WPS4368

Policy Research Working Paper 4368

Trade Costs, Barriers to Entry, and Export Diversification in Developing Countries

Allen Dennis
Ben Shepherd

The World Bank
Financial and Private Sector Development
Doing Business
&
Development Research Group
Trade Team
September 2007

Abstract

This paper finds that a 1 percent reduction in the cost of exporting or the cost of international transport is associated with an export diversification gain of 0.3 percent or 0.4 percent respectively. Lower domestic market entry costs can also promote diversification, but the elasticity is weaker (-0.1). To obtain these results, the authors construct new measures of export diversification for 118 developing countries using highly detailed 8-digit mirror data from the European Union. The analysis also incorporates new export cost data from the World Bank's Doing Business database, covering document preparation, inland transport, administrative fees, and port/customs charges. Findings are highly robust, including to the use of geography and colonial history as instruments for trade and entry costs. Both the signs and relative magnitudes of these effects are consistent with predictions from a heterogeneous firms model of trade with asymmetric costs.

This paper—a joint product of Doing Business and the Trade Team, Development Research Group—is part of a larger project on trade costs, trade facilitation, and economic development, supported through a Trust Fund of the U.K. Department for International Development (http://econ.worldbank.org/projects/trade_costs). Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at adennis@worldbank.org and bshepherd@worldbank.org.

Trade Costs, Barriers to Entry, and Export Diversification in Developing Countries

Allen Dennis and Ben Shepherd*
The World Bank†

JEL codes: F12; F13; O24.
Keywords: International trade; Economic development; Product variety; Trade policy.

1 Introduction

In this paper, we show that lower export costs, international transport costs, and domestic market entry costs are all robustly associated with export diversification in a sample of 118 developing countries. With elasticities of -0.3, -0.4, and -0.1, these effects are of real economic significance to

*The authors are respectively Economist (Doing Business) and Consultant (Trade Team, Development Research Group). This work is part of a project on trade facilitation and development supported through a Trust Fund of the U.K. Department for International Development (http://econ.worldbank.org/projects/trade_costs). We are grateful to the following for helpful discussions and suggestions: Simeon Djankov, Matthias Helble, Bernard Hoekman, Will Martin, Beata Smarzynska Javorcik, and Christian Volpe Martincus. Comments to: bshepherd@worldbank.org or adennis@worldbank.org.

the many developing countries that view diversification of production and exports as a key development objective. Indeed, Imbs and Wacziarg (2003) show that a higher level of per capita income tends to be associated with a more diverse production structure, at least until relatively late in the development process when specialization effects begin to dominate. Funke and Ruhwedel (2001) find that export diversification is positively related to per capita GDP and TFP growth in OECD countries. Similarly, Hausmann et al. (2006) report that exports of particular sets of products are associated with higher national income levels.

If diversification can have a positive impact on a country's growth and development prospects, what are the policy options available to support that process? Traditional inward-looking development policies (e.g., Prebisch, 1959) used import protection and industrial policy to induce diversification into non-traditional goods. In line with the increasing relevance of outward-looking policies, it has recently been suggested that more effective trade preferences covering non-traditional exports might be part of the answer (Collier and Venables, 2007). However, the historical experience with such schemes has been at best mixed (Ozden and Reinhardt, 2005). An alternative approach recently suggested by Rodrik (2004) includes a type of "new age" industrial policy, emphasizing the provision of public goods and measures to facilitate innovation and entrepreneurship.

The novelty of the present paper lies in providing a comprehensive framework that makes it possible to analyze and compare policies such as these from an export diversification standpoint. By characterizing diversification in terms of export growth at the extensive (or "new products") margin, we can draw on the recent heterogeneous firms literature following Melitz (2003), as well as empirical work that has started charting the pattern of extensive margin growth. One example of this recent empirical literature is Broda and Weinstein (2006), who use US import data at the 7- and 10-digit levels to show that import variety increased approximately threefold over a 30 year period, leading to substantial national welfare gains. Hummels and Klenow (2005) examine extensive margin growth across 126 exporting countries and 59 importing countries, relying primarily on trade data at the 6-digit level. They find that larger, richer countries systematically trade in a relatively wider range of goods. Kehoe and Ruhl (2003) present similar evidence in respect of variable trade costs, with episodes of tariff liberalization across 18 countries found to be associated with extensive margin growth. Their findings are corroborated for China and Mexico by Feenstra and Kee (2007), while Debaere and Mostashari (2005) find further evidence that changes in tariff rates and preferences can impact the extensive margin. In the context of examining the role of export variety in explaining cross-country productivity differences, Feenstra and Kee (2006) show that tariffs, international distance, and resource endowments can serve as instruments for export variety—and thus by implication, these factors impact exports at the extensive margin.

Our key contribution in this paper is to expand the range of factors believed to influence export growth at the extensive margin—or export diversification—to include a range of policy variables related to the costs of trading and domestic market entry. We also compare the sensitivity of diversification to changes in these parameters, giving an indication of where the highest policy payoffs might lie. We motivate the selection of this variable set using a version of the Melitz (2003) heterogeneous firms model of trade, which allows for potential asymmetries in country size, market entry costs, and trade costs. It provides insights into the effects of both reciprocal and unilateral changes in each of these variables. Numerical simulations suggest that for given sector characteristics, reductions in export-specific trade costs (fixed and variable), symmetric trade costs (such as international transport), and domestic market entry costs, as well as increases in import-specific trade costs, are positively associated with export diversification. They also provide evidence of the relative magnitudes of these effects: we expect transport costs to have the largest impact, followed by export costs and destination country tariffs, then source country tariffs, and finally domestic market entry costs.

To test these predictions, we construct new measures of export diversification for 118 developing countries using 8-digit mirror (import) data from the European Union, which distinguish among 10,753 distinct products. The World Bank's *Doing Business* database provides previously unexploited information on the costs of exporting, as well as data on the costs of domestic market entry. For applied tariffs that take full account of preferential rates, we use the ITC-CEPII Market Access Map (MAcMap) database.

Our core empirical strategy relies on a Poisson model in which the dependent variable is a count of the number of 8-digit product lines exported in each 2-digit sector. The advantage of this approach is that our estimates remain unbiased and consistent even in the presence of large numbers of fixed effects, whereas many other non-linear panel data models—including Tobit and Heckit—do not. Moreover, our formulation links directly back to the theoretical model, thereby simplifying interpretation.

Our empirical results are largely consistent with the model's predictions, both in terms of signs and relative magnitudes of the estimated elasticities. Only for source and destination country tariffs are the observed effects less robust than expected. We subject our baseline empirical model to a battery of robustness checks including alternative definitions of export diversification (including the measure due to Feenstra, 1994), an expanded set of explanatory variables, more restricted definitions of what constitutes a "developing country", and use of data on the number of administrative procedures involved in exporting and starting a business in place of cost. Our results prove remarkably consistent across the wide range of specifications used. Finally, we check model identification using an instrumental variables approach based on the association between our cost measures and European colonization and distance from the equator. These estimates do not substantially alter our results.

The paper proceeds as follows. In the next section, we take a first look at the data on export diversification, trade costs, and market entry costs. We then use non-parametric methods to derive some stylized facts that motivate our subsequent work. Section 3 develops a theoretical framework, and uses numerical simulations to make predictions as to the signs and relative magnitudes of the elasticities of diversification with respect to trade costs and market entry costs. Our empirical strategy is discussed in section 4, and estimation results presented. Section 5 presents our conclusions and puts forward some suggestions for future research in this area.

2 Data and Stylized Facts

Given the potential importance of export diversification for developing countries, what do the data say about its extent and cross-country variation? What might be some of the explanatory factors behind these observations? After a brief discussion of our approach to measurement issues, this section presents some initial impressions from the cross-country data. We then use non-parametric techniques to examine the links between export diversification and its potential determinants, such as country size and level of development, trade costs, international distance, and the costs of domestic market entry.

2.1 Measuring Export Diversification

At its simplest, export diversification means broadening the range of products that a country exports. We therefore equate it with export growth at the extensive margin, and draw on the recent literature in that area for guidance as to the ways in which diversification can be measured.[1] As a starting point, we use a direct measure, namely a count of the number of products that a country exports. In practice, this approach is not as simple as it seems because individual "products" identified in the trade data usually map in reality to a number of distinct varieties. The most detailed trade data available on a worldwide basis are at the 6-digit level of the Harmonized System (HS), and distinguish amongst 5,000 or so different products. Casual empiricism suggests, however, that counts based on 6-digit data are likely to understate the true level of export diversity due to aggregation effects.

In this paper, we improve on the level of detail provided by the 6-digit HS classification by using an 8-digit classification that provides roughly twice as much product-level detail: 10,753 distinct product lines.[2] We extract these data from a freely available Eurostat database covering

[1] We limit consideration to growth in the number of products exported, and do not examine the number of markets to which they are exported. This is an important question in its own right, and we expect that the methods used in this paper could be adapted to build on the insights developed in this area by Evenett and Venables (2002).

[2] The data are classified using the EU's Combined Nomenclature (CN), which is based on the Harmonized System (HS) but contains additional subdivisions at the 8-digit level.

exports from and imports into the European Union.[3] Our strategy is to use EU data on imports from developing countries to construct new "mirror" measures of export diversification in those same countries. Although it is true that we thereby measure export diversification vis-à-vis the EU and not the world as a whole, we believe that our measures remain highly relevant since the EU is one of the most important outlets for developing country exports. Moreover, this approach offers two concrete advantages over the use of cross-country export data at the 6-digit level. First, the aggregation problem is reduced (although not eliminated) due to the much greater level of detail in which products are defined and flows recorded. Second, import data from the EU are likely to be more reliable than the corresponding export data from developing countries due in part to stronger governance and customs agency capacity.

Taking 2005 as our base year, we start with a dataset of 470,035 observations across 246 countries and customs areas (including EU members), and 10,753 distinct products.[4] In this paper, we focus only on the developing country component of that dataset, namely countries that are neither members of the EU-25 nor the OECD. (We return to this definition in the context of robustness checks below.) Our first measure of export diversification, *lines*, is a count of the number of 8-digit product lines in which a given country exported to the EU-15 in 2005. It has one observation per country. To provide greater detail, we also construct *lines_cn2* following the same pattern as for *lines*, but with counts by 2-digit sector rather than aggregated to the country level. *Lines_cn2* therefore has 97 observations per country (the number of 2-digit Chapters in the CN classification). Given that the CN 8-digit classification scheme is inconsistent in the level of detail (i.e., individual "products") it accords each sector, we will need to take care to correct for this when using *lines_cn2* as an indicator of export diversification.

In Table 1, we provide a list of the countries included in our sample divided up according to the quintiles of *lines*. On average, they exported 1,138 8-digit product lines to the EU in 2005.

[3] See http://fd.comext.eurostat.cec.eu.int/xtweb/. To our knowledge, these data have not previously been used in product variety work. However, Fontagné et al. (1998), Henry de Frahan and Vancauteren (2006), and Manchin (2006) have exploited them in other contexts.

[4] Although these data are available for a number of years, limited availability of trade and entry cost data means that we must work in a purely cross-sectional context.

However, the range is extremely wide: from 9 lines (Palau) to 8053 (China), out of a possible maximum of 10,753. In broad terms, the country rankings accord with the sensible prior that larger, more developed countries tend to have more diversified export bundles (see Hummels and Klenow, 2005). Thus, we find China, India, and Brazil at the top of the table, while Palau, Micronesia, and the Comoros are at the opposite end.

It is important to see how our measure compares with alternative approaches. We start with the measure of relative variety proposed by Feenstra (1994) and used with modifications by, for instance, Hummels and Klenow (2005), and Broda and Weinstein (2006). We adopt the formulation used by Feenstra and Kee (2006), indexing varieties by i and using J^H and J^W to refer to the sets of varieties exported by country H and the world respectively:

$$\Lambda = \frac{\sum\limits_{i \in J^H} p_i^w q_i^w}{\sum\limits_{i \in J^W} p_i^w q_i^w} \tag{1}$$

The numerator in this measure is therefore the total value of world exports in product lines exported by country H, and the denominator is the total value of world exports across all products. When we calculate Λ, it turns out to be very strongly correlated ($\rho = 0.95$) with *lines*.

A second alternative is a Herfindahl-Hirschman index of export concentration (i.e., inverse diversification), which is simply the sum of the squared export product market shares:

$$hh\_index = \sum_{i=1}^{J} \left(\frac{p_i q_i}{\sum\limits_{j=1}^{J} p_j q_j} \right)^2 \tag{2}$$

Although lacking a strong theoretical basis of the type provided by Feenstra (1994) for Λ, this measure is frequently used in policy circles as a summary measure of export diversification (e.g., UNCTAD, 2006). It is also correlated with our count measure ($\rho = -0.52$), although since it mixes effects stemming from price, quantity, and variety changes, it is unsurprising that this correlation should be lower than for Λ.

Compared with the ready alternatives, our count data approach to measuring export diversification would appear to provide acceptable results. Given its ease of interpretation and direct link to the theoretical model we develop in the next section, we focus on that measure in the remainder of this paper. We are conscious, however, that its robustness is an important issue, and we therefore return to it below in the context of our empirical model.

2.2 Measuring the Costs of Trade and Domestic Market Entry

Full details of the remaining elements in our dataset are presented in Table 2. Our bilateral applied tariff data from the ITC-CEPII Market Access Map (MAcMap) database take full account of preference schemes, and are available for 2004 only. We aggregate them from the HS 6-digit level to the 2-digit level using the reference group methodology due to Laborde et al. (2007). This aggregation scheme minimizes endogeneity bias by calculating weights based on import shares in groups of similar countries, rather than in each country individually. We also draw historical and geographical data from CEPII (Mayer and Zignago, 2006), while our macroeconomic data are extracted from the World Development Indicators.

In addition to applied tariffs, we use new data from the World Bank's *Doing Business* database to measure trade costs. For the first time in 2006, the "Trading Across Borders" component of *Doing Business* captures the total official cost for exporting a standardized cargo of goods ("Export Cost"), excluding ocean transit and trade policy measures such as tariffs. The four main components of the costs that are captured are: costs related to the preparation of documents required for trading, such as a letter of credit, bill of lading, etc.; costs related to the transportation of goods to the relevant sea port; administrative costs related to customs clearance, technical controls, and inspections; and ports and terminal handling charges. The indicator thus provides a useful cross-section of information in relation to a country's approach to trade facilitation, in the broad sense in which that term is used by Wilson et al. (2005). The data are collected from local freight forwarders, shipping lines, customs brokers, and port officials, based on a standard set of assumptions, including: the traded cargo travels in a 20ft full container load; the cargo is valued at $20,000; and the goods do not require any special phytosanitary, environmental, or safety standards beyond what is required internationally. They disclose a considerable range of country experiences: these export operations cost as little as $300-$400 in Tonga, China, Israel, Singapore, and UAE, whereas they run at nearly ten times that level in Gabon and Tajikistan. On average, the cost is around $1278 per container (excluding OECD and EU countries). To our knowledge, these data have not previously been used in empirical work, although Djankov et al. (2006) use closely related *Doing Business* series on the amount of time and number of administrative procedures required to export and import. As expected, they find that these factors impact negatively on bilateral trade.

To measure domestic market entry costs, we use the "Starting a Business" component of *Doing Business* (see Djankov et al., 2002). This source includes indicators on the costs, time, and number of procedures required for an entrepreneur to start-up and formally operate a local limited liability company with general industrial or commercial activities. This includes legally required pre-registration, registration, and post-registration activities. Only official costs are considered, based on information gathered from the company law, commercial code, and specific regulation and fee schedules. Together, we refer to these as the costs of domestic market entry ("Entry Cost"). As far as we are aware, this is the most comprehensive source of cross-country information on business start-up costs, and has previously been used in the trade context by Helpman et al. (2007): they find that higher entry costs are negatively associated with the probability that two countries engage in trade.

2.3 Non-Parametric Evidence

Based on the theoretical approach that we develop in the next section, our main contentions in this paper concern the policy determinants of export diversification in developing countries. Using the sources just described, we will argue that lower costs of exporting, international transport, and domestic market entry, as well as higher import costs, are all potentially associated with greater diversification. The basic intuition for this is simple: existing domestic producers move into export markets whenever the returns they expect are sufficiently high and the costs they face are sufficiently low; thus, policy changes that impact the profitability of exporting can affect the rate of export market entry, which in a differentiated goods context is expressed through changes in the number of varieties exported.

Before entering into the detail of our approach, it is useful to have a first look at the data using non-parametric methods. Like Imbs and Wacziarg (2003), we use the locally weighted scatterplot smoother (Lowess) for this purpose.[5] We use *lines* as the dependent variable in all cases, resulting in a single observation per country. It is important to be clear that this analysis is intended to be exploratory only, since it is conducted one variable at a time, and does not control for the impact of other factors.

Results are presented in Figure 1. The first two plots verify the analysis of Hummels and Klenow (2005) to the effect that larger (GDP), richer (GDP per capita) countries tend to have a more diverse export bundle.[6] We do not observe the inverse-U relationship between *lines* and GDP per capita that would be expected on the basis of the results in Imbs and Wacziarg (2003), undoubtedly because our sample includes only developing countries, which may not yet have reached the relatively high level of income (approximately that of Ireland) at which specialization effects come to dominate. The remainder of Figure 1 discloses generally negative relationships between export variety and entry costs (upper right corner), great circle distance from the EU (as a proxy for transport costs), and export costs (lower right). In the case of distance, however, the relationship is a complex one involving a mid-sample "hump".

The data therefore provide support for an important stylized fact: higher trade costs and barriers to domestic market entry are associated with a less diversified export bundle. In the next section, we develop a comprehensive theoretical framework that allows us to examine the mechanisms giving rise to these observations, before conducting more rigorous, parametric tests in section 4.

[5] The Lowess smoother runs separate regressions of *lines* on trade or entry costs at each data point, using 80% of the total sample. For each regression, observations are downweighted according to their distance from the central data point around which the regression takes place. All calculations are performed in Stata 9.2SE.

[6] The outlier in the upper right corner of the first plot is China. Results do not change if it is dropped.

3 Theoretical Model

We now show that a heterogeneous firms trade model similar to Melitz (2003) can explain the stylized facts presented in the previous section.[7] The major substantive difference between our approach and Melitz's is that we allow for potentially asymmetric countries, trade costs, and market entry costs. We also derive some explicit results in relation to the extensive margin of trade, which we interpret in terms of export diversification. While the basic mechanics of the model remain unchanged, the introduction of these asymmetries prevents us from obtaining an analytical solution (Baldwin and Harrigan, 2007). We therefore proceed by numerical simulation using essentially the same parameter set as in Bernard et al. (2007). We consider unilateral changes in the fixed and variable costs of exporting and importing, as well as in the costs of domestic market entry. We then consider a reciprocal change in the variable costs of trade, which we assimilate with a variation in international transport costs. Simulation results not only accord with the preliminary empirical evidence presented in section 2, but also suggest an ordering of export diversification elasticities with respect to the different types of costs we have identified. We test both sets of predictions in section 4.

3.1 Consumption Block

The world consists of two possibly asymmetric regions, Home (H) and the Rest of the World (R). Labor is the only factor of production, and the regions are endowed with L^H and L^R units respectively. Each region has $S + 1$ productive sectors, of which one produces a freely traded homogeneous good under constant returns to scale with one unit of labor required for one unit of output. Wages are therefore equal to unity in equilibrium. The other S sectors each produce a continuum of differentiated goods under increasing returns to scale and costly trade. Absolute specialization is excluded from consideration.

Identical consumers in both markets maximize the two-tier utility function (3). The first tier is

[7] Our presentation of the model draws heavily on Helpman et al. (2004), and Baldwin and Harrigan (2007).

Cobb-Douglas, with expenditure share β_s on the differentiated product sector s and $\left(1-\sum_{s=1}^{S}\beta_s\right)$ on the homogeneous sector. The second tier for each differentiated goods sector is Dixit-Stiglitz with elasticity of substitution σ_s across the set V_s of available varieties.

$$U = q^{\left(1-\sum_{s=1}^{S}\beta_s\right)} \prod_{s=1}^{S} \left[\int_{v\in V_s} x_s(v)^{1-\frac{1}{\sigma_s}} dv \right]^{\beta_s\left(\frac{\sigma_s}{\sigma_s-1}\right)} \tag{3}$$

We use $d_s^H = \frac{\beta_s E^H}{\int_{v\in V_s^H} p(v)^{1-\sigma_s} dv}$ to denote the home country's demand shift parameter in sector s, where V_s^H is the set of sector s varieties available in the home country, and $E^H = L^H$ is total expenditure by home consumers. It is therefore possible to write the demand function x facing a given differentiated goods producer as:

$$x\left[p\left(v\right)\right] = d_s^H \left[p\left(v\right)\right]^{-\sigma_s} \tag{4}$$

3.2 Production Block

As usual in the Dixit-Stiglitz context, firms engage in "mill pricing" with a constant markup over marginal cost:

$$p\left(v\right) = \frac{\sigma_s}{\sigma_s - 1} c \tag{5}$$

For goods produced and consumed in the home country, the consumer price is simply given by the above expression. However, sector s varieties that are produced in R and shipped to H are subject to "iceberg" trade costs $\tau_s^{HR} > 1$, including factors such as import tariffs in H, export-specific costs in R relating to customs clearance or similar procedures, and international transport costs between the two countries. This means that an R firm must ship τ_s^{HR} units of a traded variety in order for one unit to arrive in H. Equivalently, the effective marginal cost for the exporting firm is $\tau_s^{HR}c$ and the consumer price in the importing country H is $\tau_s^{HR}p\left(v\right) = \frac{\sigma_s}{\sigma_s-1}\tau_s^{HR}c$. Iceberg trade costs are allowed to vary by direction of trade and sector, so we do not assume $\tau_s^{HR} = \tau_s^{RH}$ or $\tau_s^{HR} = \tau_{r\neq s}^{RH}$.

In addition to the variable costs of production c and of trade τ_s^{HR}, firms also face fixed market entry, or "beachhead" costs. We assume that these costs are country-specific, but that they do not vary by sector. Thus, entering the domestic market in H requires a payment of f_d^H, which means that a typical firm producing in that country will have domestic profits $\pi_{d,s}^H$ given by the following expression:

$$\pi_{d,s}^H = \underbrace{d_s^H \left(\frac{\sigma_s}{\sigma_s - 1} c\right)^{1-\sigma_s}}_{\text{revenue}} - \underbrace{d_s^H \left(\frac{\sigma_s}{\sigma_s - 1} c\right)^{-\sigma_s} c}_{\text{variable costs}} - \underbrace{f_d^H}_{\text{fixed costs}}$$

$$\equiv \frac{d_s^H}{\sigma_s} \left(\frac{\sigma_s}{\sigma_s - 1} c\right)^{1-\sigma_s} - f_d^H \quad (6)$$

Setting this expression equal to zero and solving for c establishes a maximum marginal cost $c_{d,s}^H$ above which it is not possible for producers in H to profitably supply the domestic market in sector s.

$$c_{d,s}^H = \left(\frac{\sigma_s f_d^H}{d_s^H}\right)^{\frac{1}{1-\sigma_s}} \left(\frac{\sigma_s - 1}{\sigma_s}\right) \quad (7)$$

Repeating the exercise for R provides a second cutoff of the same form:

$$c_{d,s}^R = \left(\frac{\sigma_s f_d^R}{d_s^R}\right)^{\frac{1}{1-\sigma_s}} \left(\frac{\sigma_s - 1}{\sigma_s}\right) \quad (8)$$

The same analysis applies to export market entry, for which an additional fixed cost of $f_x > f_d$ is payable. These costs are again allowed to vary by country. The additional firm profits due to exporting from H to R are equal to:

$$\pi_{x,s}^{HR} = \underbrace{d_s^R \left(\frac{\sigma_s}{\sigma_s - 1} \tau_s^{RH} c\right)^{1-\sigma}}_{\text{revenue}} - \underbrace{d_s^R \left(\frac{\sigma_s}{\sigma_s - 1} \tau_s^{RH} c\right)^{-\sigma} \tau_s^{RH} c}_{\text{variable costs}} - \underbrace{f_x^H}_{\text{fixed costs}}$$

$$\equiv \frac{d_s^R}{\sigma_s} \left(\frac{\sigma_s}{\sigma_s - 1} \tau^{RH} c\right)^{1-\sigma} - f_x^H \quad (9)$$

and thus the maximum marginal cost $c_{x,s}^{HR}$ above which it is not possible to profitably export from H in sector s is:

$$c_{x,s}^{HR} = \left(\frac{\sigma_s f_x^H}{d_s^R}\right)^{\frac{1}{1-\sigma_s}} \left(\frac{\sigma_s - 1}{\sigma_s \tau_s^{RH}}\right) \quad (10)$$

The corresponding condition for firms based in R is:

$$c_{x,s}^{RH} = \left(\frac{\sigma_s f_x^R}{d_s^H}\right)^{\frac{1}{1-\sigma_s}} \left(\frac{\sigma_s - 1}{\sigma_s \tau_s^{HR}}\right) \quad (11)$$

Firm heterogeneity is introduced via marginal production costs c, which follow a Pareto distribution with support $[0, \overline{c}]$.[8] For simplicity, this distribution is taken to be identical in both countries and across all sectors. The most productive firm in any sector therefore produces with zero marginal cost, while the least productive firm produces with the highest possible marginal cost $\overline{c}$.[9] This setup implies a cumulative distribution function and probability density function of the form:

$$G(c) = \left(\frac{c}{\overline{c}}\right)^k \quad (12)$$

and

$$g(c) \equiv \frac{dG(c)}{dc} = \frac{kc^{k-1}}{\overline{c}^k} \quad (13)$$

where $k > \sigma - 1$ is a "shape" parameter that in this context is effectively an index of firm heterogeneity: a higher value of k is associated with less pronounced heterogeneity.[10]

To close the production block of the model, we assume that free entry equates expected operating profits from domestic and export market sales with the entry cost f_e which firms must pay to enter the marginal cost "lottery". Home therefore has S free entry conditions of the following

[8] It is immaterial whether firm heterogeneity is introduced in terms of marginal costs, as here, or in terms of labor productivity as in Melitz (2003). See Baldwin and Harrigan (2007).

[9] The possibility of zero marginal cost production can be excluded by assuming random draws from a Pareto distribution truncated at a point $\underline{c} > 0$ (Helpman et al., 2007). This would make no material difference to our results.

[10] As usual, we assume $k > \sigma - 1$ for two reasons. First, since $\sigma > 1$, this assumption is sufficient to assure the existence of second moments for the marginal cost distribution. Second, it ensures the convergence of a number of integrals used below, for instance in the free entry conditions.

form:

$$\underbrace{\int_0^{c_{d,s}^H}\left(\frac{d_s^H}{\sigma_s}\left(\frac{\sigma_s}{\sigma_s-1}c\right)^{1-\sigma_s}-f_d^H\right)dG\left(c\right)}_{\text{Expected profits from domestic sales}}$$
$$+\underbrace{\int_0^{c_{x,s}^{HR}}\left(\frac{d_s^R}{\sigma_s}\left(\frac{\sigma_s}{\sigma_s-1}\tau_s^{RH}c\right)^{1-\sigma_s}-f_x^H\right)dG\left(c\right)}_{\text{Expected profits from domestic sales}}=f_e \quad (14)$$

The corresponding set of S conditions for R is simply:

$$\underbrace{\int_0^{c_{d,s}^R}\left(\frac{d_s^R}{\sigma_s}\left(\frac{\sigma_s}{\sigma_s-1}c\right)^{1-\sigma_s}-f_d^R\right)dG\left(c\right)}_{\text{Expected profits from domestic sales}}$$
$$+\underbrace{\int_0^{c_{x,s}^{RH}}\left(\frac{d_s^H}{\sigma_s}\left(\frac{\sigma_s}{\sigma_s-1}\tau_s^{HR}c\right)^{1-\sigma_s}-f_x^R\right)dG\left(c\right)}_{\text{Expected profits from domestic sales}}=f_e \quad (15)$$

3.3 Characterizing Export Diversification and Simulating the Model

As set out above, the model has $6S$ unknowns, namely $c_{d,s}^H$, $c_{d,s}^R$, $c_{x,s}^{HR}$, $c_{x,s}^{RH}$, d_s^H, and d_s^R. The free entry conditions (14) and (15) along with the domestic thresholds (7) and (8), and the export thresholds (10) and (11), constitute a system of $6S$ equations. Melitz (2003) and Bernard et al. (2007) provide details on the equilibrium properties of models of this type. Excluding corner solutions, equilibrium will involve both countries engaging in two way trade in each differentiated goods sector. On a sectoral level, firms in each country will self-select into three groups based on their marginal cost draw. For instance, H firms with $c > c_{d,s}^H$ will immediately exit without producing, while those with $c_{x,s}^H < c \leq c_{d,s}^H$ will produce for the domestic market only. Only those firms with $c \leq c_{x,s}^H$ will engage in export activity. This outcome fits well with the important stylized fact that only a small percentage of active firms in a country produce for overseas markets (see Bernard et al., Forthcoming, for a review of the evidence).

To characterize export diversification in equilibrium, we expand the price index in each country's demand shifter to make explicit its connection to the mass of locally produced and imported varieties:

$$d_s^H = \frac{\beta_s E^H}{\int\limits_{v \in V_s^H} p(v)^{1-\sigma_s} dv}$$

$$\equiv \frac{\beta_s E^H}{\underbrace{n_s^H \int_0^{c_{d,s}^H} \left(\frac{\sigma_s}{\sigma_s - 1} c\right)^{1-\sigma_s} dG(c)}_{\text{Locally produced varieties}} + \underbrace{n_s^R \int_0^{c_{x,s}^R} \left(\frac{\sigma_s}{\sigma_s - 1} c\tau_s^{HR}\right)^{1-\sigma_s} dG(c)}_{\text{Imported varieties}}} \tag{16}$$

$$d_s^R = \frac{\beta_s E^R}{\int\limits_{v \in V_s^R} p(v)^{1-\sigma_s} dv}$$

$$\equiv \frac{\beta_s E^R}{\underbrace{n_s^R \int_0^{c_{d,s}^R} \left(\frac{\sigma_s}{\sigma_s - 1} c\right)^{1-\sigma_s} dG(c)}_{\text{Locally produced varieties}} + \underbrace{n_s^H \int_0^{c_{x,s}^H} \left(\frac{\sigma_s}{\sigma_s - 1} c\tau_s^{RH}\right)^{1-\sigma_s} dG(c)}_{\text{Imported varieties}}} \tag{17}$$

Once the equilibrium values of d_s^H and d_s^R have been obtained by solution of the system of free entry conditions and cutoffs, the two above expressions have only two additional unknowns: n_s^H and n_s^R, i.e. the mass of firms operating in a particular sector in either country. Given n_s^H and n_s^R, it is straightforward to obtain expressions for the number of export producers in each country and sector ($n_{x,s}^H$ and $n_{x,s}^R$). In equilibrium, the proportion of exporting firms in the total is simply equal to the proportion of actively producing firms with $c \leq c_{x,s}^H$ in H and $c \leq c_{x,s}^R$ in R:

$$n_{x,s}^H = \frac{G\left(c_{x,s}^H\right)}{G\left(c_{d,s}^H\right)} n_s^H \tag{18}$$

$$n_{x,s}^R = \frac{G\left(c_{x,s}^R\right)}{G\left(c_{d,s}^R\right)} n_s^H \tag{19}$$

Since each firm produces a distinct product variety in each differentiated goods sector, we equate an increase in $n_{x,s}^H$ or $n_{x,s}^R$ with export diversification in that country. This definition of export diversification is equivalent to viewing that process as growth at the extensive, or "new products", margin of trade in terms of the recent work by Hummels and Klenow (2005) and Broda and Weinstein (2006) referred to above.

To examine the impact of various policy factors on export diversification as we have defined it, we analyze the model using numerical simulations. For simplicity, we use two countries of the same size (labor endowment) and consider a single sector only. For the underlying model parameters, we assume substantially the same values as in Bernard et al. (2007).[11] We then conduct the simulations by substituting the threshold conditions (7), (8), (10), and (11), along with the marginal cost cdf (12) and pdf (13), into the free entry conditions (14) and (15). The system is solved numerically for the demand shifters. Solutions are then substituted back to give equilibrium values for the production and export thresholds. The full set of solutions is then substituted into the demand shifter expressions (16) and (17), which are again solved numerically for the equilibrium mass of firms in each country. This in turn provides solutions for each country's level of export diversification through (18) and (19). We repeat this process separately over a pre-defined range for each of the key policy parameters, namely the fixed and variable costs of trade, and domestic market entry costs.[12]

Figure 2 presents summary results for this set of simulations in terms of approximate elasticities at each point in the simulation space.[13] Our choice of simulation scenarios is driven by the desire to provide a solid motivation for the empirical work that we undertake in the next section, so we concentrate on parameter changes that can easily be mapped to cross-country differences in observed policies. Thus, the first scenario we consider (solid line) is a unilateral increase in H's import-specific trade costs (τ_s^{HR}), which could be brought about, for example, by an increase in H's own tariff rate in sector s. By insulating H's market to some degree from international competition,

[11]We initially use: $\sigma = 3.8$, $k = 3.4$, $L^H = L^R = 1000$, $\beta = 0.5$, $\overline{c} = 0.2$, $f^e = 2$, and $f_d^H = f_d^R = 0.1$. To conform with the stylized facts that trade is costly and that only some firms export, we set $f_x^H = f_x^R = 0.15$ and $\tau^{HR} = \tau^{RH} = 1.1$.

[12]We simulate over the following ranges: $1.1 \leq \tau^{HR} \leq 1.2$, $0.15 \leq f_x^H \leq 0.25$, and $0.10 \leq f_d^H \leq 0.145$. Full results, including Maple worksheets, are available on request.

[13]For example, for the simulation based on changes to f_d^h, we calculate $\left(\Delta n_{x,s}^h / n_{x,s}^h\right) * \left(f_d^h / \Delta f_d^h\right)$ at each simulation point.

the overall mass of active firms increases and so, therefore, does the number of exporters–thus, diversification in H increases. The dashed line shows the impact of this policy change from R's perspective (i.e., a unilateral change in a partner country's tariffs): predictably, a smaller effective market for firms in that country leads to a decrease in diversification there.

Next we turn to changes to H's fixed costs of exporting ($f_{x,s}^{H}$) and domestic market entry ($f_{d,s}^{H}$). Changes to these parameters could be brought about by, for example, more complex export procedures that impose significant setup costs on firms due to the need to establish "boilerplate" documentation, or in the case of $f_{d,s}^{H}$ more complex and burdensome regulations governing firm establishment and registration. The dotted line ($f_{x,s}^{H}$) and the dashed-dotted line ($f_{d,s}^{H}$) show that the expected impact of increases along these two dimensions is to decrease export diversification in H, due to the higher barriers that potential producers and exporters must overcome in order to access foreign markets.

Finally, we consider a reciprocal increase in import-specific trade costs in both countries (τ_s^{HR} and τ_s^{RH}), such as could be brought about by a change in international transport costs (under the assumption that such costs are roughly symmetric). Alternatively, this scenario could be seen as representing the impact of an agreement for reciprocal tariff cuts. Results in Figure 2 (crossed line) show that the diversification impact of higher symmetric trade costs is unambiguously negative.

Putting all of these results together, we see that an increase in import-specific trade costs (such as a country's own tariffs) can promote diversification, while it is impeded by increases in export-specific trade costs (such as a country's own export formalities, or a partner country's tariffs), symmetric trade costs (such as those related to international transport), and domestic market entry costs (such as charges involved in establishing a commercial operation). Moreover, our simulations provide some suggestive evidence of the likely relative magnitudes involved: at moderate cost levels, the largest elasticity in absolute value is associated with symmetric trade costs, followed by export-specific trade costs, then import-specific trade costs, and finally domestic market entry costs. This is an intuitively sensible ordering, since the first two policies act directly on the incentives facing exporters, whereas the second two act only indirectly. Whilst the differences in magnitudes are relatively small for the first three categories, we nonetheless expect there to be a noticeable difference in magnitude between effects stemming from changes along those dimensions, and those associated with domestic market entry costs.

4 Empirical Model and Results

Our baseline empirical strategy to test the above predictions is a straightforward one. Since our diversification measure takes the form of count data–i.e., the number of 8-digit product lines exported in every 2-digit sector–we postulate that it follows a Poisson distribution with mean and variance equal to μ_{es} (where e indexes exporters and s indexes sectors).[14] Its density conditional on a set of independent variables $\mathbf{X}_{es}$ is:

$$f\left(lines\_cn2_{es} \mid \mathbf{X}_{es}\right) = \frac{\exp\left(-\mu_{es}\right)\mu_{es}^{lines\_cn2_{es}}}{lines\_cn2_{es}!} \tag{20}$$

We specify the conditional mean function μ_{es} in terms of the parameters of the theoretical model developed above and a set of sector fixed effects (δ_s) to control for unobserved heterogeneity affecting all exporters in a particular sector in the same way.[15] Thus:

$$\mu_{es} = \delta_s \exp\left[\begin{array}{c} \beta_1 \ln\left(entry_e\right) + \beta_2 \ln\left(export_e\right) + \beta_3 \ln\left(1 + t_{es}^{EU}\right) + \beta_4 \ln\left(dist_e\right) \\ +\beta_5 \ln\left(1 + t_{es}^{own}\right) + \beta_6 \ln\left(gdp\_hs2_e\right) + \beta_7 \ln\left(gdppc_e\right) \end{array}\right] \tag{21}$$

We use $entry_e$ to refer to restrictions on entry in a given exporting country (f_d^H), which we proxy using the *Doing Business* entry costs data referred to above. Export-specific trade costs (τ_s^{RH} and f_x^H) are proxied using MAcMap bilaterally disaggregated applied tariff data for the EU (t_{es}^{EU}) and *Doing Business* data on the costs of exporting ($export_e$). While ad valorem tariffs impose

[14] We estimate the model at the sectoral (2-digit) level, since three of the variables of interest–EU and own tariffs, and sectoral expenditures–vary at that level. However, we are conscious that the other variables of interest vary only at the country level, and so we adjust all standard errors for clustering by exporting country.

[15] We expect the sector fixed effects to account for influences such as trade-related measures applied on an MFN basis within our country sample (e.g., product standards), as well as the different numbers of 8-digit product lines included in each 2-digit sector within the CN classification.

only variable trade costs, the types of costs captured by *Doing Business* include both fixed and variable components: for instance, export documentation needs to be agreed and drafted prior to any export activity taking place (fixed cost), but then needs to be copied and slightly adapted for each shipment (variable cost). We use the distance between the exporting country and Europe ($dist_e$) as an indicator of the extent of symmetric trade costs ($\tau_s^{RH} = \tau_s^{HR}$) such as international transport charges, while sectoral expenditures ($\beta_s E^H$) and technology ($\bar{c}$ and k) are proxied by (respectively) GDP multiplied by a sectoral (2-digit) import share, and GDP per capita.[16] We can therefore summarize the core contentions of the paper as derived in the previous section in terms of restrictions on the parameter space of the conditional mean function: $\beta_5 > 0; \beta_1, \beta_2, \beta_3, \beta_4 < 0$ and $|\beta_4| > |\beta_2|, |\beta_3| > |\beta_5| >> |\beta_1|$.

The model can be estimated using standard maximum likelihood techniques.[17] The Poisson estimator has at least three desirable properties that make it particularly useful in this context. First, it has identical first order conditions to those obtained by running weighted non-linear least squares on (21) with $lines\_cn2$ in place of μ_{es} (Gourieroux et al., 1984). Thus, the various β coefficients can be interpreted as elasticities, but the model does not suffer from the usual limitation of log-log models in relation to zero observations: these can simply be included in the estimation sample as usual. Santos Silva and Tenreyro (2006) exploit this property as a means of dealing with zero trade flows in the gravity model context.

Second, the Poisson model has been shown to produce consistent estimates under relatively weak assumptions. It is not even necessary that the dependent variable in fact follow a Poisson distribution, so long as the conditional mean function is correctly specified (Gourieroux et al., 1984; Wooldridge, 1997). Other count data models, such as the negative binomial (Klinger and Lederman, 2004 and 2006), do not have this property and require the data to follow a particular distribution in order to ensure consistency. Third, Poisson is one of relatively few non-linear maximum likelihood

[16] To construct our tariff measures and sectoral expenditure proxies, we use the reference group approach of Laborde et al. (Forthcoming), which relies on observed imports for a group of similar countries so as to avoid the endogeneity inherent in using a simple import weighted average. Little turns on this choice, however, since our results do not change substantially if aggregate GDP or simple average tariffs are used instead. (Results available on request.)

[17] See generally: Cameron and Trivedi (2001), Wooldridge (1997), and Winkelmann (2000).

models in which standard, unconditional fixed effects do not lead to bias and inconsistency due to the "incidental parameters problem" (Greene, 2004). This is an important practical advantage in relation to the alternative empirical strategy of analyzing the extensive margin via a standard gravity model using a Heckman sample selection estimator (Helpman et al., 2007), a probit model (Debaere and Mostashari, 2005) or the Tobit estimator.[18] While it may be possible to avoid this problem by using a random effects estimator (Baldwin and Harrigan, 2007), the price that must be paid is a strong assumption as to the distribution of country- or sector-specific unobserved heterogeneity.

Poisson estimates of our baseline model are presented in column 1 of Table 3.[19] Overall, the model provides a close fit to the data, with R^2 equal to 0.92.[20] All coefficients carry the expected signs: entry costs, export costs, distance, and EU tariffs are negative, while own tariffs, GDP, and GDP per capita are positive. The coefficients on entry costs, export costs, distance, and GDP are all statistically significant at the 1% level, but the remaining coefficients are not significant at the 10% level. These results are broadly consistent with previous work examining the diversification impacts of GDP, per capita income, and trade costs, such as Hummels and Klenow (2005) and Feenstra and Kee (2006). Similarly, our finding that lower domestic market entry costs are associated with greater export diversification is consistent with the results of Helpman et al. (2007), who find a negative and statistically significant coefficient on market entry costs in the first stage of a sample selection gravity model.[21]

In terms of the magnitudes of the estimated coefficients on the trade and entry cost variables, Table 3 indicates that EU tariffs have the largest elasticity in absolute value terms (-0.6), followed

[18]Based on the evidence in Greene (2004), Debaere and Mostashari (2005) argue that the bias or inconsistency of their results is likely to be relatively minor in practice.

[19]We estimate the model cross-sectionally, using data for a single year (2005) only. While it would be desirable to estimate over multiple time periods, this is not currently possible given limited availability of our explanatory variables.

[20]We follow Wooldridge (1997) in using $R^2 = 1 - ESS/TSS$ as a convenient summary measure of fit. We prefer it to the more common (for count data) pseudo-R^2 measures due to its ease of interpretation. We are conscious, however, that it should be used with caution since the Poisson model does not contain a residual as such.

[21]Our results in respect of domestic market entry costs might appear at first glance to contrast with those of Klinger and Lederman (2004, 2006). However, those authors use a different dependent variable, namely a count of the number of 3-digit SITC categories in which a country exports for the first time over a given period. Thus, they view diversification through the lens of export "discoveries", rather than as an increase in the proportion of nationally produced varieties that make it to the international market. Their explanatory variable set is also different, including a wider set of entry cost data such as labor market rigidity and contract enforceability indicators, but they do not consider trade costs.

by distance and own tariffs (-0.4 and 0.4 respectively), then export costs (-0.3), and finally entry costs (-0.1). These magnitudes are clearly of economic significance, with (for example) a 10% decrease in export costs being associated with a 3% increase in export diversification. While this ordering of effects should clearly be interpreted with caution due to the relative imprecision with which the two tariff effects are estimated, it nonetheless moves in generally the same direction as the predictions from our theoretical model. The only major difference with the expected ordering based on simulation results is that EU tariffs have a larger elasticity (in absolute value terms) than does distance, but a Wald test of equality between the two coefficients does not reject the null (prob. = 0.89). However, when we test the hypothesis that the coefficients on entry costs, export costs, EU tariffs, and distance are equal, we find that the null is rejected at the 5% level (prob. = 0.03). This supports our earlier contention that the impact of entry costs on diversification is likely to be considerably weaker than for trade costs.

4.1 Instrumental Variables Strategy

One potential criticism of the results in Table 3 column 1 could be that our measures of trade and entry costs are potentially endogenous to export diversification. One reason might be political economy: a more diverse export base means a larger lobby in favor of lower export costs and higher import costs. Similarly, the results presented by Imbs and Wacziarg (2003), Funke and Ruhwedel (2001), or Feenstra and Kee (2006) could be interpreted as suggesting endogeneity of per capita income with respect to export diversification.

To deal with these possibilities, we adopt two complementary approaches. First, we use five year lags of GDP and per capita GDP, since these measures should be genuinely exogenous with respect to the current level of export diversification. Next, we use an instrumental variables strategy to identify exogenous variation in our measures of entry and export costs (including own tariffs).[22] To estimate the Poisson model with instrumental variables, we follow the procedure set out in Wooldridge (2002, pp.663-665), which is analogous to two-stage least squares. The first stage con-

[22] We treat distance and EU tariffs as exogenously determined.

sists of OLS regressions in which the potentially endogenous variables–entry costs, export costs, and own tariffs–are used sequentially as the dependent variable, while the exogenous variables from the baseline model along with the instruments are used as the independent variables. The residuals from these regressions are retained, and entered into the baseline Poisson formulation (21) as additional regressors (i.e., retaining the potentially endogenous measures of trade and entry costs). Conditional on the use of appropriate instruments, inclusion of the first stage residuals removes any endogeneity bias. A simple test of the joint significance of the residual terms can be used as an indication of the endogeneity of the variables being instrumented for: rejection of the null hypothesis indicates that there is a serious endogeneity problem.

As usual, the most difficult aspect of this approach lies in identifying variables that satisfy the twin conditions for instrument validity: strong correlation with the potentially endogenous trade and entry costs, but orthogonality to current export diversification. Our strategy exploits variation in geography and colonial history to identify exogenous movements in the variables of interest. The idea is that a country's economic and social institutions should be correlated with distance from the equator (Hall and Jones, 1999), as well as with the marks left by colonial rule on legal regimes and institutions (Djankov et al., 2002).[23] Thus we expect both instruments to be correlated with the currently observed levels of trade and entry costs, while being exogenous to current export diversification. In our first stage regressions (Table 4), we therefore include distance from the equator and binary dummies for colonization by Great Britain, France, Spain, Portugal, the Netherlands, and Russia in addition to the exogenous variables from the baseline model. The first stage R^2 is 0.4 for entry costs, and 0.3 for export costs and own tariffs, while F-tests of the null hypothesis that all instruments jointly equal zero are 5.47***, 1.88*, and 4.65*** respectively. While our instruments do a fair job of explaining the observed pattern of entry and trade costs, it would obviously be desirable in future work to try and identify stronger instruments. This is particularly true in the case of export costs, for which the F-test only rejects the null hypothesis at the 10% level. Nonetheless,

[23] In fact, Djankov et al. (2002) use legal origin dummies, not colonization, as instruments for market entry costs. However, those measures display insufficient variation in our sample, and so we prefer to use closely related data on colonization drawn from Mayer and Zignago (2006).

we continue with geography and colonial history here in order to get a first impression of the extent to which instrumental variables estimates might have the potential to alter our baseline results.

Results from the second stage regression are in column 2 of Table 3, with the coefficients on the three sets of first stage residuals suppressed for brevity. The coefficients on entry costs, export costs, distance, and GDP all carry the expected sign, and (with the exception of entry costs) are statistically significant at the 5% or 1% levels. EU and own tariffs, as well as GDP per capita, carry unexpected coefficient signs, but are not statistically significant at the 10% level. While we are cautious in interpreting changes in coefficient magnitude given the relative imprecision with which some of them are estimated, we observe that the elasticities of export diversification with respect to export costs and distance increase in absolute value terms, but the reverse is true for entry costs. Finally, we test the joint significance of the three first-stage residuals using a standard Wald test, and find that the null hypothesis that they equal zero cannot be rejected at the 10% level (prob. = 0.24). This suggests that the empirical impact of endogeneity would appear to be relatively minor in this particular case.

Taking all these results together, we conclude that the data generally support our core contention in so far as it relates to trade and entry costs: we find evidence of a negative impact of both factors on export diversification, with the estimated elasticity being larger in absolute value for export costs and international transport (proxied by distance) than for entry costs. The evidence in relation to own and EU tariffs is much more mixed: the estimated parameters are not statistically significant in either formulation, and they change sign from one model to the other. While there could be many possible reasons for this, one likely candidate is the lack of observed variance in the EU's applied tariffs. Since our sample is composed exclusively of non-OECD and non-EU25 countries, a very large number of applied bilateral tariffs (over 70%) are in fact equal to zero once full account is taken of preferential rates. It is therefore difficult to draw any firm conclusions in respect of importer tariffs from this dataset, and it will be important to return to that question in future work.

In addition to examining potential endogeneity problems, it is also important to ensure that our results are robust in other respects. We now perform checks along four additional dimensions:

alternative measures of export diversification and entry/trade costs, additional dependent variables, changes in the country sample, and application of alternative estimators.

4.2 Alternative Measures of Diversification, Entry Costs, and Trade Costs

The results discussed above use *lines_cn2* as the dependent variable, i.e. a count of the number of 8-digit product lines exported to the EU in each 2-digit sector. Table 5 shows that our results are not sensitive to this choice. The first two columns use a more narrowly defined dependent variable, in which 8-digit export flows are only considered to be non-zero if they are greater than $100,000 or $1,000,000 respectively. The third column uses the original definition, but excludes agriculture (chapters 1-24) from consideration. Although there are some small changes in coefficient estimates as a result of these alterations to the dependent variable, Table 5 shows that the substance of our results is unchanged: entry and export costs, as well as distance, exert a negative and statistically significant impact on diversification, while the positive impact of own tariffs and the negative impact of EU tariffs are statistically insignificant.

The final two columns of Table 5 take completely different approaches to measuring diversification. Column 4 uses the relative variety measure of Feenstra and Kee (2006), which is closely related to the measures used by Hummels and Klenow (2005) and Broda and Weinstein (2006). Column 5, on the other hand, replaces *lines_cn2* with the Herfindahl index of concentration that is frequently used as an inverse indicator of export diversification in policy circles (e.g., UNCTAD, 2006; and Amin Gutiérrez de Piñeres and Ferrantino, 1997). Both measures are calculated at the 2-digit sector level using 8 digit data, applying the formulae set out in section 2. As can be seen from the table, both alternative measures of diversification produce qualitatively identical results to those in our baseline formulation. The estimated signs are the same in all cases except EU tariffs, while entry costs, export costs, distance, and GDP remain statistically significant at the 5% or 1% level.

There are suggestions in the literature (e.g., Klinger and Lederman, 2006) that measures of the number of administrative procedures required to enter the domestic market may be more accurate than the corresponding cost data. We therefore re-estimate our baseline formulation using alternative *Doing Business* data on the number of procedures required for market entry, and the number of documents required for exports. Results in the third column of Table 3 are qualitatively identical to the baseline, although the coefficient on export costs is now marginally significant at the 10% level (prob. = 0.100). The signs on EU tariffs and GDP per capita change, but these coefficients remain statistically insignificant.

4.3 Additional Independent Variables

While the set of explanatory variables used in our baseline formulation accords well with the theoretical model developed in this paper, we are conscious that the literature discloses a number of additional factors that might be expected to impact export diversification. For instance, Feenstra and Kee (2006) include measures of factor endowments as instruments for export variety. Given that the mechanism driving diversification in our theoretical model is investment related, it may also be appropriate to take account of macroeconomic conditions. We therefore augment our baseline model to include data on the percentage of manufacturing and agriculture in GDP–as a proxy for economic structure as determined by factor endowments–plus the GDP deflator and real interest rate as indicators of macroeconomic performance. Results are presented in the last two columns of Table 3, and are qualitatively identical to those from our baseline formulation as regards trade and market entry costs. Moreover, results in the last two columns indicate that a larger manufacturing sector and a lower GDP deflator are both associated with greater export diversity, and that these effects are statistically significant at the 1% level. The two other additional variables are statistically insignificant. While the size of the agricultural sector carries the expected negative sign in column 4, both that variable and the real interest rate have unexpected positive signs in column 5.

4.4 Narrower Country Samples

The definition of developing country that we have used thus far is a wide one: all countries that are not members of the OECD or the EU-25. It is important to ensure that our results continue to hold using a more narrow approach that might take better account of the differing situations of developing countries according to their income level. Progressively narrower definitions are applied moving from left to right across the first three columns of Table 6, based on country income groups as defined by the World Bank. The first column excludes high income countries, the second excludes in addition upper-middle income countries, while the third includes only the low income group. There is very little substantive change from our baseline results. Interestingly, the estimated elasticities for entry and export costs become larger in absolute value as the income group becomes poorer, which suggests that these factors may be particularly important for low income developing countries–exactly the group with the most significant policy interest in diversification. While own tariffs are positive (but statistically insignificant) in all three columns, EU tariffs only have the expected negative sign in the first two columns. We suspect that the counter-intuitive result in column 3 (low income countries only) is due to the very high proportion of zeros in the EU's applied tariff matrix: over 85% vis-à-vis the low income group, versus 70% for the full sample. More puzzling is the distance coefficient in column 3, which carries an unexpected sign and is statistically insignificant. This result varies starkly with all others that we report, which show distance as having a negative and 1% significant impact on diversification. We can only surmise that it is a function of greatly reduced sample size–about 1/3 of the full dataset–and the relative lack of variance this introduces into the distance data.

4.5 Alternative Estimators

As a final check, we examine the robustness of our results to the use of alternative estimators. As suggested above, Poisson is consistent under relatively weak assumptions, and is not adversely affected by the inclusion of unconditional fixed effects. However, the possibility remains of gains in efficiency through the use of alternative estimators that have been employed in the literature. Columns 4-6 of Table 6 present results using standard OLS, Tobit, and the negative binomial model.[24] Once again, there is no substantive change from our baseline results in column 1 of Table

[24] Tobit estimates should be interpreted with caution since we include unconditional fixed effects. For the negative binomial estimates, we employ the conditional fixed effects approach of Hausman et al. (1984). Robust standard errors

3. Entry costs, export costs, and distance all impact export diversification negatively and significantly (at the 1% level), while own tariffs have a positive impact which is statistically significant at the 5% level under OLS and at the 1% level using the negative binomial. EU tariffs enter with an unexpected positive sign, but are not statistically significant in any of the three models.

5 Conclusions and Policy Implications

We have used highly disaggregated mirror data from the EU, along with new information from the World Bank's *Doing Business* database, to show that trade costs and entry restrictions are inversely related to export diversification in a sample of 118 developing countries. We find that export costs, international transport costs (proxied by distance), and domestic market entry costs all impact negatively and significantly on export diversification, with the strongest effects working through transportation (with an elasticity of -0.4) and export costs (-0.3). By contrast, the elasticity on market entry costs is only -0.1. We find much more mixed evidence in support of the expected positive impacts of lower EU tariffs and higher own tariffs.

Our core results turn out to be quite robust to alternative diversification measures, model specifications, and country samples. Although we are cautious in drawing causal inferences based on cross-sectional data for a single year, instrumental variables estimates using geography and colonial history to identify exogenous variation in trade and entry costs would appear to provide preliminary support for such a conclusion. This interpretation would moreover be in accordance with the theoretical structure we have developed in the paper, incorporating heterogeneous firms and asymmetric trade costs.

Subject to that caveat, our results can be used to sketch out some preliminary implications for development policy. On the one hand, policies aimed at reducing trade costs related specifically to exporting, as well as international transport costs, would appear to be promising avenues for developing countries aiming to promote export diversification. Trade facilitation measures, in the broad sense in which Wilson et al. (2005) use that term, would therefore be attractive. Similarly,

are obtained using interval regression for Tobit, while boostrapping (200 replications) is used for the negative binomial.

measures aimed at reducing the costs of domestic market entry are also promising, although our results suggest that their impact will be less strong than for export costs or international transport costs. The unifying factor behind both sets of measures is that they envisage diversification not as a result of governments "picking winners" through industrial policy, but as a natural outcome of winners picking themselves through an intensification of the Schumpeterian process at the heart of the Melitz (2003) model.

It is interesting that the empirical evidence in favor of lower EU tariffs or higher home country tariffs is considerably weaker than for the other trade cost factors we have identified. This suggests that circumspection may be required in pursuing either more generous trade preferences or stronger import protection in the hopes of promoting export diversification. It may be, as Collier and Venables (2007) suggest, that it is not necessarily more generous preferential rates that are required, but complementary policies–such as rules of origin–that make preferences easier and cheaper to access. However, this is a point that we leave for future research to address more fully in the light of additional evidence comparing the extensive margin impacts of different preferential schemes. Since *Doing Business* data on entry and trade costs are being updated annually, we are hopeful that future research will also be able to exploit the availability of panel data both to assist with achieving identification, and to provide greater clarity on the dynamics involved in the diversification process.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson, 2001, "The Colonial Origins of Comparative Development: An Empirical Investigation", *The American Economic Review*, 91(5), 1369-1401.

Amin Gutiérrez de Piñeres, Sheila, and Michael Ferrantino, 1997, "Export Diversification and Structural Dynamics in the Growth Process: The Case of Chile", *Journal of Development Economics*, 52(2), 375-391.

Baldwin, Richard E., and James Harrigan, 2007, "Zeros, Quality, and Space: Trade Theory and

Trade Evidence", Working Paper No. 13214, NBER.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott, Forthcoming, "Firms in International Trade", *Journal of Economic Perspectives*.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott, 2007, "Comparative Advantage and Heterogeneous Firms", *Review of Economic Studies*, 74, 31-66.

Broda, Christian, and David E. Weinstein, 2006, "Globalization and the Gains from Variety", *The Quarterly Journal of Economics*, 121(2), 541-585.

Cameron, A. Colin, and Pravin K. Trivedi, 2001, "Essentials of Count Data Regression" in Badi H. Baltagi (ed.), *A Companion to Theoretical Econometrics*, Malden, Ma.: Blackwell.

Collier, Paul, and Anthony J. Venables, 2007, "Rethinking Trade Preferences: How Africa Can Diversify its Exports", *The World Economy*, 30(8), 1326-1345.

Debaere, Peter, and Shalah Mostashari, 2005, "Do Tariffs Matter for the Extensive Margin of International Trade? An Empirical Analysis", Discussion Paper No. 5260, CEPR.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer, 2002, "The Regulation of Entry", *The Quarterly Journal of Economics*, 117(1), 1-37.

Djankov, Simeon, Caroline Freund, and Cong S. Pham, 2006, "Trading on Time", Policy Research Working Paper No. 3909, The World Bank.

Evenett, Simon J., and Anthony J. Venables, 2002, "Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows", Mimeo.

Feenstra, Robert C., 1994, "New Product Varieties and the Measurement of International Prices", *The American Economic Review*, 84(1), 157-177.

Feenstra, Robert C., and Hiau Looi Kee, 2006, "Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity", Mimeo.

Feenstra, Robert C., and Hiau Looi Kee, 2007, "Trade Liberalization and Export Variety: A Comparison of Mexico and China", *The World Economy*, 30(1), 5-21.

Fontagné, Lionel, Michael Freudenberg, and Nicolas Péridy, 1998, "Intra-Industry Trade and the Single Market: Quality Matters", Discussion Paper No. 1959, CEPR.

Funke, Michael, and Ralf Ruhwedel, 2001, "Product Variety and Economic Growth: Empirical Evidence for the OECD Countries", *IMF Staff Papers*, 48(2), 225-242.

Gourieroux, C., A. Monfort, and A. Trognon, 1984, "Pseudo Maximum Likelihood Methods: Applications to Poisson Models", *Econometrica*, 52(3), 701-720.

Greene, William, 2004, "The Behavior of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects", *Journal of Econometrics, 7, 98-119.*

Hall, Robert E., and Charles I. Jones, 1999, "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, 114(1), 83-116.

Hausman, Jerry, Bronwyn H. Hall and Zvi Griliches, 1984, "Econometric Models for Count Data with an Application to the Patents-R&D Relationship", *Econometrica*, 52(4), 909-938.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik, 2006, "What You Export Matters", Working Paper, http://ksghome.harvard.edu/~drodrik/hhr.pdf.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, 2007, "Trading Partners and Trading Volumes", Mimeo, http://post.economics.harvard.edu/faculty/helpman/papers/TradingPartners.pdf.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple, 2004, "Export Versus FDI With Heterogeneous Firms", *The American Economic Review*, 94(1), 300-316.

Henry de Frahan, Bruno, and Mark Vancauteren, 2006, "Harmonization of Food Regulations and Trade in the Single Market: Evidence from Disaggregated Data", *European Review of Agricultural Economics*, 33(3), 337-360.

Hummels, David, and Peter J. Klenow, 2005, "The Variety and Quality of a Nation's Exports", *The American Economic Review*, 95(3), 704-723.

Imbs, Jean and Romain Wacziarg, 2003, "Stages of Diversification", *The American Economic Review*, 93(1), 63-86.

Kehoe, Timothy J., and Kim J. Ruhl, 2003, "How Important is the New Goods Margin in International Trade", Research Department Staff Report No. 324, Federal Reserve Bank of Minneapolis.

Klinger, Bailey, and Daniel Lederman, 2004, "Discovery and Development: An Empirical Exploration of New Products", Policy Research Working Paper No. 3450, The World Bank

Klinger, Bailey, and Daniel Lederman, 2006, "Diversification, Innovation and Imitation Inside the Global Technological Frontier", Policy Research Working Paper No. 3872, The World Bank

Laborde, David, Houssein Boumelassa, and Maria C. Mitaritonna, Forthcoming, "A Consistent Picture of the Protection Across the World in 2004: MAcMap HS6 V2", Working Paper, CEPII.

Mayer, Thierry, and Soledad Zignago, 2006, "Notes on CEPII's Distance Measures", Working Paper, CEPII.

Melitz, Marc J., 2003, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71(6), 1695-1725.

Ozden, Caglar, and Eric Reinhardt, 2005, "The Perversity of Preferences: GSP and Developing Country Trade Policies, 1976-2000", *Journal of Development Economics*, 78(1), 1-21.

Prebisch, Raúl, 1959, "Commercial Policy in the Underdeveloped Countries", *The American Economic Review*, 49(2), 251-273.

Rodrik, Dani, 2004, "Industrial Policy for the Twenty-First Century", Mimeo, http://ksghome.harvard.edu/~drodrik/UNIDOSep.pdf.

Santos Silva, J.M.C. and Silvana Tenreyro, 2006, "The Log of Gravity", *Review of Economics and Statistics*, 88(4), 641-658.

UNCTAD, 2006, *UNCTAD Handbook of Statistics*, Geneva: United Nations.

Wilson, John S., Catherine L. Mann, and Tsunehiro Otsuki, 2005, "Assessing the Benefits of Trade Facilitation: A Global Perspective", The World Economy, 28(6), 841-871.

Winkelmann, Rainer, 2000, *Econometric Analysis of Count Data*, 3rd edition, Berlin: Springer.

Wooldridge, Jeffrey M., 1997, "Quasi-Likelihood Methods for Count Data" in M.H. Pesaran and P. Schmidt (eds.), *Handbook of Applied Econometrics*, Vol. 2, Oxford: Blackwell.

Wooldridge, Jeffrey M., 2002, *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

Tables

Table 1: Countries included in the sample, sorted by quintile of lines.

Range	Countries
0-118	Belize, Bhutan, Burundi, Central African Republic, Chad, Comoros, Djibouti, Dominica, Eritrea, Grenada, Guinea-Bissau, Haiti, Kiribati, Lesotho, Marshall Islands, Micronesia, Palau, Rwanda, Samoa, Serbia, Solomon Islands, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, São Tomé and Principe, Tajikistan, Tonga, Vanuatu
131-281	Afghanistan, Antigua and Barbuda, Armenia, Benin, Botswana, Burkina Faso, Equatorial Guinea, Fiji, Gambia, Guinea, Guyana, Iraq, Kyrgyz Republic, Malawi, Maldives, Mali, Mauritania, Mozambique, Nicaragua, Niger, Papua New Guinea, Seychelles, Sierra Leone, Sudan, Swaziland, Togo, Zambia
296-685	Angola, Azerbaijan, Bolivia, Cambodia, Cameroon, Cape Verde, Congo, Rep., El Salvador, Ethiopia, Gabon, Georgia, Guatemala, Honduras, Jamaica, Kazakhstan, Lao PDR, Mongolia, Namibia, Panama, Paraguay, Suriname, Tanzania, Trinidad and Tobago, Uganda, Uzbekistan, Yemen, Zimbabwe
746-1815	Albania, Algeria, Bangladesh, Belarus, Bosnia and Herzegovina, Colombia, Costa Rica, Côte d'Ivoire, Dominican Republic, Ecuador, Ghana, Jordan, Kenya, Kuwait, Lebanon, Macedonia, FYR, Madagascar, Mauritius, Moldova, Nepal, Nigeria, Oman, Peru, Senegal, Syria, Uruguay, Venezuela
1876-8053	Argentina, Brazil, Bulgaria, Chile, China, Croatia, Egypt, Hong Kong, China, India, Indonesia, Iran, Israel, Malaysia, Morocco, Pakistan, Philippines, Romania, Russia, Saudi Arabia, Singapore, South Africa, Sri Lanka, Thailand, Tunisia, Ukraine, United Arab Emirates, Vietnam

Table 2: Data and sources.

Variable	Description	Units (Yr.)	Source
Ag. % GDP	Agriculture value added as % GDP.	% (2005)	WDI
Col_*	Equal to one if a country was colonized by * (UK, France, Spain, Portugal, the Netherlands, or Russia).	NA	CEPII
Distance	Average of the great circle distances between the main cities of the exporting country and Germany, weighted by population shares.	Km	CEPII
Entry Cost	Official cost of starting up and formally operating an industrial or commercial business in the exporting country.	USD (2006)	Doing Business
Export Cost	Official fees levied on a 20 foot container leaving the exporting country. Includes document preparation costs, administrative fees for customs clearance and technical control, terminal handling charges, and inland transit.	USD (2006)	Doing Business
GDP	Gross domestic product.	USD (2005)	WDI
GDP Defl.	GDP deflator.	% (2005)	WDI
GDPPC	Per capita GDP.	USD (2005)	WDI
Lat.	Latitude of the main city in the exporting country (absolute value).	Deg.	CEPII
Lines	Number of 8-digit product lines in which a country has strictly positive exports to the EU.	NA (2005)	Eurostat
Lines_CN2	Number of 8-digit product lines in a 2-digit sector for which a country has strictly positive exports to the EU.	NA (2005)	Eurostat
Manuf. % GDP	Manufacturing value added as % GDP.	% (2005)	WDI
Real Int. Rate	Real interest rate.	% (2005)	WDI
Tariffs	Average applied ad valorem tariff by HS2 sector. Aggregated from 6-digit data using the reference group methodology of Laborde et al. (Forthcoming).	% (2005)	MAcMap

Table 3: Baseline estimation results.

	Baseline	IV	Alt. 1	Alt. 2	Alt. 3
Entry Cost	-0.142***	-0.06		-0.128**	-0.098*
	[0.055]	[0.102]		[0.054]	[0.051]
Export Cost	-0.322***	-0.918***		-0.424***	-0.280***
	[0.092]	[0.351]		[0.094]	[0.087]
Entry Procs.			-0.394**		
			[0.156]		
Export Docs.			-0.408		
			[0.248]		
Distance	-0.418***	-0.512***	-0.364***	-0.444***	-0.455***
	[0.088]	[0.096]	[0.099]	[0.078]	[0.060]
Tariff (EU)	-0.639	0.414	0.335	-1.207	-0.217
	[1.541]	[1.595]	[1.514]	[1.555]	[1.398]
Tariff (Own)	0.431	-1.608	0.436	0.424	0.319
	[0.418]	[2.674]	[0.429]	[0.312]	[0.296]
GDP	0.452***		0.504***	0.462***	0.487***
	[0.028]		[0.025]	[0.026]	[0.020]
GDPPC	0.087		-0.004	0.037	0.022
	[0.054]		[0.051]	[0.067]	[0.060]
GDP 2000		0.434***			
		[0.045]			
GDPPC 2000		-0.086			
		[0.094]			
Manuf. %				0.385***	0.345***
				[0.105]	[0.084]
Ag. %				-0.04	0.01
				[0.080]	[0.063]
GDP Defl.					-0.247***
					[0.071]
Real Int. Rate					0.024
					[0.041]
Constant	0.466	6.822	8.136	0.674	0.103
	[1.213]	[4.192]	[9.352]	[1.175]	[1.046]
Obs.	11328	11328	11328	10752	9024

1. Estimation is by Poisson, with dependent variable lines_cn2. Independent variables are in logarithms. All models include fixed effects by 2-digit sector. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%).

2. Specification IV uses colonization dummies and distance from the equator as instruments for entry and export costs, and own tariffs. First stage R^2 is 0.4, 0.3, and 0.3 respectively.

Table 4: First stage instrumental variables regressions.

	Entry Cost	Export Cost	Own Tariff
Distance	-0.551***	-0.005	-0.004
	[0.204]	[0.081]	[0.014]
Tariff (EU)	-1.734	-1.466*	-0.145
	[1.678]	[0.795]	[0.131]
GDP 2000	-0.149***	-0.084***	0.001
	[0.050]	[0.022]	[0.003]
GDPPC 2000	0.415***	-0.100**	-0.014***
	[0.085]	[0.038]	[0.005]
Col. GBR	0.306	-0.225*	0.012
	[0.276]	[0.134]	[0.016]
Col. FRA	0.746***	-0.161	0.026
	[0.262]	[0.138]	[0.021]
Col. ESP	1.532***	-0.227	-0.012
	[0.481]	[0.184]	[0.021]
Col. PRT	1.328**	0.039	-0.015
	[0.668]	[0.162]	[0.019]
Col. NLD	1.494***	-0.315	-0.029
	[0.544]	[0.251]	[0.019]
Col. RUS	-0.572**	0.353**	-0.050***
	[0.284]	[0.178]	[0.016]
Latitude	-0.197*	-0.047	-0.002
	[0.101]	[0.063]	[0.005]
Constant	10.191***	9.134***	0.302***
	[2.131]	[0.801]	[0.112]
Obs.	11328	11328	11328
R2	0.43	0.33	0.29

1. Estimation is by OLS with dependent variables as indicated. All variables are in logarithms. All models include fixed effects by 2-digit sector.

2. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%).

Table 5: Estimation results using alternative measures of export diversification.

	>$100k	>$1m	Manuf. Only	Lambda	HH Index
Entry Cost	-0.254***	-0.314***	-0.142**	-0.017**	0.052**
	[0.086]	[0.094]	[0.056]	[0.007]	[0.022]
Export Cost	-0.506***	-0.644***	-0.327***	-0.039***	0.161***
	[0.107]	[0.088]	[0.095]	[0.012]	[0.034]
Distance	-0.537***	-0.569***	-0.456***	-0.050***	0.165***
	[0.127]	[0.141]	[0.089]	[0.012]	[0.034]
Tariff (EU)	-1.413	-0.518	-0.406	0.309	-0.678
	[2.355]	[2.850]	[1.947]	[0.213]	[0.736]
Tariff (Own)	0.424	0.462	0.335	0.075	-0.092
	[0.727]	[0.845]	[0.709]	[0.056]	[0.174]
GDP	0.604***	0.672***	0.461***	0.058***	-0.166***
	[0.039]	[0.036]	[0.029]	[0.004]	[0.012]
GDPPC	0.058	-0.004	0.094	0.014*	-0.021
	[0.075]	[0.071]	[0.058]	[0.008]	[0.023]
Constant	-0.102	-0.125	1.188	0.021	-0.512
	[1.602]	[1.612]	[1.305]	[0.149]	[0.438]
Observations	11328	11328	8496	7659	7642

1. Estimation in columns 4-5 is by OLS with dependent variables lambda_cn2 and hh_ index_cn2. Estimation in columns 1-3 is by Poisson, with dependent variable lines_cn2. Columns 1 and 2 only count export flows greater than $100k and $1m respectively, while column 3 excludes agricultural products (HS chapters 1-24). All models include fixed effects by 2-digit sector.

2. Only the independent variables are in logarithms in columns 1-3, while both the dependent and independent variables are transformed in columns 4-5.

3. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%).

Table 6: Additional robustness checks.

	Low + Middle	Low + Low. Mid.	Low	OLS	Tobit	Neg. Bin.
Entry Cost	-0.117*	-0.141*	-0.207**	-0.090**	-0.124***	-0.133***
	[0.064]	[0.075]	[0.102]	[0.035]	[0.046]	[0.008]
Export Cost	-0.329***	-0.365**	-0.563**	-0.287***	-0.396***	-0.352***
	[0.115]	[0.149]	[0.233]	[0.069]	[0.088]	[0.023]
Distance	-0.460***	-0.359***	0.228	-0.262**	-0.383***	-0.326***
	[0.086]	[0.115]	[0.342]	[0.102]	[0.121]	[0.024]
Tariff (EU)	-0.158	-0.539	2.91	2.287	1.634	0.088
	[1.643]	[1.683]	[3.528]	[1.383]	[1.628]	[0.649]
Tariff (Own)	0.206	0.01	0.106	0.595**	0.522	0.499***
	[0.462]	[0.581]	[0.707]	[0.273]	[0.379]	[0.161]
GDP	0.461***	0.434***	0.520***	0.352***	0.485***	0.390***
	[0.031]	[0.045]	[0.045]	[0.021]	[0.026]	[0.012]
GDPPC	0.113*	0.172*	0.084	0.086**	0.085*	0.101***
	[0.066]	[0.100]	[0.227]	[0.040]	[0.049]	[0.012]
Constant	0.591	0.332	-4.376	0.361	0.256	-1.816***
	[1.304]	[1.594]	[4.126]	[1.177]	[1.499]	[0.386]
Observations	10656	8256	3744	11328	11328	11328

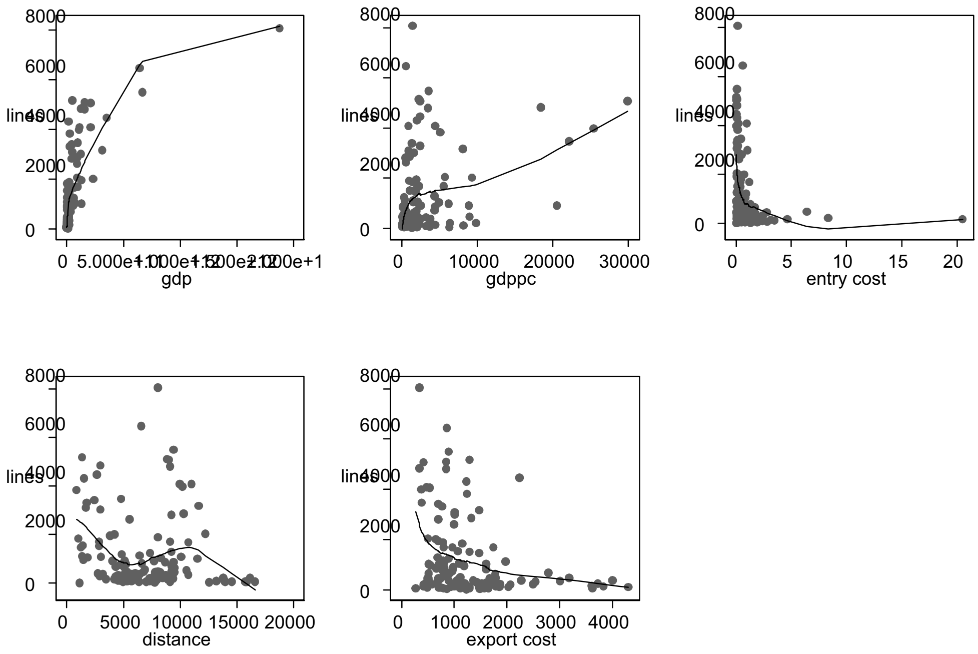
1. Estimation in columns 1-3 is by Poisson, using data on low and middle income, low and lower-middle income, and low income countries only. Estimation in columns 4-6 is by OLS, Tobit, and negative binomial, using the full sample. All models include fixed effects by 2-digit sector.

2. All independent variable are in logarithms. The dependent variable is in logarithms in columns 4-5 only.

3. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%).

Figures

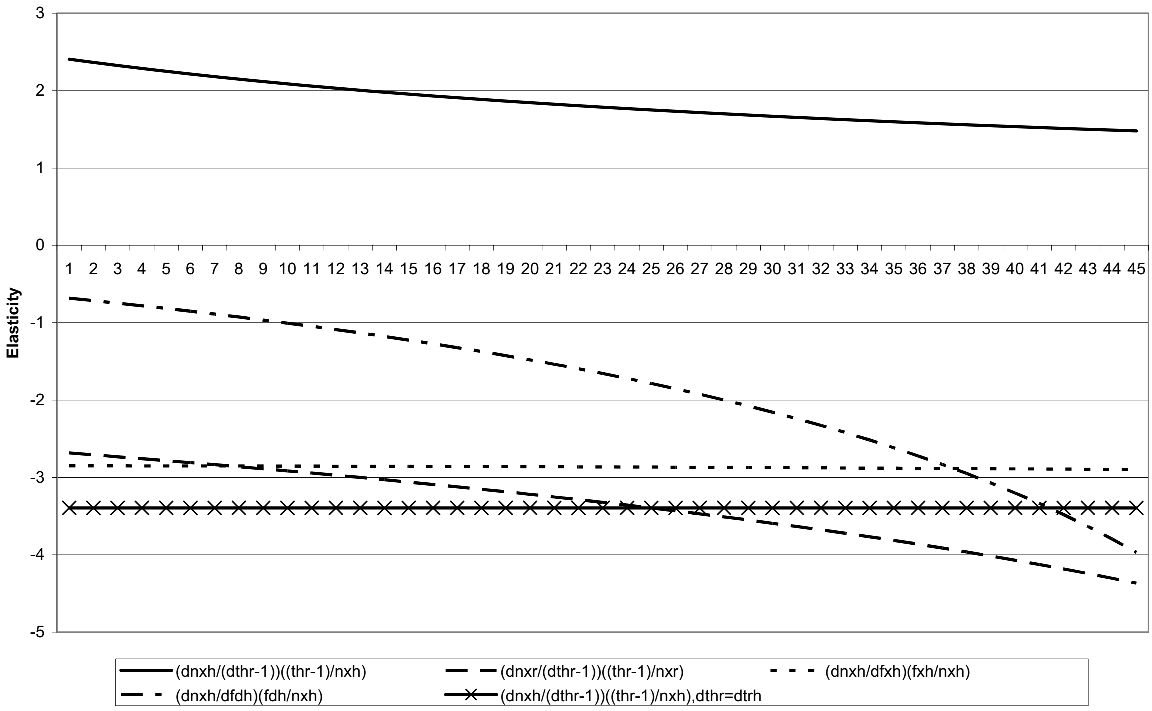
Figure 1: Non-parametric regression results.



1. The y-axis of each plot is lines, and the x-axis is as marked. Entry cost is expressed in percent per capita GNI terms, and all other independent variables are in simple levels.

2. All results were obtained by Lowess regression with bandwidth=0.8.

Figure 2: Simulation results, expressed as approximate elasticities.



1. Simulations were conducted as set out in the main text. Full results (including Maple worksheets) are available on request.

2. In the terminology of the main text, the solid line equates to an increase in import-specific costs, the dotted and dashed lines refer to increases in export-specific costs, the crossed line is an increase in symmetric trade costs, and the dashed-dotted line is an increase in the costs of domestic market entry.