Working Paper DTC-2009-2

Trade Facilitation and Export Diversification

Ben Shepherd, Principal.

Allen Dennis, Economist, World Bank.

24 June 2009.

developing trade consultants
policy • research • capacity building

349 5th Avenue
New York, NY 10016
Ben@Developing-Trade.com

Trade Facilitation and Export Diversification

Allen Dennis* and Ben Shepherd† ‡

This version dated: 24 June 2009

Abstract

This paper shows that improved trade facilitation can help promote export diversification in developing countries, where this issue has long been an important concern of policymakers. We find that 10% reductions in the costs of international transport and exporting (documentation, inland transport, and port/customs charges) are associated with export diversification gains of 4% and 3% respectively in a sample of 118 developing countries. Lower market entry costs can also promote diversification, but the effect is weaker (1%). Our results are highly robust to estimation using alternative dependent and independent variables, different country samples, and alternative econometric techniques. We link these findings to recent advances in trade theory that emphasize firm heterogeneity, and trade growth at the extensive margin.

JEL codes: F12; F13; O24.

Keywords: International trade; Economic development; Product variety; Trade policy.

1 Introduction

Export diversification has long been a policy concern for developing countries. Dependence of export revenues on just a handful of products–often primary commodities–is perceived as creating

*Economist (Doing Business), World Bank, 1818 H. St. NW, Washington, DC 20433. Contact: adennis@worldbank.org.

†Postdoctoral Research Associate, Niehaus Center for Globalization and Governance, Princeton University, Princeton, NJ 08544. Contact: bashepherd@gmail.com (corresponding author).

‡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 comments: Simeon Djankov, Matthias Helble, Bernard Hoekman, Will Martin, Beata S. Javorcik, Christian Volpe Martincus, and John S. Wilson. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent.

excessive short-run volatility in national income. In addition, commodity dependence is frequently associated with lower growth rates over the long-run, and stagnation at relatively low levels of per capita income. Recent empirical evidence lends substance to these concerns. 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. Recent work by Cadot et al. (2007) finds similar evidence for exports. Funke and Ruhwedel (2001) find that export diversification is positively related to per capita GDP and TFP growth in OECD countries.

If diversification is positively associated with a country's growth and development prospects, what are the policy options available to support that process? This paper shows that trade facilitation can be highly effective in promoting export diversification in developing countries. We use a broad definition of trade facilitation, in line with the approach taken by forums such as the Asia-Pacific Economic Cooperation (APEC): trade facilitation is any policy that reduces the transaction costs of international trade (Wilson et al., 2005). To measure it, we use data from the World Bank's *Doing Business* project that capture all official costs that must be paid in order to export a standardized container-load of goods (Djankov et al., Forthcoming). We use data from the same source on the official costs of starting a business to proxy the cost of market entry (Helpman et al., 2008). Concretely, we find that reducing by 10% the costs of exporting, international transport, or market entry, can increase export diversification by 3%, 4%, and 1% respectively. Our results prove to be highly consistent across a series of robustness checks covering alternative samples, alternative independent and dependent variable sets, and alternative estimation methods.

These findings build on and extend two recent literatures. On the one hand, existing work on trade facilitation has focused exclusively on its potential to promote growth in existing trade flows (Wilson et al., 2005; Djankov et al., Forthcoming). One implication of our findings is that existing analysis tends to underestimate the potential of trade facilitation measures, since it does not take account of their potential to promote export diversification as well. In addition, our characterization of export diversification in terms of trade growth at the extensive ("new products") margin, we can develop some important policy implications that flow from recent theoretical work on heterogeneous firms in international trade (Melitz, 2003). Existing empirical work using such models shows that larger, richer countries systematically trade in a wider range of goods (Hummels and Klenow, 2005), and that growth in import variety can lead to substantial national welfare gains (Broda and Weinstein, 2006). However, the only trade costs that have been considered by the empirical literature on extensive margin growth are international transport charges and tariffs (Debaere and Mostashari, 2005; and Feenstra and Kee, 2008).[1]

The paper proceeds as follows. In the next section, we develop a definition of export diversification in terms of trade growth at the extensive margin. We draw on recent theoretical and empirical work to develop a series of hypotheses that we proceed to test in the remainder of the paper. Section 3 presents our dataset and sets out our approach to measuring export diversification, export costs, and market entry costs. Our empirical strategy is discussed in Section 4, and estimation results are presented. Section 5 concludes and puts forward some suggestions for future research in this area.

2 Conceptualizing Export Diversification

Empirical work on export diversification has tended to use intuitively appealing, but theoretically ad hoc, measures of diversification, such as a Herfindahl-Hirschman Index of export values across a given range of products or sectors (e.g., Amin Gutiérrez de Piñeres and Ferrantino, 1997). The intent behind such measures is clear: they are designed to pick up the extent to which different countries rely more or less heavily on a small range of products for the bulk of their export earnings. From a theoretical perspective, there are two mechanisms by which such measures of diversification can increase (Cadot et al., 2007). One possibility is that export goods at low aggregate values grow more quickly than those goods at high aggregate values. This kind of diversification can be thought of as occurring at the intensive margin, in the sense that it does not involve export-

[1]After the working paper version of the present paper appeared, Persson (2008) compared the influence of trade facilitation on the intensive and extensive margins of trade using similar data and methods to ours, but focusing on the distinction between homogeneous and differentiated goods.

ing any new products. A second way in which diversification can occur is through the extensive margin: introducing new products into a country's export bundle necessarily increases its level of diversification, keeping all else constant.

From a development policy perspective, extensive margin diversification is particularly important. Indeed, we would argue that it forms the core of the concept. Export diversification is often linked with a shift in export composition from primary commodities to manufactured goods (Collier and Venables, 2007). Such a process necessarily involves extensive margin diversification. It is for this reason that we focus exclusively on the extensive, or "new products", margin in this paper.

In addition to policy relevance, our focus on the extensive margin also has another important advantage: it enables us to leverage the recent literature to better understand the policy factors behind export diversification. It is now well established, for instance, that not all countries export all products to all other countries, and thus bilateral trade matrices tend to contain a large number of zero entries (Haveman and Hummels, 2004). The broader the country sample, and the greater the level of product disaggregation, the more common zero entries become. There is thus considerable scope for export diversification at the extensive margin, in particular in developing countries.

In an important contribution, Helpman et al. (2008) have recently shown that a heterogeneous firms model in the spirit of Melitz (2003) provides a powerful explanation for the presence of zeros in the trade matrix, and thus, conversely, for the fact that some countries have more diversified export bundles than others. The basic mechanism behind their model is straightforward. In a world with multiple countries and multiple sectors, each firm manufactures a unique product under Dixit-Stiglitz monopolistic competition. Firms can sell in their home market without paying transport costs, but if they sell overseas they face "iceberg" trade costs, such as transport charges, and tariffs in the importing country. Firms differ in their marginal cost of production, which can be thought of as having been drawn at random from a given probability distribution. By assuming that productivity is drawn from a Pareto distribution, it is possible to ensure that low-productivity firms are relatively common, and high-productivity firms are relatively uncommon, in equilibrium. This accords well with the available empirical evidence.

Fixed market entry costs differ from country to country. Firms that enter only one market need only pay one entry cost, whereas firms that enter multiple markets must pay multiple entry costs. Based on the combination of market entry costs, trade costs, and productivity distributions, individual firms self-select into three groups: those with very low productivity cannot profitably produce for any market, and so drop out of the model altogether; those with intermediate productivity draws service the domestic market only; and those with high productivity serve the domestic market, and, in addition, at least one export market. Because of the assumption that each firm manufactures a distinct product, the more high productivity firms a given country has, the wider the range of products that it exports to at least one foreign market. Another way of putting this condition is in terms of productivity cutoffs: the lower the threshold level of productivity required in order to profitably export to at least one overseas market, the more firms become exporters in equilibrium, and the greater the aggregate range of products the country exports.

The productivity cutoff for a given country-pair depends intimately on the level of trade and market entry costs facing producers in both markets. Helpman et al. (2008) show that higher market entry costs tend to increase the productivity cutoff, thereby making it more difficult for domestic firms to access foreign markets. Similarly, trade costs facing exporters–such as international transport charges, export costs, and import country tariffs–make it more difficult for firms to start exporting. On the other hand, tariffs in the exporting country can potentially protect local producers from foreign competition, and in so doing make it easier for them to export.[2]

From a policy point of view, the Helpman et al. (2008) model thus suggests two main hypotheses of interest:

1. Lower levels of trade costs are associated with a more diverse export bundle; and

2. Lower market entry costs are associated with a more diverse export bundle.

In the empirical part of their paper, Helpman et al. (2008) do not test either of these hypotheses.

[2]We are not suggesting that "infant industry" protection is a sensible policy option to promote export diversification. The analysis we have presented is based on comparative statics, and thus does not take account of the many dynamic inefficiencies that are associated with such policies in practice.

Instead, they estimate a two-stage gravity model which shows that the factors listed above indeed affect the probability that two countries trade with each other (i.e., that their entry in the bilateral trade matrix is not zero). However, they use aggregate trade data for their regressions. Thus, although their results are consistent with a positive association between both sets of factors and export diversification, it is nonetheless impossible to draw any strong conclusions about product mix.[3] Part of the value added of the present paper lies exactly in its focus on the new products margin: by dealing explicitly with export diversification, it extends the results of Helpman et al. (2008) in a policy-relevant way. This paper also complements recent work by Kee and Feenstra (2008) showing that trade costs such as tariffs and distance affect export diversification in the sense in which we are using that term. It does so by expanding the range of policy variables beyond tariffs to include market entry costs, and trade facilitation.

3 Data and Measurement

Full details of our dataset, for which many of the sources are standard, are presented in Tables 1 and 2. In this section, we focus on two aspects of our data that are novel: measures of export diversification covering 118 developing countries, and measures of the costs of exporting and domestic market entry for those same countries.

3.1 Measuring Export Diversification

In terms of the characterization adopted in the previous section, export diversification means an increase in the range of products a country exports.[4] The most obvious approach to measurement is to simply count the number of exported products for each country. (We return to the question of

[3] Similar comments apply to the results of Debaere and Mostashari (2005): although they use disaggregated trade data, their use of a logit model makes it impossible to distinguish between trade growth at the "new products" and "new markets" margins.

[4] We limit consideration to growth in the number of products exported, and do not examine the separate question of the number of markets to which countries export.

alternative approaches below, when we check the robustness of our empirical results.) 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. However, 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.[5] We extract these data from a freely available Eurostat database covering exports from and imports into the European Union.[6] To our knowledge, these data have not previously been used in product variety work, where the focus has been either on cross-country data at the 6-digit level, or on more detailed US data.

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.[7] In this paper, we focus

[5] 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.

[6] See http://fd.comext.eurostat.cec.eu.int/xtweb/.

[7] We use only a single year of data due to limited availability of our explanatory variables, in particular export costs (see below). Combined cross-section time-series estimates are not currently feasible, but will become possible as more trade cost data are published.

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_cn*2 following the same pattern as for *lines*, but with counts by 2-digit sector rather than aggregated to the country level. The *lines_cn*2 variable 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_cn*2 as an indicator of export diversification.

In Table 3, 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. 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.

3.2 Measuring Export and Market Entry Costs

We use new data from the World Bank's *Doing Business* database to measure trade costs, which we conceptualize as an inverse measure of broad sense trade facilitation. 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. To our knowledge, these data have not previously been used in empirical work, although Djankov et al. (Forthcoming) 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 bilateral trade negatively.

These *Doing Business* 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 $1,278 per container (excluding OECD and EU countries).

To measure 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 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. (2008): they find that higher entry costs are negatively associated with the probability that two countries engage in trade.

4 Empirical Model and Results

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

$$f\left(lines\_cn2_{es} \mid \mathbf{X}_{es}\right) = \frac{\exp\left(-\mu_{es}\right)\mu_{es}^{lines\_cn2_{es}}}{lines\_cn2_{es}!} \tag{1}$$

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.[9] Thus:

$$\mu_{es} = \delta_s \exp\left[\begin{array}{c} \beta_1 \ln\left(entry_e\right) + \beta_2 \ln\left(export_e\right) + \beta_3 \ln\left(1+t_{es}^{EU}\right) + \beta_4 \ln\left(dist_e\right) \\ +\beta_5 \ln\left(1+t_{es}^{own}\right) + \beta_6 \ln\left(gdp\_hs2_{es}\right) + \beta_7 \ln\left(gdppc_e\right) \end{array}\right] \tag{2}$$

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

[8]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. Estimating at the aggregate (country) level does not change our conclusions materially (results available on request).

[9]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.

for each shipment (variable cost). We use the distance between the exporting country and Europe ($dist_e$) as an indicator of the extent of other international trade costs such as international transport charges, while sectoral expenditures and technology are proxied by (respectively) GDP multiplied by a sectoral (2-digit) import share (gdp_hs2_{es}), and GDP per capita ($gdppc_e$).[10] We can therefore summarize the core contentions of this paper as follows: $\beta_1, \beta_2, \beta_3, \beta_4 < 0$.

Maximum likelihood estimates of our baseline Poisson model are presented in column 1 of Table 4.[11] Overall, the model provides a close fit to the data, with R^2 equal to 0.92.[12] All coefficients–which can be interpreted as elasticities–carry the expected signs: entry costs, export costs, distance, and EU tariffs are all negative, while sectoral expenditures, and GDP per capita are positive. The coefficients on entry costs, export costs, distance, and sectoral expenditures are all statistically significant at the 1% 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), Debaere and Mostashari (2005), and Feenstra and Kee (2008). Similarly, our finding that lower market entry costs are associated with greater export diversification is consistent with the results of Helpman et al. (2008), who find a negative and statistically significant coefficient on market entry costs in the first stage of a sample selection gravity model.[13]

A number of factors may contribute to the lack of statistical significance of the remaining coefficient estimates. GDP per capita is not a perfect proxy for the state of technology, and does not allow

[10]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. Our results do not change substantially if aggregate GDP or simple average tariffs are used instead. (Results available on request.)

[11]The Poisson estimator has identical first order conditions to those obtained by running weighted non-linear least squares on (2) with $lines\_cn2$ in place of μ_{es} (Wooldridge, 1997). Thus, 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. This is an important point, since $lines\_cn2$ contains a relatively high proportion of zeros (4,825 out of 13,192 observations, or 37%).

[12]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.

[13]Our results for market entry costs might appear to contrast with those of Klinger and Lederman (2006). However, those authors take a different approach in which 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, but excluding trade costs.

us to distinguish between sector-specific and economy-wide dimensions of technological progress. In the case of EU tariffs, the extensive availability of preferences for developing countries results in a data series with a very large number of zero entries (9,686 out of 13,056 observations, or 74%). The resulting lack of variation makes it difficult to obtain precise estimates. Data on exporter tariffs exhibit greater variation, but may conflate two separate effects: protection of final production can potentially increase diversification through a kind of import substitution mechanism, but protection of intermediate inputs tends to increase production costs, and thereby reduce diversification. Moreover, the Lerner symmetry suggests that, over time, an anti-import bias can in fact inhibit exports. Given the lack of detailed data on effective rates of protection, however, it is not currently possible to identify these effects separately.

In terms of the magnitudes of the estimated coefficients on the trade and entry cost variables, Table 4 indicates that EU tariffs have the largest elasticity in absolute value terms (-0.6), followed 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 suggests that trade facilitation–which reduces export and market entry costs–has considerable potential to boost export diversification.

4.1 Additional Exporter Country Characteristics

While the set of explanatory variables used in our baseline formulation accords well with intuition and the theoretical models discussed in Section 2, we are conscious that the literature discloses a number of additional factors that might be expected to impact export diversification.[14] For instance, Feenstra and Kee (2008) include measures of factor endowments as instruments for export variety.

[14] As new data become available, it will be possible to use fixed effects to account for time-invariant country characteristics. At this stage, however, the best that can be done is to add more variables covering potentially important dimensions such as economic structure and macroeconomic policy. In additional results (available on request), we add random effects by exporter to the baseline specification and find results that are largely similar to those reported here.

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 two additional sets of control variables. First, we include data on the percentage of manufacturing and agriculture in GDP as proxies for economic structure as determined by factor endowments, institutions, and development history. Then in addition, we include the GDP deflator and real interest rate as indicators of macroeconomic performance. Results are presented in columns 2 and 3 of Table 4, and are qualitatively identical to those from our baseline formulation as regards trade and market entry costs. Moreover, these results 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. These additional findings accord well with basic intuition. Although the size of the agricultural sector carries the expected negative sign in column 2, both that variable and the real interest rate have unexpected positive signs in column 3. However, neither coefficient is statistically significant.

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

In addition to checking the robustness of our results to the inclusion of additional exporter characteristics, it is also important to ensure that they are not unduly sensitive to the way in which our primary variables of interest are measured. For instance, there are suggestions in the literature (e.g., Klinger and Lederman, 2006) that measures of the number of administrative procedures required to enter the market may be more accurate than the corresponding cost data. To take account of this possibility, Table 5 column 1 re-estimates our baseline model using these alternative data. Specifically, we use *Doing Business* data on the number of procedures required for market entry, and the number of documents required for exports. Results are qualitatively identical to the baseline, although the coefficient on export costs is now borderline significant at the 10% level (prob. = 0.100). The signs on EU tariffs and GDP per capita change, but these coefficients remain statistically insignificant.

The results discussed thus far use *lines_cn*2 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 the way in which we have operationalized our definition of export diversification. Columns 2-3 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. Although there are some small changes in coefficient estimates as a result of these alterations to the dependent variable, 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 both statistically insignificant.

In column 4, we show that our results are largely unchanged when we run the regression using data on the manufacturing sector only, i.e. excluding all data from HS Chapters 1-24. We do this to deal with the possibility that the theoretical models discussed in Section 2–which rely on product differentiation and monopolistic competition–apply more naturally to the industrial sector than to agriculture and commodities. Thus, our measures of diversification may be more appropriate to the former than to the latter.[15] As an additional robustness check in this direction, we re-estimate the baseline model separately for each HS 2-digit sector. Results in Table 7 generally support our hypotheses: the average estimates of entry costs and export costs across all 96 sectoral regressions are negative and of similar magnitude to those reported in Table 4 column 1. Entry costs are 10% significant in 57 regressions, while export costs are statistically significant in 65 regressions out of a total of 96.

The final two columns of Table 5 take different approaches to measurement of the dependent variable, again using 8-digit EU import data to construct diversification measures by 2-digit chapter. We start in column 5 with the theory-consistent measure of relative variety proposed by Feenstra (1994) and used with modifications by, for instance, Hummels and Klenow (2005), and Broda and

[15] In additional robustness checks, available on request, we have also interacted entry and export costs with 2-digit average elasticities of substitution from Broda and Weinstein (2006). Coefficients for both cost terms remain negative, statistically significant, and close to their baseline values. The interaction terms are positive, and statistically significant in the case of export costs. These results strengthen our conclusions, and highlight their particular relevance for strongly differentiated products.

Weinstein (2006). We adopt the formulation used by Feenstra and Kee (2008), 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:

$$\Lambda = \frac{\sum_{i \in J^H} p_i^w q_i^w}{\sum_{i \in J^W} p_i^w q_i^w} \tag{3}$$

The numerator in this measure is 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. In a cross-sectional setting like this one, Λ can only change due to differences in export composition across countries.

Column 6, on the other hand, replaces $lines\_cn2$ with a Herfindahl index of export concentration (using J as the set of products exported by a given country):

$$hh\_index = \sum_{i=1}^{J} \left(\frac{p_i q_i}{\sum_{j=1}^{J} p_j q_j} \right)^2 \tag{4}$$

Although lacking a theoretical basis of the type provided by Feenstra (1994) for Λ, this measure is sometimes used in policy and applied work (e.g., Amin Gutiérrez de Piñeres and Ferrantino, 1997; Cadot et al., 2007).

As can be seen from the table, both alternative measures of diversification produce qualitatively identical results to those in our baseline formulation. (Since the Herfindahl index is a measure of concentration, i.e. an inverse measure of diversification, we expect the pattern of signs to be reversed from the baseline.) The estimated signs are as expected in all cases except EU tariffs, while entry costs, export costs, distance, and GDP remain statistically significant at the 5% or 1% level. This is unsurprising, given how closely correlated our $lines\_cn2$ measure is with the two alternatives: $\rho = 0.95$ for Λ and $\rho = -0.52$ for the Herfindahl index.

4.3 Alternative Country Samples

The definition of a 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 hold using a more focused 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 World Bank country income groups. 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 again 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.

4.4 Alternative Estimators

As a final set of robustness checks, we examine the consistency of our results across a range of alternative econometric estimators. Columns 4-6 of Table 6 present results using standard OLS, Tobit, and the negative binomial model. Once again, there is no substantive change from our baseline results in column 1 of Table 4. 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.

To deal with the possibility that our measures of trade and market entry costs might be endogenous–for example, due to political economy considerations–we re-estimate the baseline model with two changes. 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). 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 consists 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 (2) as additional regressors. Standard errors for the second stage are obtained by bootstrapping (500 replications) 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.

We exploit 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).[16] Thus we expect both in-

[16] In fact, Djankov et al. (2002) use legal origin dummies, not colonization, as instruments for market entry costs.

struments 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 8 columns 2-4), we 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. F-tests of the null hypothesis that all instruments jointly equal zero are 5.47***, 1.88*, and 4.65*** respectively. Results from the second stage regression are in column 1 of Table 8, 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. Distance and GDP have coefficients that are statistically significant at the 1% level, and the export costs coefficient is significant at the 15% level. Entry costs, however, has a statistically insignificant coefficient. 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, and should not call into question the results of the baseline model discussed above. Given the inevitable loss of precision involved in instrumental variables estimation, we therefore tend to prefer the baseline results.

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).

5 Conclusion

The results presented in Section 4 show that the data strongly support the core contention of this paper, namely that export costs and market entry costs, as well as international transport costs, impact negatively on developing country export diversification. Thus, improved trade facilitation represents a set of policy options that would appear to have significant scope to promote export diversification. Interestingly, the evidence in relation to own and EU tariffs is much more mixed: the estimated parameters are not statistically significant in most formulations, and they change sign from one model to the other. This tends to suggest that the type of improved access to developed country markets that developing countries currently receive may be more limited as an export diversification policy than is trade facilitation (cf. Collier and Venables, 2007). Similarly, "infant industry" protection seems to be a riskier proposition than is trade facilitation in terms of promoting export diversification. The unifying factor behind the set of policies we have considered in this paper 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.

Our results can be used to sketch out areas that future work could explore in greater depth. On the one hand, additional work on complementary market access policies, such as relaxing rules of origin (de Melo and Portugal-Perez, 2008), would be needed before drawing a strong conclusion as to the efficacy of developed country trade policy reforms as a means of helping developing countries pursue export diversification (Collier and Venables, 2007). Since *Doing Business* data on entry and trade costs are being updated annually, we are also hopeful that future research will be able to exploit the availability of panel data to assist with achieving identification and controlling for unobserved cross-country heterogeneity, and to provide greater clarity on the dynamics involved in the diversification process.

References

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

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

Cadot, Olivier, Céline Carrère, and Vanessa Strauss-Kahn, "Export Diversification: What's Behind the Hump?" CEPR Discussion Paper 6590 (2007).

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

Debaere, Peter, and Shalah Mostashari, "Do Tariffs Matter for the Extensive Margin of International Trade? An Empirical Analysis," CEPR Discussion Paper 5260 (2005).

De Melo, Jaime A.P., and Alberto Portugal-Perez, "Rules of Origin, Preferences, and Diversification in Apparel: African Exports to the US and to the EU," CEPR Discussion Paper 7072 (2008).

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

Djankov, Simeon, Caroline Freund, and Cong S. Pham, "Trading on Time," *The Review of Economics and Statistics* (Forthcoming).

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

Feenstra, Robert C., and Hiau Looi Kee, "Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity," Journal of International Economics, 74:2 (2008), 500-518.

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

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

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics* 123:2 (2008), 441-487.

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

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

Klinger, Bailey, and Daniel Lederman, "Diversification, Innovation and Imitation Inside the Global Technological Frontier," World Bank Policy Research Working Paper 3872 (2006).

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

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

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

Persson, Maria, "Trade Facilitation and the Extensive and Intensive Margins of Trade," Working Paper, Lund University Department of Economics (2008).

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

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

Tables

Table 1: 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 HS2	Sectoral expenditure, proxied by gross domestic product multiplied by sectoral import shares derived using the reference groups methodology of Laborde et al. (Forthcoming).	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 2: Descriptive statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
Ag. %	12610	17.92	12.76	0.07	61.83
Col. ESP	13192	0.13	0.34	0	1
Col. FRA	13192	0.21	0.4	0	1
Col. GBR	13192	0.36	0.48	0	1
Col. NLD	13192	0.02	0.15	0	1
Col. PRT	13192	0.04	0.21	0	1
Distance	13095	6825.18	3405.65	853.23	16597.13
Entry Cost	13192	3193.04	25367.69	43.03	297400.9
Entry Procs.	13192	10.21	3.22	3	20
Export Cost	12804	1278.33	804.92	265	4300
Export Docs.	12804	7.76	2.39	3	16
GDP HS2	11520	6.24E+08	5.50E+09	6.69E+03	2.88E+11
GDP HS2 2000	11520	4.77E+08	3.82E+09	6.75E+03	1.84E+11
GDP Defl.	13192	9.3	20.87	-8.27	240.27
GDPPC	12998	2761.85	4635.01	104.64	29944.97
GDPPC 2000	12901	2451.49	4193.23	109.32	25319.42
Lat. Exp.	13192	11.76	21.41	-34.92	55.75
Lines_CN2	13192	11.74	37.37	0	944
Manuf. %	12319	12.63	7.48	0.38	37.18
Real Int. Rate	10864	7.78	8.04	-12.15	44.93
Tariff (EU)	13056	0.01	0.02	0	0.39
Tariff (Own)	11520	0.14	0.17	0	11.44

1. Lines_CN2 contains 4825 observations equal to zero. The corresponding figure for Tariffs (EU) is 9686, and for Tariffs (Own) is 684.

Table 3: Countries 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 4: Baseline estimation results.

	Baseline	+ Controls 1	+ Controls 2
Entry Cost	-0.142***	-0.128**	-0.098*
	[0.055]	[0.054]	[0.051]
Export Cost	-0.322***	-0.424***	-0.280***
	[0.092]	[0.094]	[0.087]
Entry Procs.			
Export Docs.			
Distance	-0.418***	-0.444***	-0.455***
	[0.088]	[0.078]	[0.060]
Tariff (EU)	-0.639	-1.207	-0.217
	[1.541]	[1.555]	[1.398]
Tariff (Own)	0.431	0.424	0.319
	[0.418]	[0.312]	[0.296]
GDP HS2	0.452***	0.462***	0.487***
	[0.028]	[0.026]	[0.020]
GDPPC	0.087	0.037	0.022
	[0.054]	[0.067]	[0.060]
GDP HS2 2000			
GDPPC 2000			
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	0.674	0.103
	[1.213]	[1.175]	[1.046]
Obs.	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%).

Table 5: Estimation results using alternative measures of trade costs, and export diversification.

	*lines_cn*2	*lines_cn*2 >$100k	*lines_cn*2 >$1m	*lines_cn*2 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]
Entry Procs.	-0.394**					
	[0.156]					
Export Docs.	-0.408^					
	[0.248]					
Distance	-0.364***	-0.537***	-0.569***	-0.456***	-0.050***	0.165***
	[0.099]	[0.127]	[0.141]	[0.089]	[0.012]	[0.034]
Tariff (EU)	0.335	-1.413	-0.518	-0.406	0.309	-0.678
	[1.514]	[2.355]	[2.850]	[1.947]	[0.213]	[0.736]
Tariff (Own)	0.436	0.424	0.462	0.335	0.075	-0.092
	[0.429]	[0.727]	[0.845]	[0.709]	[0.056]	[0.174]
GDP HS2	0.504***	0.604***	0.672***	0.461***	0.058***	-0.166***
	[0.025]	[0.039]	[0.036]	[0.029]	[0.004]	[0.012]
GDPPC	-0.004	0.058	-0.004	0.094	0.014*	-0.021
	[0.051]	[0.075]	[0.071]	[0.058]	[0.008]	[0.023]
Constant	8.136	-0.102	-0.125	1.188	0.021	-0.512
	[9.352]	[1.602]	[1.612]	[1.305]	[0.149]	[0.438]
Observations	11328	11328	11328	8496	7659	7642

1. Estimation in columns 5-6 is by OLS with dependent variables lambda_cn2 and hh_ index_cn2. Estimation in columns 1-4 is by Poisson, with dependent variable lines_cn2. Columns 2 and 3 only count export flows greater than $100k and $1m respectively, while column 4 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 1-4, while both the dependent and independent variables are transformed in columns 5-6.

3. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%). ^ indicates borderline significance at the 10% level (prob. = 0.100).

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 HS2	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%).

Table 7: Regression results by sector.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.	10% Signif.
Entry Cost	96	-0.17	0.09	-0.55	-0.01	57
Export Cost	96	-0.31	0.22	-0.80	0.32	65
Distance	96	-0.43	0.27	-1.05	0.25	73
Tariffs (EU)	96	-3.83	22.31	-104.11	97.05	21
Tariffs (Own)	96	-0.17	1.83	-6.39	7.28	26
GDP HS2	96	0.47	0.17	-0.06	1.02	92
GDPPC	96	-0.14	1.28	-10.18	1.51	42

1. Each line in the table presents summary statistics for a single coefficient estimate from regressions conducted by individual HS-2 sectors. The specification is the same as the baseline model in Table 4 column 1, but without sector fixed effects.

Table 8: Instrumental variables regressions.

	Stage 2		Stage 1	
	$lines_{cn2}$	Entry Cost	Export Cost	Own Tariff
Entry Cost	-0.06			
	[0.146]			
Export Cost	-0.918+			
	[0.613]			
Distance	-0.512***	-0.551***	-0.005	-0.004
	[0.122]	[0.204]	[0.081]	[0.014]
Tariff (EU)	0.414	-1.734	-1.466*	-0.145
	[2.077]	[1.678]	[0.795]	[0.131]
Tariff (Own)	-1.608			
	[4.111]			
GDP HS2 2000	0.434***	-0.149***	-0.084***	0.001
	[0.080]	[0.050]	[0.022]	[0.003]
GDPPC 2000	-0.086	0.415***	-0.100**	-0.014***
	[0.133]	[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	11328
R2		0.43	0.33	0.29

1. Estimation in column 1 is by Poisson with the residuals from columns 2-4 included as additional regressors. Standard errors are obtained by bootstrapping (500 replications). Estimation in columns 2-4 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 + (15%), * (10%), ** (5%), and *** (1%).

Additional Results for Referees (Not For Publication)

	Aggregate	Random Effects	Sigma
Entry Cost	-0.133**	-0.373***	-0.183***
	[0.059]	[0.076]	[0.067]
Entry Cost * Sigma			0.017
			[0.013]
Export Cost	-0.282***	-1.010***	-0.560***
	[0.101]	[0.156]	[0.120]
Export Cost * Sigma			0.100***
			[0.022]
Distance	-0.434***	-0.403***	-0.418***
	[0.087]	[0.126]	[0.088]
Tariffs (EU)	0.089	-1.424	-0.65
	[6.187]	[0.921]	[1.543]
Tariffs (Own)	-0.233	0.548**	0.433
	[1.099]	[0.260]	[0.418]
GDP	0.470***		
	[0.038]		
GDP HS2		0.018	0.452***
		[0.044]	[0.028]
GDPPC	0.089	0.408***	0.087
	[0.062]	[0.104]	[0.054]
Constant	1.74	10.343***	1.087
	[1.424]	[1.777]	[1.306]
Observations	118	11328	11328

1. Dependent variable in column 1 is *lines*. Estimation is by Poisson, with robust standard errors in brackets.

2. Dependent variable in columns 2-3 is *lines_cn2*. Estimation in column 2 is by Poisson with fixed effects by 2-digit sector, and random effects by exporter. Standard errors are estimated by bootstrapping (500 replications). Estimation in column 3 is by Poisson with fixed effects by 2-digit sector, with robust standard errors corrected for clustering by exporter. Sigma is the 2-digit sectoral average of the 10-digit elasticities of substitution estimated by Broda and Weinstein (2006).

3. All independent variables are in logarithms. Statistical significance is indicated by * (10%), ** (5%), and *** (1%).