with increasing returns and monopolistic competition, to motivate the analysis. The models suggest that trade in homogeneous goods will be more sensitive to changes in tariffs and costs than trade in differentiated goods. We use highly disaggregated data on US imports to investigate the relationship between the product type and the duration of trade.

Nonparametric estimates of survival and hazard functions reveal that trade in differentiated goods has longer survival time (and hence lower hazard rate) than trade in homogeneous products. We also show that the results are not driven by a lot of relationships of small trade value. To the contrary, we find the differences among the product types increase as we drop observations whose initial trade value was small.

We employ Cox's semiparametric proportional hazards model to investigate how size and cost factors influence the duration. We find that shorter distance, common language, common border, higher GDP, higher tariffs, and depreciation of the source country's currency all lead to longer durations. We again find that differentiated goods experience longer survival than homogeneous goods. Allowing for each variable to have a separate impact for each product type only strengthens our results. The addition of region dummies does not affect the results. Estimates for the set of OECD countries are similar to estimates for all countries with the exception that the exchange rate variable has a much bigger impact.

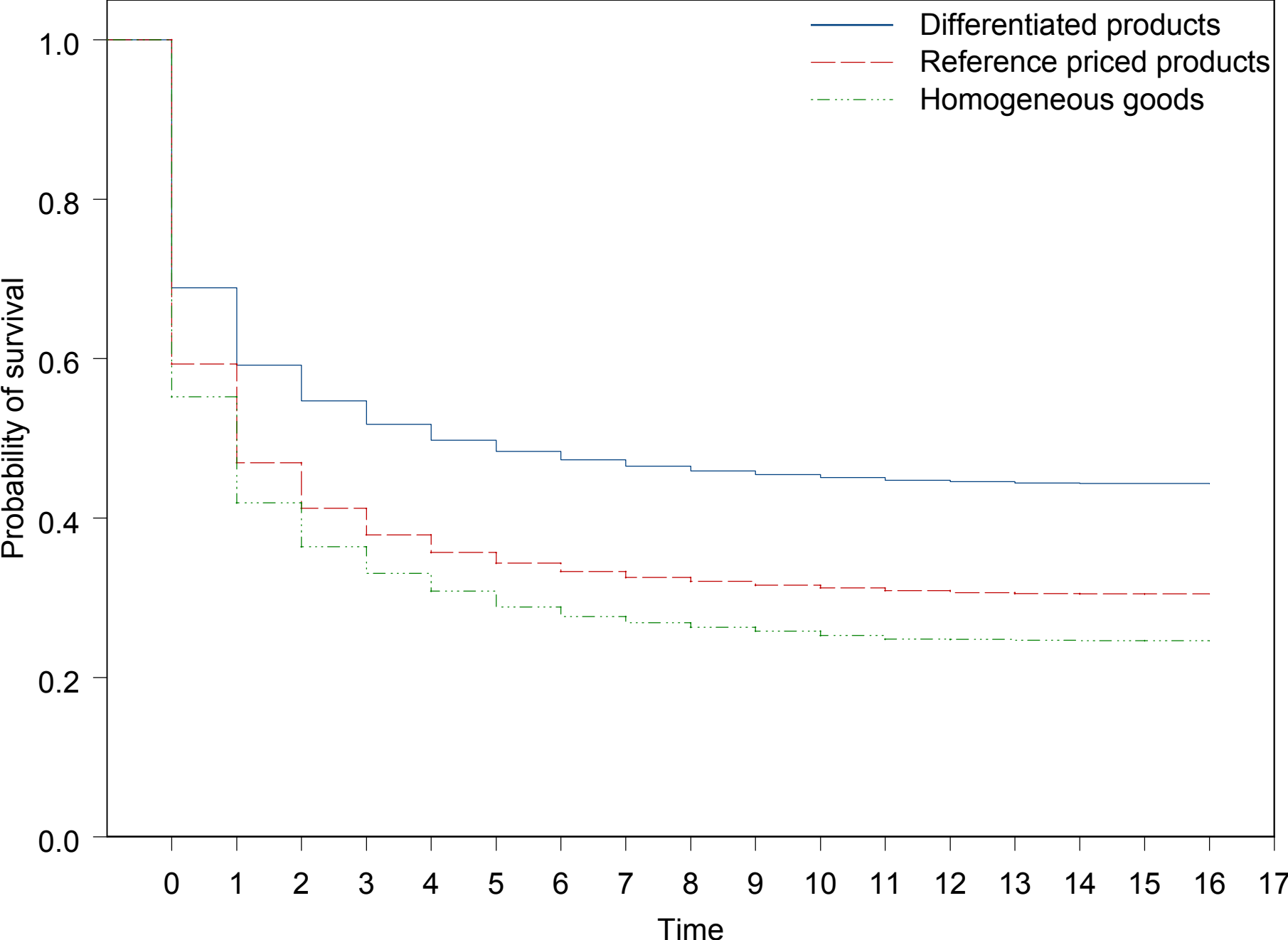
Our results show that duration of trade relationships critically depends on the product type involved. Differentiated goods have the longest duration, followed by reference priced, and then homogeneous goods.

## References

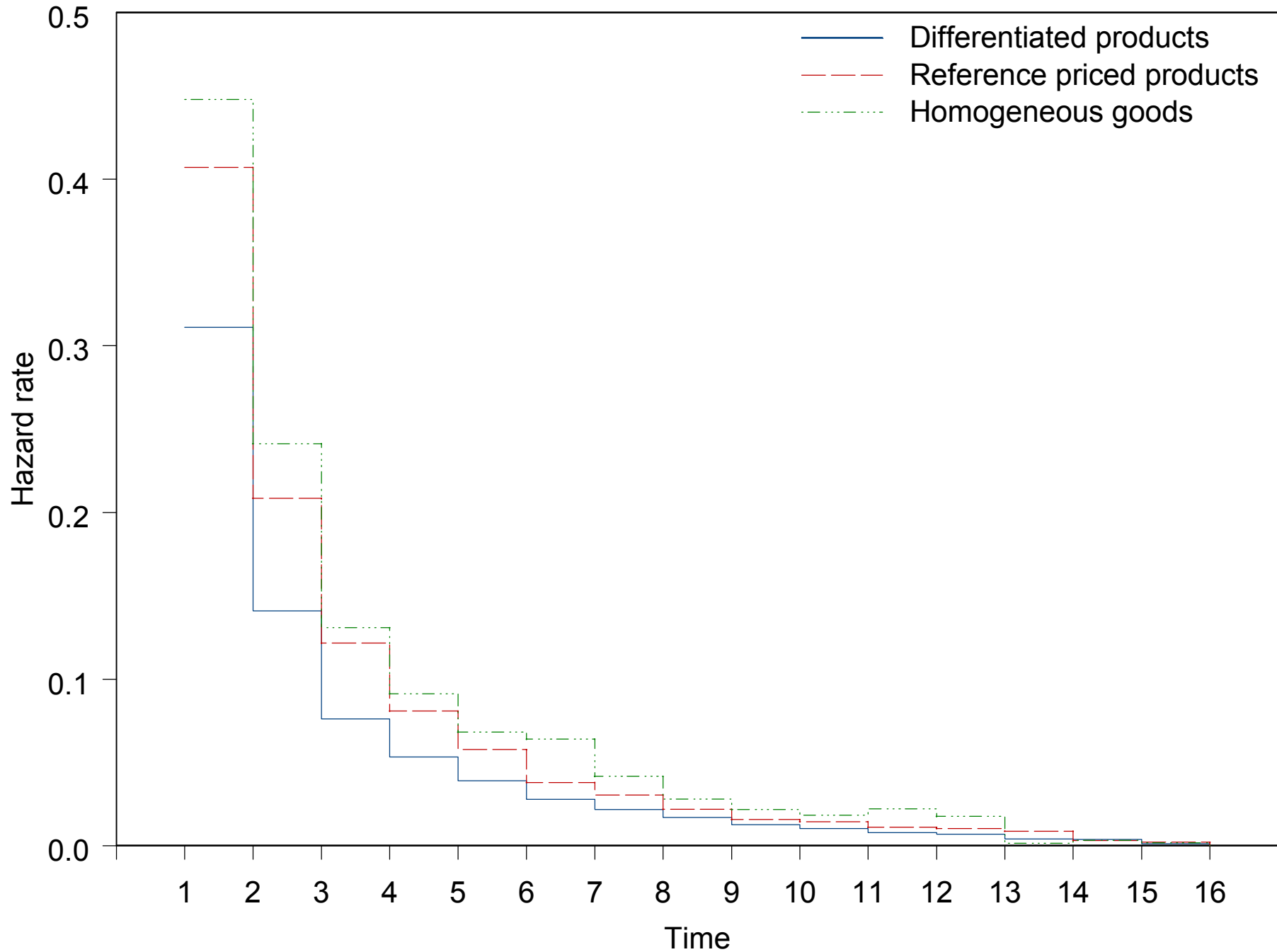
- [1] Besedeš T. and T.J. Prusa (2002), “Ins and Outs and the Duration of Trade,” Rutgers University Working Paper.
- [2] Cox, D.R. (1972), “Regression Models and Life-Tables,” *Journal of the Royal Statistical Society B* 34:187:220.
- [3] Dixit, A.K. and J. Stiglitz (1977), “Monopolistic competition and optimum product diversity,” *American Economic Review* 67:297–308.
- [4] Efron, B. (1977), “The Efficiency of Cox’s Likelihood Function for Censored Data,” *Journal of the American Statistical Association* 72(359):557-65.
- [5] Feenstra, R.C. (1996), “US Imports, 1972–1994: Data and Concordances,” NBER Working Paper 5515.
- [6] Greene, W.H. (1997), *Econometric Analysis*, 3ed. Upper Saddle River: Prentice Hall.
- [7] Grossman, G. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge: MIT Press.
- [8] Helpman, E. and P.R. Krugman (1985), *Market Structure and Foreign Trade*, Cambridge: MIT Press.
- [9] Hosmer, D.W. and S. Lemeshow (1998), *Applied Survival Analysis: Regression Modeling of Time to Event Data*, New York: John Wiley & Sons.
- [10] Keifer, N.M. (1988), “Economic Duration Data and Hazard Functions,” *Journal of Economic Literature* 26:646-79.
- [11] Klein, J.P. and M.L. Moeschberger (1997), *Survival Analysis: Techniques for Censored and Truncated Data*, New York: Springer Verlag New York.
- [12] Krugman, Paul R. (1979), “Increasing returns, monopolistic competition, and international trade,” *Journal of International Economics* 9(4):469–79.
- [13] Krugman, Paul R. (1980), “Scale economies, product differentiation, and the pattern of trade,” *American Economic Review* 70:950–59.
- [14] Krugman, Paul R. (1981), “Intraindustry specialization and the gains from trade,” *Journal of Political Economy* 89(5): 959–74.

- [15] Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.
- [16] Lee, E.T. (1992), *Statistical Methods for Survival Data Analysis*, New York: John Wiley & Sons.
- [17] Melitz, M. (2002), "The impact of trade on intra-industry reallocations and aggregate industry productivity," NBER working paper No. 8881.
- [18] Neary, J.P., (2000), "Monopolistic competition and international trade theory," paper prepared for the conference on "The Monopolistic Competition Revolution after Twenty-Five Year."
- [19] Porter, Michael E., *Competitive Advantage: Creating and Sustaining Superior Performance*, New York: The Free Press, 1985. Republished with a new introduction, 1998.
- [20] Porter, Michael E., *The Competitive Advantage of Nations*. New York: The Free Press, 1990. Republished with a new introduction, 1998.
- [21] Proudman, J. and S. Redding (1998), "Persistence and Mobility of International Trade," Chapter 2 in *Openness and Growth*, Bank of England, 1998.
- [22] Rauch, J.E. (1999), "Networks versus Markets in International Trade," *Journal of International Economics*, 48:7-35.
- [23] Salop, S. (1979), "Monopolistic competition with outside goods," *Bell Journal of Economics*, 10:141–56.
- [24] Schott, P.K. (2001), "Do Countries Specialize?," NBER Working Paper No. 8492.

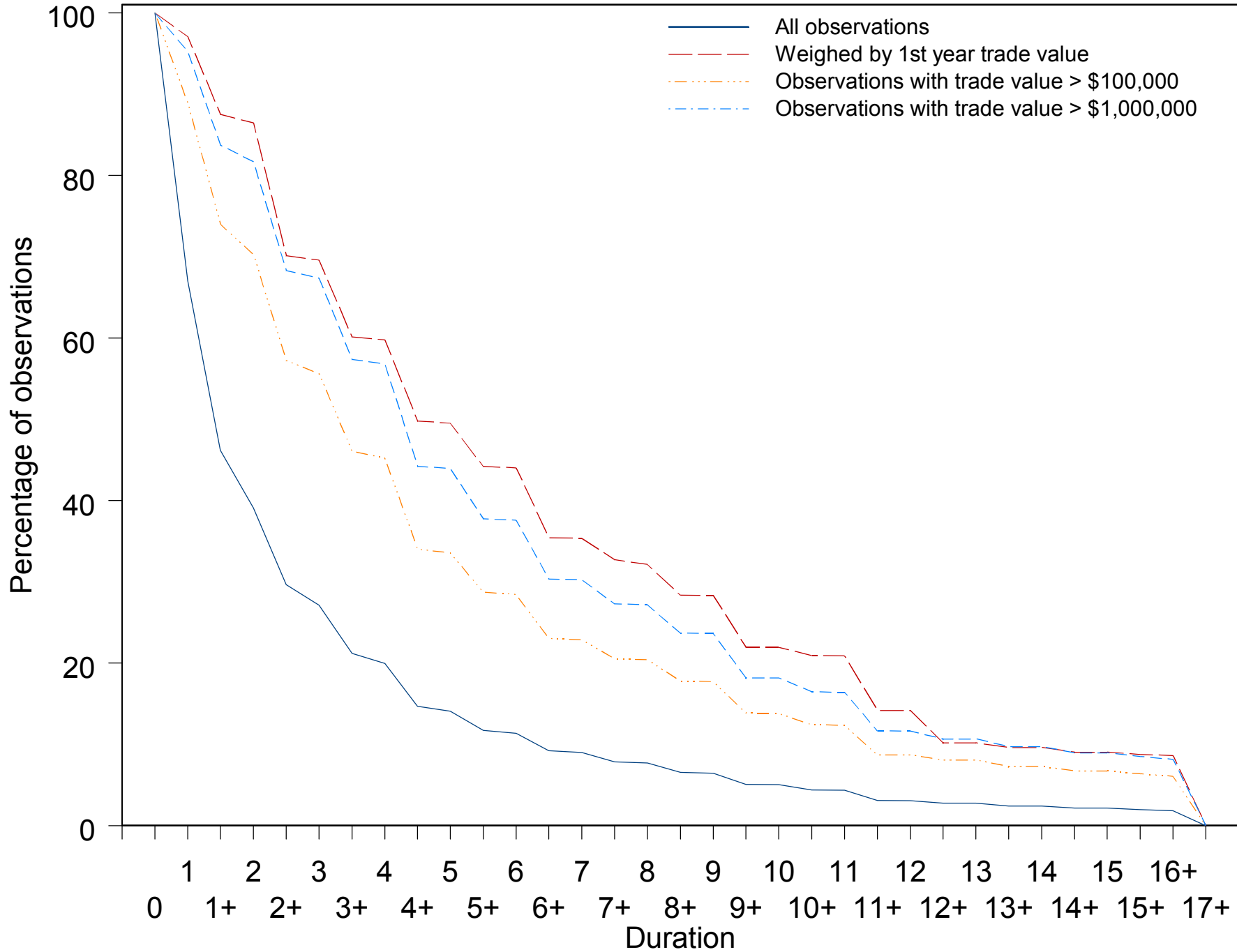
Figure 1 - Survival Function for Rauch's Product Classification



# Figure 2 - Hazard Function for Rauch's Product Classification

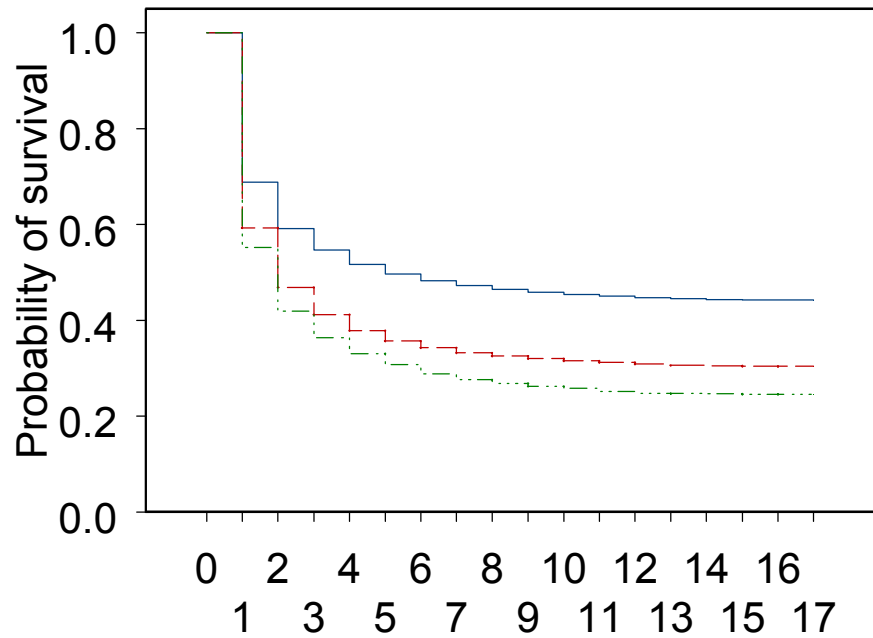


# Figure 3 - Distribution of Spell Lengths

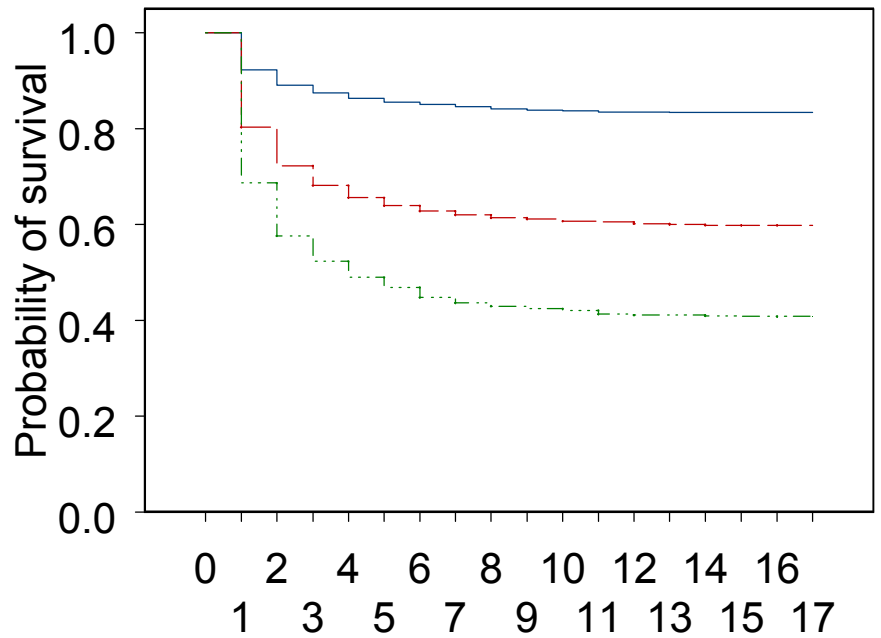


# Figure 4 - Survival Function for Rauch's Product Classification

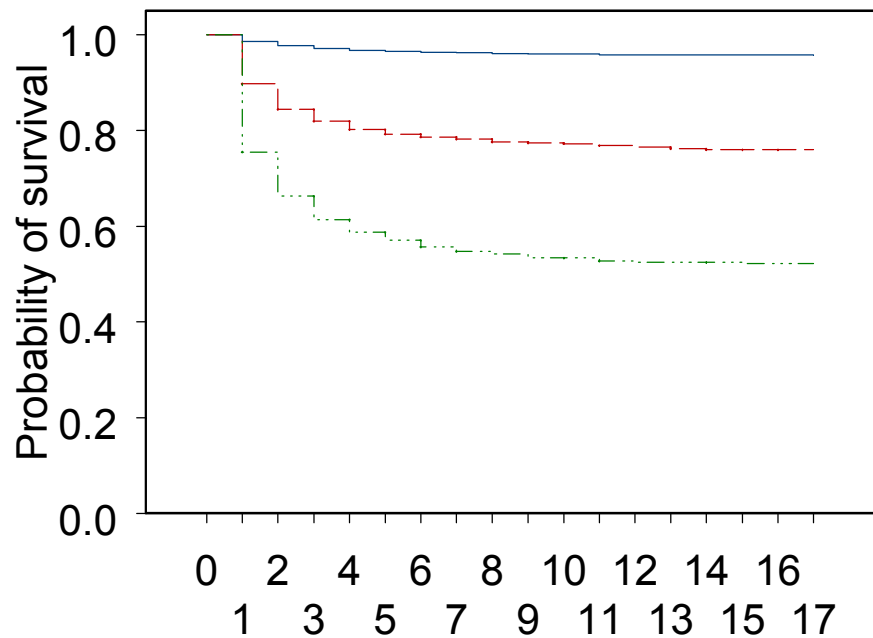
## All observations



## Observations with trade value > \$100,000



## Observations with trade value > \$1,000,000



- Differentiated goods
- - Reference priced products
- . Homogeneous goods

Table 1 - Semiparametric Cox Estimates (Hazard Ratios)

All countries		Basic	With relative ex rate		Separating (de-) appreciation		
Distance		1.01006	1.00147	1.00839	1.00247	1.01122	1.00043
	(1 unit increase = 1,000 kilometers)	0.000	0.416	0.000	0.171	0.000	0.812
Language Dummy		0.95949	0.95560	0.95730	0.95451	0.97445	0.96321
	(=1 if English)	0.000	0.000	0.000	0.000	0.000	0.000
Contiguous with USA		0.84575	0.74998	0.83504	0.75188	0.84968	0.74950
	(=1 if share border with USA)	0.000	0.000	0.000	0.000	0.000	0.000
GDP (\$100 billion)		0.97497	0.97919	0.97558	0.97998	0.97480	0.97867
	(1 unit increase = \$100 billion)	0.000	0.000	0.000	0.000	0.000	0.000
4 digit SIC tariff rate		0.98218	0.98226	0.98210	0.98223	0.98221	0.98229
		0.000	0.000	0.000	0.000	0.000	0.000
Percentage change in relative real exchange rate		0.90497	0.90136	0.91640	0.91252		
	(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
Relative real exchange rate				0.78573	0.79640		
				0.000	0.000		
Percentage change in relative real exchange rate>0						0.82557	0.83935
	(1 unit increase = 10 percentage points)					0.000	0.000
Percentage change in relative real exchange rate<0						0.97895	0.95950
	(1 unit increase = 10 percentage points)					0.000	0.000
Coefficient of variation of unit values		0.95275	0.94408	0.95170	0.94307	0.94971	0.94220
		0.000	0.000	0.000	0.000	0.000	0.000
Dummy if reference priced		1.16370	1.16143	1.16409	1.16143	1.16801	1.16490
		0.000	0.000	0.000	0.000	0.000	0.000
Dummy if homogeneous goods		1.24468	1.19115	1.24207	1.18851	1.24524	1.19506
		0.000	0.000	0.000	0.000	0.000	0.000
Dummy for spell >= 2		1.70203	1.72369	1.69561	1.71788	1.69463	1.71666
		0.000	0.000	0.000	0.000	0.000	0.000
Observations		866531	866531	866531	866531	866531	866531
Subjects		320635	320635	320635	320635	320635	320635
Estimated log likelihood		-1658722.88	-1657287.14	-1658623.74	-1657195.15	-1658281.21	-1657032.41
Region Dummies		no	yes	no	yes	no	yes

p-values are reported below the coefficients

Table 2 - Semiparametric Cox Estimates (Hazard Ratios)

OECD countries		Basic	With relative ex rate		Separating (de-) appreciation		
Distance		1.03667	1.01521	1.03668	1.01544	1.03729	1.01610
	(1 unit increase = 1,000 kilometers)	0.000	0.031	0.000	0.028	0.000	0.022
Language Dummy		0.95302	0.95774	0.95080	0.96088	0.95323	0.96040
	(=1 if English)	0.000	0.002	0.000	0.005	0.000	0.004
Contiguous with USA		0.98081	0.73499	0.97699	0.73776	0.98971	0.74654
	(=1 if share border with USA)	0.409	0.003	0.322	0.004	0.661	0.005
GDP (\$100 billion)		0.97726	0.97217	0.97733	0.97103	0.97713	0.97158
	(1 unit increase = \$100 billion)	0.000	0.000	0.000	0.000	0.000	0.000
4 digit SIC tariff rate		0.98991	0.98995	0.98999	0.99005	0.98998	0.99003
		0.000	0.000	0.000	0.000	0.000	0.000
Percentage change in relative real exchange rate		0.83899	0.83952	0.84461	0.84750		
	(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
Relative real exchange rate				0.88964	0.84719		
				0.000	0.000		
Percentage change in relative real exchange rate>0						0.82075	0.81903
	(1 unit increase = 10 percentage points)					0.000	0.000
Percentage change in relative real exchange rate<0						0.87064	0.87524
	(1 unit increase = 10 percentage points)					0.000	0.000
Coefficient of variation of unit values		0.90413	0.90402	0.90335	0.90290	0.90347	0.90328
		0.000	0.000	0.000	0.000	0.000	0.000
Dummy if reference priced		1.17087	1.16945	1.17137	1.16992	1.17204	1.17068
		0.000	0.000	0.000	0.000	0.000	0.000
Dummy if homogeneous goods		1.32301	1.32119	1.32389	1.32224	1.32474	1.32306
		0.000	0.000	0.000	0.000	0.000	0.000
Dummy for spell >= 2		2.12558	2.12894	2.11609	2.11634	2.11788	2.12064
		0.000	0.000	0.000	0.000	0.000	0.000
Observations		550213	550213	550213	550213	550213	550213
Subjects		182149	182149	182149	182149	182149	182149
Estimated log likelihood		-850180.11	-850159.02	-850172.23	-850143.97	-850169.56	-850145.77
Region Dummies		no	yes	no	yes	no	yes

p-values are reported below the coefficients

Table 3 - Semiparametric Cox Estimates by Product Types (Hazard Ratios)

All countries		Basic		With relative ex rate		Separating (de-) appreciation	
Differentiated	Distance	1.00615	0.99977	1.00415	1.00069	1.00731	0.99909
	(1 unit increase = 1,000 kilometers)	0.000	0.902	0.000	0.710	0.000	0.626
	Language Dummy	0.96564	0.96079	0.96324	0.95962	0.98329	0.97177
	(=1 if English)	0.000	0.000	0.000	0.000	0.034	0.001
	Contiguous with USA	0.83458	0.74186	0.82221	0.74297	0.83760	0.74280
	(=1 if share border with USA)	0.000	0.000	0.000	0.000	0.000	0.000
	GDP (\$100 billion)	0.97073	0.97517	0.97147	0.97607	0.97061	0.97475
	(1 unit increase = \$100 billion)	0.000	0.000	0.000	0.000	0.000	0.000
	4 digit SIC tariff rate	0.98217	0.98198	0.98207	0.98192	0.98215	0.98199
		0.000	0.000	0.000	0.000	0.000	0.000
	Percentage change in relative real exchange rate	0.89665	0.89321	0.91068	0.90684		
	(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
	Relative real exchange rate			0.74910	0.76201		
				0.000	0.000		
	Percentage change in relative real exchange rate>0					0.81336	0.82465
(1 unit increase = 10 percentage points)					0.000	0.000	
Percentage change in relative real exchange rate<0					0.98015	0.96261	
(1 unit increase = 10 percentage points)					0.000	0.000	
Coefficient of variation of unit values	0.92703	0.91597	0.92572	0.91472	0.92376	0.91382	
	0.000	0.000	0.000	0.000	0.000	0.000	
Dummy (=1) for spell >= 2	1.81437	1.83537	1.80604	1.82784	1.80849	1.82927	
	0.000	0.000	0.000	0.000	0.000	0.000	
Reference priced	Distance	1.02754	1.01219	1.02694	1.01417	1.02806	1.01111
	(1 unit increase = 1,000 kilometers)	0.000	0.000	0.000	0.000	0.000	0.000
	Language Dummy	0.95257	0.95338	0.95128	0.95272	0.95870	0.95138
	(=1 if English)	0.002	0.003	0.001	0.003	0.007	0.002
	Contiguous with USA	0.94009	0.80050	0.93584	0.80859	0.94188	0.80070
	(=1 if share border with USA)	0.028	0.000	0.019	0.000	0.033	0.000
	GDP (\$100 billion)	0.98343	0.98719	0.98366	0.98774	0.98330	0.98677
	(1 unit increase = \$100 billion)	0.000	0.000	0.000	0.000	0.000	0.000
	4 digit SIC tariff rate	0.97847	0.98085	0.97844	0.98089	0.97881	0.98082
		0.000	0.000	0.000	0.000	0.000	0.000
	Percentage change in relative real exchange rate	0.93299	0.92812	0.93667	0.93255		
	(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
	Relative real exchange rate			0.91847	0.90763		
				0.027	0.009		
	Percentage change in relative real exchange rate>0					0.88897	0.91313
(1 unit increase = 10 percentage points)					0.000	0.000	
Percentage change in relative real exchange rate<0					0.96726	0.94020	
(1 unit increase = 10 percentage points)					0.000	0.000	
Coefficient of variation of unit values	1.02884	1.02540	1.02848	1.02489	1.02728	1.02511	
	0.000	0.001	0.000	0.001	0.000	0.001	
Dummy (=1) for spell >= 2	1.42731	1.45305	1.42529	1.45063	1.42217	1.44994	
	0.000	0.000	0.000	0.000	0.000	0.000	
Homogeneous	Distance	1.01073	1.00393	1.00959	1.00542	1.01180	1.00291
	(1 unit increase = 1,000 kilometers)	0.008	0.363	0.019	0.212	0.004	0.501
	Language Dummy	0.88593	0.86094	0.88665	0.86188	0.89510	0.86428
	(=1 if English)	0.000	0.000	0.000	0.000	0.001	0.000
	Contiguous with USA	0.76539	0.71533	0.75776	0.71911	0.77404	0.71645
	(=1 if share border with USA)	0.000	0.000	0.000	0.000	0.000	0.000
	GDP (\$100 billion)	0.99475	0.99995	0.99520	1.00054	0.99475	0.99949
	(1 unit increase = \$100 billion)	0.006	0.979	0.012	0.772	0.006	0.783
	4 digit SIC tariff rate	0.98559	0.98637	0.98574	0.98644	0.98577	0.98648
		0.000	0.000	0.000	0.000	0.000	0.000
	Percentage change in relative real exchange rate	0.92955	0.92649	0.93601	0.93115		
	(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
	Relative real exchange rate			0.86562	0.90611		
				0.061	0.186		
	Percentage change in relative real exchange rate>0					0.86024	0.88460
(1 unit increase = 10 percentage points)					0.000	0.000	
Percentage change in relative real exchange rate<0					0.97856	0.95593	
(1 unit increase = 10 percentage points)					0.145	0.003	
Coefficient of variation of unit values	1.09634	1.10311	1.09586	1.10273	1.09162	1.09989	
	0.000	0.000	0.000	0.000	0.000	0.000	
Dummy (=1) for spell >= 2	1.21921	1.23194	1.21562	1.22915	1.20789	1.22423	
	0.000	0.000	0.000	0.000	0.000	0.000	
Dummy (=1) reference priced	0.95013	0.99072	0.76243	0.82082	1.00158	1.05507	
	0.038	0.704	0.000	0.000	0.951	0.038	
Dummy (=1) homogeneous goods	1.17405	1.10776	1.00298	0.92224	1.21062	1.15797	
	0.000	0.016	0.975	0.379	0.000	0.001	
Observations	866531	866531	866531	866531	866531	866531	
Number of subjects	320635	320635	320635	320635	320635	320635	
Estimated log likelihood	-1658230.90	-1656807.86	-1658119.59	-1656705.19	-1657767.49	-1656521.12	
Region Dummies	no	yes	no	yes	no	yes	
p values below the coefficients							

Table 4 - Semiparametric Cox Estimates by Product Types (Hazard Ratios)

OECD countries		Basic		With relative ex rate		Separating (de-) appreciation		
Differentiated	Distance	1.04405	1.02037	1.04402	1.02030	1.04505	1.02137	
		(1 unit increase = 1,000 kilometers)	0.000	0.004	0.000	0.004	0.000	0.003
	Language Dummy	0.95584	0.95354	0.95105	0.95298	0.95668	0.95623	
		(=1 if English)	0.001	0.003	0.000	0.002	0.001	0.005
	Contiguous with USA	1.06393	0.77037	1.05409	0.76556	1.08004	0.78390	
		(=1 if share border with USA)	0.027	0.013	0.060	0.011	0.006	0.020
	GDP (\$100 billion)	0.97258	0.96810	0.97271	0.96718	0.97234	0.96746	
		(1 unit increase = \$100 billion)	0.000	0.000	0.000	0.000	0.000	0.000
	4 digit SIC tariff rate	0.99093	0.99097	0.99114	0.99120	0.99110	0.99115	
			0.000	0.000	0.000	0.000	0.000	0.000
	Percentage change in relative real exchange rate	0.81389	0.81438	0.82643	0.82889			
		(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
	Relative real exchange rate			0.77158	0.74081			
				0.000	0.000			
	Percentage change in relative real exchange rate>0					0.78293	0.78161	
		(1 unit increase = 10 percentage points)				0.000	0.000	
	Percentage change in relative real exchange rate<0					0.87158	0.87569	
	(1 unit increase = 10 percentage points)				0.000	0.000		
Coefficient of variation of unit values	0.86594	0.86553	0.86424	0.86352	0.86489	0.86441		
		0.000	0.000	0.000	0.000	0.000		
Dummy (=1) for spell >= 2	2.30580	2.30725	2.28435	2.28276	2.29300	2.29390		
		0.000	0.000	0.000	0.000	0.000		
Reference priced	Distance	1.01977	0.99536	1.01940	0.99464	1.01844	0.99403	
		(1 unit increase = 1,000 kilometers)	0.000	0.555	0.000	0.496	0.000	0.448
	Language Dummy	0.97268	0.97723	0.97917	0.98907	0.97369	0.98040	
		(=1 if English)	0.229	0.345	0.360	0.652	0.246	0.417
	Contiguous with USA	0.83299	0.59498	0.83892	0.59907	0.81774	0.58514	
		(=1 if share border with USA)	0.000	0.000	0.000	0.000	0.000	
	GDP (\$100 billion)	0.98803	0.98337	0.98775	0.98193	0.98827	0.98318	
		(1 unit increase = \$100 billion)	0.000	0.000	0.000	0.000	0.000	
	4 digit SIC tariff rate	0.97662	0.97673	0.97651	0.97665	0.97651	0.97663	
			0.000	0.000	0.000	0.000	0.000	
	Percentage change in relative real exchange rate	0.90543	0.90622	0.89205	0.89534			
		(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
	Relative real exchange rate			1.33981	1.27394			
				0.000	0.000			
	Percentage change in relative real exchange rate>0					0.94350	0.94183	
		(1 unit increase = 10 percentage points)				0.000	0.000	
	Percentage change in relative real exchange rate<0					0.85296	0.85714	
	(1 unit increase = 10 percentage points)				0.000	0.000		
Coefficient of variation of unit values	1.00546	1.00621	1.00790	1.00843	1.00708	1.00778		
		0.575	0.524	0.419	0.388	0.468	0.425	
Dummy (=1) for spell >= 2	1.73349	1.73948	1.75369	1.75799	1.74652	1.75224		
		0.000	0.000	0.000	0.000	0.000		
Homogeneous	Distance	1.00710	0.98354	1.00702	0.98319	1.00706	0.98350	
		(1 unit increase = 1,000 kilometers)	0.389	0.120	0.394	0.112	0.394	0.120
	Language Dummy	0.83442	0.83829	0.83395	0.84273	0.83450	0.84030	
		(=1 if English)	0.001	0.001	0.001	0.002	0.001	0.001
	Contiguous with USA	0.73340	0.52545	0.73073	0.52292	0.73315	0.52622	
		(=1 if share border with USA)	0.005	0.000	0.005	0.000	0.005	0.000
	GDP (\$100 billion)	0.99595	0.99113	0.99602	0.99005	0.99596	0.99069	
		(1 unit increase = \$100 billion)	0.068	0.000	0.073	0.000	0.068	0.000
	4 digit SIC tariff rate	0.97760	0.97752	0.97756	0.97743	0.97760	0.97753	
			0.000	0.000	0.000	0.000	0.000	
	Percentage change in relative real exchange rate	0.92047	0.92176	0.92335	0.92729			
		(1 unit increase = 10 percentage points)	0.000	0.000	0.000	0.000		
	Relative real exchange rate			0.93958	0.88983			
				0.660	0.409			
	Percentage change in relative real exchange rate>0					0.92142	0.92035	
		(1 unit increase = 10 percentage points)				0.012	0.011	
	Percentage change in relative real exchange rate<0					0.91917	0.92399	
	(1 unit increase = 10 percentage points)				0.046	0.061		
Coefficient of variation of unit values	1.08336	1.08505	1.08266	1.08396	1.08342	1.08502		
		0.002	0.002	0.002	0.002	0.002		
Dummy (=1) for spell >= 2	1.31879	1.32276	1.31378	1.31462	1.31901	1.32264		
		0.000	0.000	0.000	0.000	0.000		
Dummy (=1) reference priced	1.41351	1.42013	0.78771	0.80125	1.54323	1.55097		
		0.000	0.000	0.004	0.007	0.000		
Dummy (=1) homogeneous goods	1.92846	1.93155	1.56806	1.59760	2.02045	2.02437		
		0.000	0.000	0.009	0.006	0.000		
Observations	550213	550213	550213	550213	550213	550213		
Number of subjects	182149	182149	182149	182149	182149	182149		
Estimated log likelihood	-849686.54	-849667.69	-849646.86	-849622.13	-849654.13	-849633.02		
Region Dummies	no	yes	no	yes	no	yes		
p values below the coefficients								

Table 5 - Tests for joint equality of coefficients, basic specification with region dummies

	All countries			OECD countries		
	All observations	Observations with trade value > \$100,000	Observations with trade value > \$1,000,000	All observations	Observations with trade value > \$100,000	Observations with trade value > \$1,000,000
Distance	25.12 0.000	13.20 0.001	10.93 0.004	47.67 0.000	13.99 0.001	3.01 0.222
Language Dummy	10.51 0.005	2.05 0.359	0.48 0.786	7.37 0.025	10.85 0.004	4.46 0.108
Contiguous with USA	4.81 0.090	0.55 0.759	0.54 0.765	32.16 0.000	14.00 0.001	13.47 0.001
GDP	266.40 0.000	96.66 0.000	49.91 0.000	261.10 0.000	32.92 0.000	28.07 0.000
4 digit SIC tariff rate	2.64 0.267	3.43 0.180	4.97 0.083	41.01 0.000	1.22 0.543	2.28 0.320
Percentage change in relative real exchange rate	52.99 0.000	2.21 0.332	1.74 0.418	166.81 0.000	4.86 0.088	1.49 0.474
Coefficient of variation of unit values	126.76 0.000	102.13 0.000	22.60 0.000	123.82 0.000	74.18 0.000	7.98 0.019
Dummy for spell >= 2	370.69 0.000	387.55 0.000	232.30 0.000	356.36 0.000	334.35 0.000	160.96 0.000

p-values are reported below test statistics

Table 6 - Semiparametric Cox Estimates (Hazard Ratios)

Countries	Observations with trade value > \$100,000		Observations with trade value > \$1,000,000		Spells merged if gap between them is one year		First spell only	
	All	OECD	All	OECD	All	OECD	All	OECD
Distance	1.01995	1.03579	1.00806	1.03966	1.01197	1.04339	1.01248	1.04913
(1 unit increase = 1,000 kilometers)	0.000	0.000	0.264	0.017	0.000	0.000	0.000	0.000
Language Dummy	0.75040	0.80920	0.71637	0.74358	0.95129	0.95055	0.93193	0.94845
(=1 if English)	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000
Contiguous with USA	0.67507	0.69025	0.40804	0.32634	0.78544	0.93067	0.77535	0.90087
(=1 if share border with USA)	0.000	0.000	0.000	0.000	0.000	0.013	0.000	0.001
GDP (\$100 billion)	0.95423	0.96501	0.95595	0.96201	0.96771	0.97190	0.96302	0.96831
(1 unit increase = \$100 billion)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4 digit SIC tariff rate	0.94454	0.94051	0.88220	0.87838	0.98060	0.98913	0.97865	0.98779
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Percentage change in relative real exchange rate	0.91779	0.83254	0.95974	0.84678	0.87935	0.78108	0.93404	0.87919
(1 unit increase = 10 percentage points)	0.000	0.000	0.025	0.000	0.000	0.000	0.000	0.000
Coefficient of variation of unit values	0.89140	0.85128	0.93978	0.89770	0.95177	0.89403	0.96176	0.89589
	0.000	0.000	0.029	0.009	0.000	0.000	0.000	0.000
Dummy if reference priced	1.54210	1.47143	2.58799	2.21425	1.18748	1.20323	1.26277	1.29001
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Dummy if homogeneous goods	1.85310	1.87826	3.34329	2.79400	1.27195	1.39213	1.42182	1.63395
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Dummy for spell >= 2	3.02537	4.75365	3.39557	5.52020	1.69352	2.14100		
	0.000	0.000	0.000	0.000	0.000	0.000		
Observations	286482	193303	109596	77715	910563	577646	670947	425171
Subjects	62258	35876	19664	12290	275923	154040	233396	128318
Estimated log likelihood	-115843.97	-53095.87	-16773.48	-7694.23	-1204366.82	-583958.41	-1003780.07	-461757.68
Region Dummies	no	no	no	no	no	no	no	no

p-values are reported below the coefficients

