research paper series

Research Paper 2015/13

Trade Costs and the Composition of Developing Countries' Exports

By

Salamat Ali and Chris Milner

The Authors

Salamat Ali is a PhD candidate in the School of Economics, University of Nottingham. Chris Milner is a Professor International Economics at the School of Economics, University of Nottingham.

Trade Costs and the Composition of Developing Countries' Exports

by

Salamat Ali and Chris Milner

Abstract

This study investigates how trade costs fashion the export composition of the developing countries. It uses the World Bank's bilateral trade cost dataset and incorporates trade flows of a large set of developing countries. We exploit the variation in trade costs across countries and the differences in trade cost sensitivity across industries in the identification strategy. Moreover, we employ a set of cost shifters as instruments to overcome the endogeneity of trade costs. The paper finds that high trade cost countries gain a relatively lower export share overall compared to lower trade cost countries, and importantly that the share of exports declines more for more trade cost sensitive exports. The policy implications are clear: reducing trade costs would both increase the manufactured exports of developing economies and alter the composition of these exports.

JEL Classification: F14

Key words: Trade costs, export composition

Outline

1. Introduction

Trade costs have long been of interest to researchers and economic policy makers. Their importance is highlighted by the mystery of 'missing trade' (Trefler, 1995; McCallum, 1995) and by some of the puzzles in international macro-economics (Obstfeld and Rogoff, 2001). Further, the main focus of research has been on the trade volume effects of trade costs. The trade- deterring effect of bilateral distance between trading partners (proxying for the effects of trade costs) is extensively identified in the gravity model literature (Helpman et al., 2008). This literature has also explored the effects of specific types of trade costs (e.g. trade policies, common borders, different languages, different currencies, port efficiency etc.) on trade volumes. For example, Limao and Venables (2001) investigate the influence of poor infrastructure on the volume of intra- Africa trade.

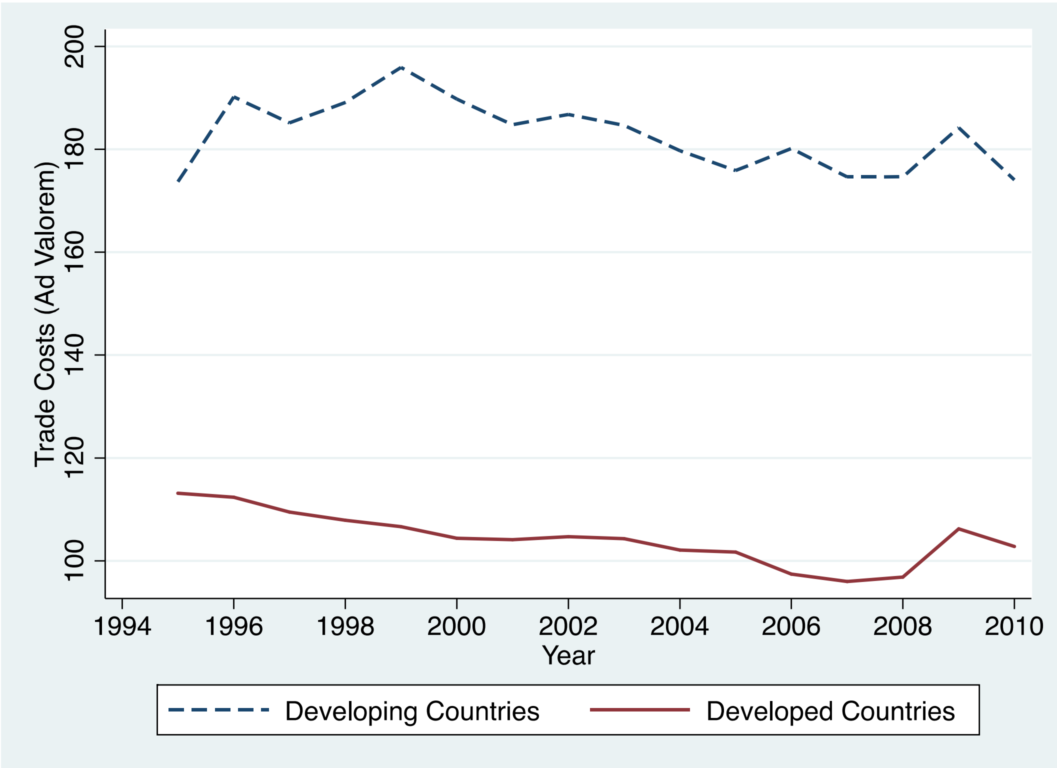
One area which has received much less attention is the impact of trade costs on the composition of trade. Recent work by Milner and McGowan (2013) finds that trade costs influence the manufacturing export mix of a set of mainly industrial countries; industries located in lower trade cost countries capturing relatively higher shares in the export of manufactured goods, the production of which is more sensitive to trade costs. Given that trade costs are considerably higher in general in developing countries (see Figure 1) and variable across and within developing country regions[1] and that the compostion of exports may well have important implications for the growth and development prospects of developing countries[2], it is important to understand if and how trade costs differences affect the export composition of dveloping countries. The present study extends therefore the earlier work to developing countries, but it also seeks to extend on the earlier empirical analysis. First, it uses

[1] Arvis et al. (2013) found that trade costs are much higher in developing countries and they are falling slowly than that in developed economies.

[2] Hausmann et al. (2007) for instance find that export product mix is one of the determinants of income level and subsequent economic development, with the countries specialising more in high-income country products tending to grow more rapidly.

a common trading partner approach (USA, with replacement by Germany for robustness checking), allowing the more precise measurement of trade costs and easier control for unobserved effects. Second, the concentration on a specific bilateral trade flow makes it necessary to deal systematically with the zero trade flows problem (a problem that was not evident for the multilateral trade of the industrial countries). Third, the study confronts the potential problem of the heterogeneity of trade costs.

Figure 1: Trade Costs in Developed and Developing Countries (% Ad Valorem Equivalent)



Source: Bilateral trade cost dataset (World Bank)

The theoretical literature captures a range of aspects of the possible role or influence of trade costs. Dornbusch, Fisher and Samuelson (1977) explore the influence of trade costs on the tradeability of goods. Similary, Markusen and Venables (2007) incorporate trade costs into an endowments model of trade, but allow trade costs only to vary across countries (trade costs being the same for goods to/from a particular country and a particular country having the same trade costs with all its trading partners). This specification allows for a clearly defined 'world price' for each good (X_i). They develop a model of three goods (produced under constant returns and competitive conditions), using two factors (capital, *K*, and labour, *L*). With zero (country) trade costs the pattern of production across countries would be indeterminate, though with full employment we can make predictions about the overall or average factor content of trade. The addition of trade costs (here country-specific trade costs, *t*) makes the commodity structure of production determinate. A key message from their modelling is that lower trade cost countries are characterised by partial or complete specialisation in production, while higher trade cost countries tend to become less specialised in production. Low trade cost countries trade all goods, while the incidence of autarky or non-tradability increases with trade costs. Indeed, reduced tradability starts to be a feature of increasing trade costs for countries close to the world average endowments. Further, at higher trade costs more extreme endowments are required to maintain a country as an exporter of the good intensively using the country's abundant factor.

In the above formulation national trade costs are a modifier of comparative advantage and the pattern of trade, through their influence on the tradability of goods. In order to capture national trade costs role as a source of comparative advantage we need to represent national trade costs as analogous to a traditional endowment. In a strand of the literature this has been done in effect by representing national trade costs as the country-specific, fixed cost or additional investment associated with the trade impediments that agents in that country need to overcome in order to transact internationally. This is an extension of Levchenko's representation (Levchenko, 2007) of national institutional differences which induce differences in international transaction impediments across countries.[3] In a world of trade cost barriers or frictions, a fraction (*f*) of the investment of factor resources (capital and/or labour) required to produce certain units of a tradable good become specific to the particular activity. In a frictionless world (f=0) agents do not need to invest specific resources (ex ante) to acquire information about how the infrastructure and institutional characteristics of the economy affect their ability to recoup their investment. Where f>0 it is harder to induce resources to enter sectors in general (harder relative to countries where trade costs are lower), and the more so in sectors that are more transactions-intensive; the ex post returns to factors being driven down relative to the frictionless case.

Analogous to Levchenko (2007), we can view trade in a two country or bloc (L – low trade cost country and H – high trade cost country) case as involving differences in national trade costs such that $f^H > f^L$; a lower fraction of factors being specific to transactions-sensitive activities in L than H.[4] In this set-up national trade cost differences act much like a normal endowment difference source of comparative advantage in an H-O framework; the low trade cost country being able to produce transaction-intensive goods, ceteris paribus, at relatively lower cost than the high trade costs country.

The core hypothesis of this study is that higher trade costs reduce the industry share of developing countries' exports, and do so more the more trade cost-intensive or sensitive is the production of goods in that industry. To test this proposition, we construct a panel dataset of trade costs and the export share for 113 developing country trading partners of the United States for the period 1995 to 2010, and investigate the relationship between these variables in a regression framework. Trade costs differ across countries and overtime, and trade cost

[3] Nunn (2007) also develops an analytical framework in which inter-country differences in institutional quality are a source of comparative advantage.
[4] Other possible sources of difference across countries, such as technological differences, are ruled out.

intensity (the proportion of imported inputs used in manufactured goods) also varies across industries as well as overtime. We exploit these variations in trade costs and trade cost sensitivity in the identification strategy. In the robustness checks we restrict the variation to variation in trade costs only, use different estimation approaches and empirical settings and include controls for factor endowments and institutional quality variables. The empirical analysis conducted by the study shows that the costs of conducting international trade are relatively high in developing economies and that these costs vary enormously across developing countries. Relatively high (low) costs are a source of comparative disadvantage (advantage), which reduces (increases) the manufacturing export shares of high (low) trade cost developing countries and reduces (increases) export shares most for industries that are most trade cost sensitive.

The remainder of the paper is organised as follows. Section 2 describes the trade cost indicator used in this work and summarises the pattern of trade costs across the developing countries. Section 3 explains the estimation methodology and introduces the dataset. Section 4 reports on the baseline estimation results. Section 5 reports on the robustness analysis, in particular endogeneity and specification issues. Finally, the conclusions and implications of the study are set out in section 6.

2. Measurement of Trade Costs

In the absence of directly measured data on trade costs, most studies compute them indirectly. The burgeoning literature on trade policy analysis uses distance between trading partners as a proxy for trade costs. Some scholars compute the CIF to FOB ratios, while others use the gravity model to isolate the effect of some specific component of trade cost. This study, however, uses a comprehensive trade cost indicator developed by the World Bank using an inverse gravity approach.

The inverse gravity model measures trade costs in terms of the trade depressing effect of national borders relative to domestic trade. It is based on the theoretical foundation developed in Novy (2013) and subsequently formalised in Arvis et al. (2013). This trade cost indicator measures the trade restricting effect of national borders from the gap between observed trade and actual trade potential. The trade potential is estimated from the patterns of production, consumption, and trade in origin and destination countries and trade costs are computed using the following approach:

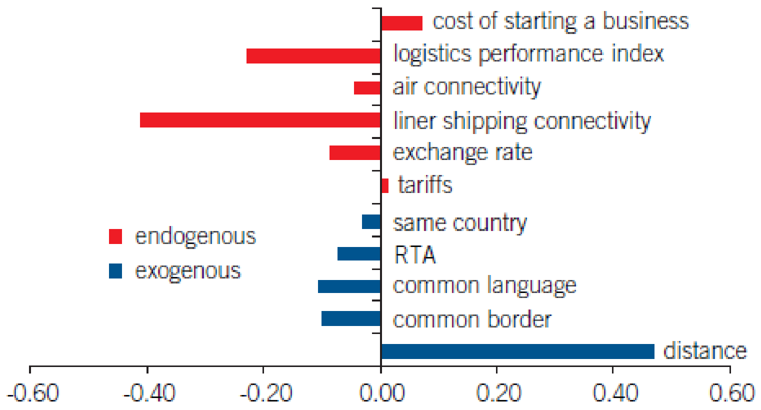
$$\tau_{ijkt} \equiv \left(\frac{t_{ijkt} t_{jikt}}{t_{iikt} t_{jjkt}} \right)^{\frac{1}{2}} - 1 = \left(\frac{x_{iikt} x_{jjkt}}{x_{ijkt} x_{jikt}} \right)^{\frac{1}{2(\sigma_k - 1)}}$$

…………(1)

where 'Ʈ' is trade cost between country 'i' and country 'j' at time 't', in sector 'k' (agriculture or manufacturing); 'x' is exports, 'σ' is elasticity of substitution, subscripts 'ii'/'jj' indicate intra-country trade and subscripts 'ij'/'ji' designate inter-country trade. Bilateral trade costs computed using this procedure are ad valorem (tariff) equivalents, and are symmetric in nature.

This methodology has micro-theoretical foundations and is devoid of omitted variables bias. Moreover, it is theoretically consistent as it includes all components of trade costs discussed in Anderson and van Wincoop (2004). Following this approach, the World Bank and UN-ESCAP have generated a bilateral trade cost dataset of 178 countries. This dataset provides trade costs in two broad economic categories, agriculture and manufacturing, and is available for the period 1995 to 2012. This relatively long time span allows the estimation of empirical trade models in a panel setting. Another feature of this indicator is its sensitivity to key components of trade costs: policy barriers (endogenous), and non-policy or natural barriers (exogenous). The former includes tariffs and non-tariff measures, shipping line connectivity, and infrastructure performance, whereas the latter comprises geographical or natural factors, such as distance and the lack of common language, etc.

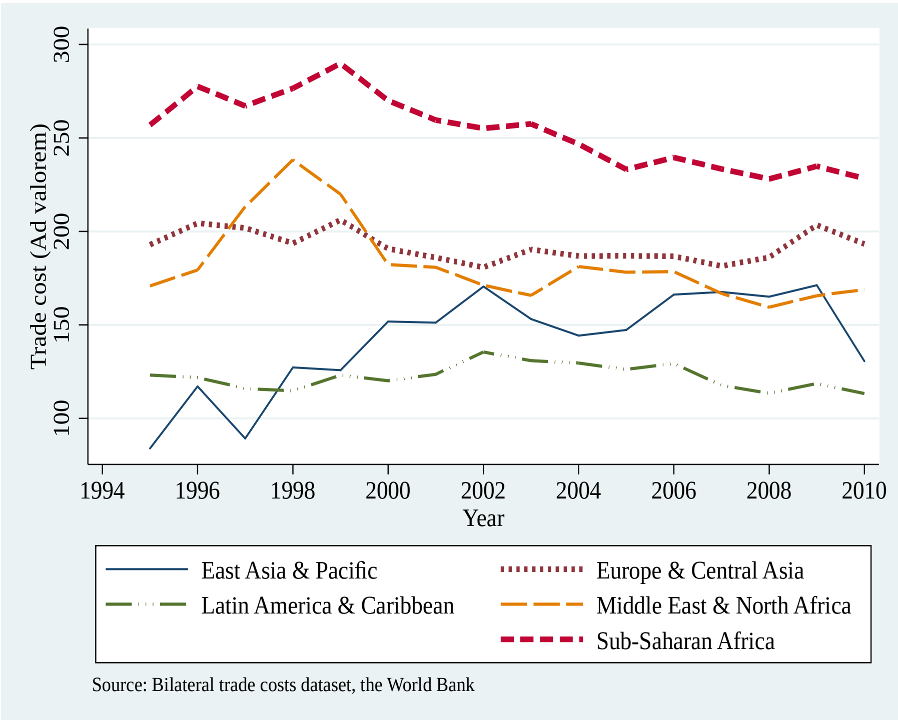
Figure 2: Sensitivity of Trade Costs to Various Components



Source: Economic Premise 104 (World Bank)

Figure 2 plots standardized regression coefficients depicting correlation of the indicator with various elements of trade costs. It shows the relative importance of various components and suggests that policy interventions can even out the effect of some of the exogenous and endogenous barriers. For example, the magnitude of the coefficient on distance (the most important exogenous barrier) is similar to that on the shipping line connectivity (an endogenous barrier), but these coefficients bear opposite signs. This means that a good logistic network can circumvent the trade depressing effect of geographical remoteness between trading partners. In Figure 3 we show the average trade costs of countries in different regions (controlling in part for the endogenous influence on the selection of trade partners by measuring the average costs of each country's trade with its largest ten trading partners). Note the marked variation in trade costs between regions, with Africa having experienced some decline in trade costs over the period but being the highest cost region throughout.

Figure 3: Evolution of Trade Costs by Region over Time



3. Identification Strategy, Data and Empirical Model

We exploit the variation in trade costs of developing countries (with a common trading partner) and the variation in trade cost sensitivity across industries in the identification strategy. The dataset of the World Bank provides trade cost indicators at a bilateral level, and to make estimation results comparable across countries and keep the sample size large we choose the United States as a common trading partner of all developing countries. Romer

(1994) illustrates that large economies import a wider variety of goods from a large number of foreign suppliers. Table 1 shows that 113 developing countries from various geographical regions export to the United States. Moreover, the substantial variation in trade costs across regions and over time provides a good setting for the econometric analysis in a panel structure.

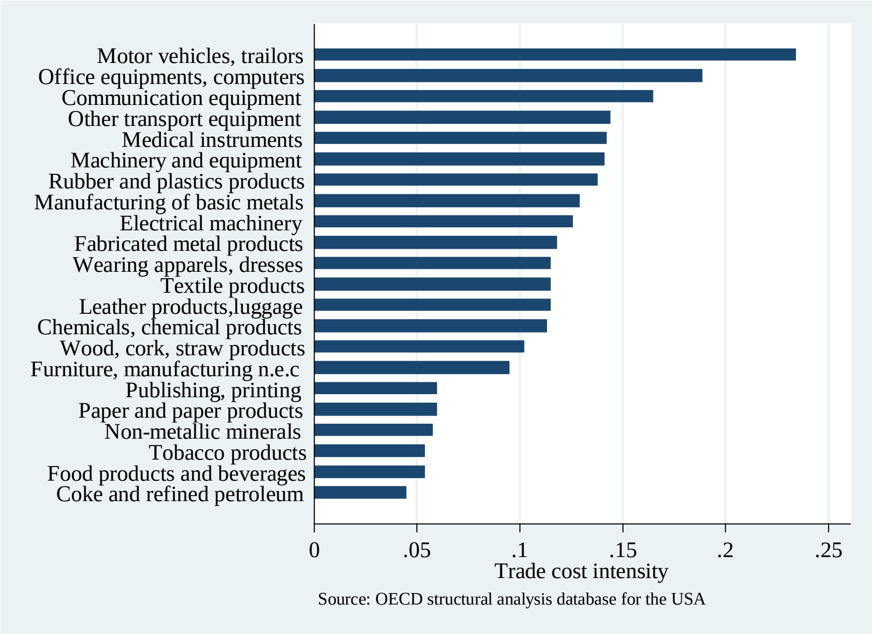
Table 1: Trade Costs and Trading Partners of the USA

Geographical Region	Trading Partners of the USA	Trade Costs (Ad Valorem)				
		Min	Mean	Max	SD	Obs
East Asia and Pacific	24	32	148	425	92	3,696
Europe and Central Asia	17	61	192	356	51	4,642
Latin America and Caribbean	25	34	123	384	46	7,546
Middle East and North Africa	12	98	182	625	72	2,640
South Asia	7	92	190	358	83	1,892
Sub-Saharan Africa	44	94	253	821	102	10,450

Source: Authors' Working using Comtrade and Trade Costs Dataset

The second pillar of the identification strategy, sensitivity to trade costs, varies across industries as well as over time. The proportion of imported inputs used in manufactured goods for each industrial category is used to proxy this variable. Figure 4 graphs this variable across industries for the year 2008. As this diagram illustrates, the motor vehicles, office equipment, and the telecommunication sectors are relatively high trade cost intense, whereas the food and mineral sectors are low trade cost intense.

Figure 4: Variation in Trade Cost Sensitivity across Industries (2008)



Data

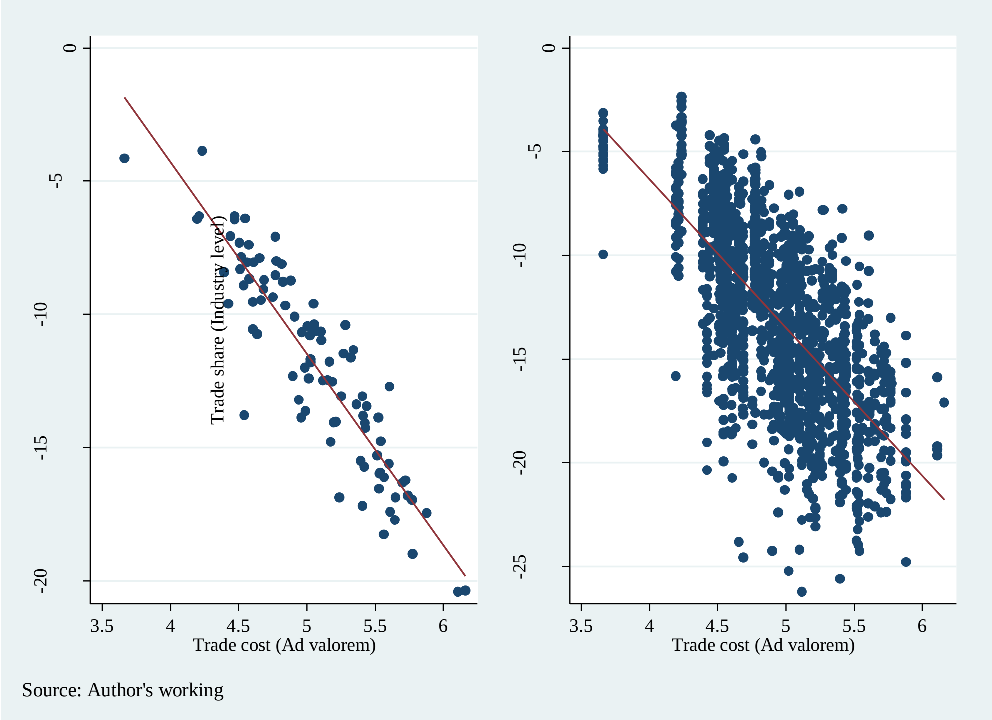
In order to investigate the relationship between trade costs and trade composition we construct a panel dataset of the low and middle income trading partners of United States by extracting information from multiple sources. Industry level trade flows are retrieved from the UN Comtrade database, and the bilateral trade cost indicators are extracted from the trade costs dataset of the World Bank. The list of developing countries and industries included in this work are contained in the annexes (Table A.1 and A.2). Trade cost intensity data is taken from the OECD structural database for the USA. We take the USA as the reference or frontier technology for this purpose, because it is a relatively open and undistorted economy and because it is exogenous to the set of developing countries whose export structure we are investigating. The information on policy-related variables and factor endowments (used in the robustness tests) has been collected from multiple sources described in annex 1 and 2. The summary statistics of all the variables are contained in Table A.3.

The dataset contains 32,868 observations: 22 industrial categories (ISIC Code 15-36) of 113 countries for the period 1995 to 2010. It has 22,253 (68%) actual trade flows and contains 10,615 (32%) missing observations. Since we observe the trade flows at a (relatively high level of aggregation) two-digit level of ISIC Rev 3 classification, it cannot be assumed that these observations are missing at random. Literature suggests that these missing observations may be structural zeroes because firms may not export to high trade cost destinations (UNCTAD, 2011, p-112). Given that the proportion of missing values is quite large and their distribution is not random, the econometric analysis must recognise this in an appropriate manner.

Figure 5 (left panel) graphs the industrial exports of the trading partners of the USA against bilateral trade costs (on a logarithmic scale) for the year 2008. This figure illustrates the negative correlation between these variables. It is evident that the trade share of these countries in the US market declines sharply with the rise in trade costs, and for relatively higher costs, these trade flows are almost negligible. In right panel of the chart, we disaggregate the country level trade share into industrial categories as per the ISIC (Rev 3.) classification system. The scatter plot of the disaggregated data (right panel) points to the heterogeneity of the impact of trade costs across industries.

Figure 5: Trade Costs and Manufacturing Trade Share (2008)

On a Logarithmic Scale



Regression Framework

To investigate our hypothesis in a regression framework, the baseline specification is as follows:

$$trade\_share_{ict} = \beta_0 + \beta_1 (trade\_cost\_intensity)_{it} * (trade\_cost)_{ct} + \gamma_c + \delta_i + \Upsilon_t + \varepsilon_{ict} \quad ...(2)$$

The dependent variable is the exports of an industry 'i' of a country 'c' to the USA, at time 't' and normalized by the world trade flow in the same industry. This transformation is similar to one adopted in Eaton and Kortum (2002) and Levchenko (2007) in order to make results comparable across countries. The description of the trade cost indicator, trade cost intensity and their data sources is described above. Fixed effects for trading partners, industries and years are also included in the estimated model.

In this framework, the effect of trade costs on trade composition is determined by the variation in trade costs across countries and time, and the variation in trade cost sensitivity across industries as well as across time. This is a sort of difference-in-difference approach, although both trade costs and trade cost intensity are continuous variables. Angrist and Pischke (2009, p-201) argue that the idea of difference-in-difference is more general and it implicitly assumes treatment and control groups. We estimate this baseline equation in a panel structure. The coefficient on the interaction terms (trade costs x trade cost intensity) is the main focus of the investigation. A negative and statistically significant value of the coefficient would mean that high trade cost intensive industries located in high trade cost countries gain a relatively smaller share in exports to the USA of manufactured industry compared to their counterparts located in low trade cost countries. This functional form is quite standard in international trade literature and similar to one employed in Rajan and Zingales (1998), Nunn (2007) and Levchenko (2007).

To mitigate any problem of serial correlation, robust standard errors are clustered at the country level since trade flows between countries tend to be highly persistent due to the presence of sunk costs. Moreover, to obviate the possibility of a unit root due to the time series nature of the panel data, the augmented Dickey-Fuller (ADF) test as well as Maddala and Wu unit root tests have been conducted. Both of which confirm the stationarity of the data series at 1% significance level. Finally, we restrict the scope of the work to manufacturing sectors only. We exclude the agriculture sector deliberately because trade in this sector is subject to a plethora of market access barriers, such as tariff-rate quotas, sanitary and phyto-sanitary and quality control measures.

4. Base Estimations

Base Results

Table 2 summarizes the baseline results of the estimated equation using the OLS estimator. The column (1) presents the results for the pooled sample, while the columns (2) through (4) contain those after incorporating controls for unobservables by including the country, industry, and time dummies. All estimations are in logs.

Table 2: Base Results (OLS Estimates)

Dependent Variable: Export Share

	(1)	(2)	(3)	(4)
Trade costs x cost intensity	-3.767***	-0.833***	-1.244***	-1.352***
	(0.036)	(0.045)	(0.105)	(0.112)
Country dummies		Y	Y	Y
Industry dummies			Y	Y
Time dummies				Y
R^2	0.29	0.63	0.68	0.69
N	22,253	22,253	22,253	22,253

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients on the fixed effects are not reported. These coefficients were obtained using the OLS estimator using Stata 13.0 SE.

These results show that the trade share of industries varies inversely with our regressor of interest - the interaction of trade costs and trade cost intensity, and that this effect is statistically significant at a 1% significance level. Since the OLS regressions are in logs, the coefficient on the interaction term can be interpreted as an elasticity. Column (1) indicates that the elasticity of exports to the interaction variable is (-) 3.76. These results hold after including fixed effects, as shown in column (2) through column (4) but their magnitude varies indicating the presence of omitted variables. The explanatory power of the model (R^2)

improves when fixed effects are included for trading partners but it changes little once industry and time dummies are added in the regressions. This suggests that the differences in country characteristics account for a large proportion of the variation in export shares.

The OLS estimator truncates the sample and drops missing trade flows, which may create a selection problem because the descriptive analysis of data in the previous section suggests the presence of a substantial fraction of zeroes. Therefore, the same analysis has been performed again by using the Pseudo Poisson Maximum Likelihood (PPML) estimator (Table 3). The PPML estimator overcomes the selection problem that may arise in this kind of analysis because all industries of these countries may not export to the common trading partner in each year. Since this entry and exit is not random, it may create a selection bias. As the PPML estimation explicitly accounts for the missing values, it can circumvent this problem. Moreover, the PPML estimator accounts for heteroskedasticity that is prevalent in the trade flow data, and it is also robust to arbitrary patterns of serial correlation (Silva and Tenreyro, 2006).

The results of the PPML estimations are reported in Table 3. They show that the change in estimation technique does not affect the baseline results: the coefficient on the interaction term remains negative and statistically significant in column (1) through column (4). However, it increases the sample size from 22,253 to 32,868 which may improve the precision of the coefficient estimates.

Table 3: Base Results (PPML Estimates)

Dependent Variable: Export Share

	(1)	(2)	(3)	(4)
Trade costs x cost intensity	-1.664***	-0.418***	-1.188***	-1.348***
	(0.034)	(0.071)	(0.156)	(0.167)
Country dummies		Y	Y	Y
Industry dummies			Y	Y
Time dummies				Y
R^2_p	0.09	0.29	0.32	0.32
N	32,868	32,868	32,868	32,868

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients on the fixed effects are not reported. These coefficients were obtained using the PPML estimator using Stata 13.0 SE

A comparison the OLS and the PPML estimations shows that the coefficients in the pooled sample are upwardly biased in the former technique, but they tend to converge when fixed effects are included in the estimations. PPML estimations are used as the preferred estimator hereafter.

The base results in tables 2 and 3 allow trade cost sensitivity to vary over time. We consider also the effect of making the 'technology' (i.e. industry level, intermediate import-intensity) invariant over time as well as country, by taking the start of sample value (1995) and interacting this measure with trade costs (which vary across country and time). This allows use to check that our key finding on the interaction term in Tables 2 and 3 is not being driven only by the change over time in trade cost sensitivity or intensity. The results for both the OLS and PPLM estimates in this case are presented in Appendix Table A.4. For both estimation methods the interaction term remains negative and highly significant, supporting the core hypothesis.

Figure 3 above shows that trade costs vary widely across regions; note the marked variation in trade costs between regions, with Africa having experienced some decline in trade costs over the period but being the highest cost region throughout. Trade cost intensity also varies markedly across industries as depicted in Figure 4, and it might be argued that the above results are being driven by some specific regions or industrial sectors. Therefore, we explore the heterogeneity of the impact of trade costs for various sub-samples. Table 4 presents the estimation results for the regional sub-samples used.

Table 4: Heterogeneity Across Developing Regions

Dependent Variable: Export Share

	(1)	(2)
Trade costs x cost intensity x		
Low trade cost	-1.334***	
	(0.170)	
High trade cost	-1.703***	
	(0.162)	
East Asia & Pacific		-1.127***
		(0.152)
Europe & Central Asia		-1.679***
		(0.184)
Latin America & Caribbean		-1.740***
		(0.240)
South Asia		-1.436***
		(0.173)
Sub-Saharan Africa		-2.165***
		(0.272)
R^2_p	0.32	0.32
N	32,868	32,868

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients on the fixed effects are not reported. These coefficients were obtained using the PPML estimator using Stata IC 13.0.

Model (1) splits the sample in two parts on the basis of mean trade costs i.e. 191% ad valorem. In the first sub-set these costs range from 32% to 190% with a mean of 129%, while in the second one they range from 190% to 861% with a mean of 278%. These sub-samples have 13,530 and 19,338 observations respectively. The estimates presented in the table directly measure the effect of trade costs for each group. They show that the core hypothesis holds for both low and high trade cost developing countries. They also show, however, that within the developing countries the effect of trade costs on export shares is greater for the higher trade cost countries; meaning that these high trade cost countries would experience a greater increase in export shares from any reduction in trade costs.

In order to further explore this heterogeneity across countries, we estimate the same equation by disaggregating the sample into regional sub-groups. The results in column (2) indicate that effect of trade costs is negative and significant for each developing region. The magnitude of the effect, however, varies slightly, with the largest negative coefficient being found for the high cost Africa region.

The base results are further decomposed to explore the heterogeneity of response across manufacturing sectors. Table 5 present the results for a few, selected industrial sectors with the complete results being given in Table A.5. We find a negatively significant effect on the key interaction term for all sectors; the magnitude of the coefficient however varies, as might be expected, in line with the degree of sensitivity of industries to trade cost. It appears that export shares of the low trade cost intensive industries (such as fruit and vegetables, wood, furniture, minerals, textiles, leather) are relatively less affected by national trade cost differences or changes compared to high trade cost intensive industries (such as electronics, telecommunication, and automobiles).

Table 5: Heterogeneity Across Industries

Dependent Variable: Export Share

	(1)
Office equipment, computers	-1.926***
	(0.215)
Chemicals, chemical products	-1.305***
	(0.175)
Communication equipment	-2.804***
	(0.198)
Electrical machinery	-3.618***
	(0.209)
Food products and beverages	-1.018***
	(0.200)
Furniture, manufacturing, etc.	-1.015***
	(0.195)
Machinery and equipment	-2.175***
	(0.198)
Manufacturing of basic metals	-0.571**
	(0.237)
Medical instruments	-2.398***
	(0.182)
Non-metallic minerals	-1.738***
	(0.193)
Textile products	-0.629***
	(0.202)
Wood, cork, straw products	-0.665***
	(0.198)
R^2_p	0.33
N	32, 868

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients on the fixed effects (countries, industries and time) are not reported. These coefficients were obtained using the PPML estimator using Stata IC 13.0.

5. Robustness Analysis

Alternative Estimator and Empirical Setting

The first-difference estimator has been suggested in many studies as it relies on weaker exogeneity assumptions and is more efficient when the error term is serially correlated. Baier and Bergstrand (2007), for example, argue for employing it to overcome the problem of serial correlation. These authors recommend including fixed effects for countries and time in addition to first-differencing the data, or applying this technique at least to test the robustness of results. Magee (2008) adopts a similar approach. The estimation results contained in column (1) of Table 6 indicate that the coefficient on the interaction term remains negative as in the earlier estimations. This variation in the estimation technique does not influence the sign or significance level on the coefficient. In column (2), we replicate the same analysis by selecting Germany as a reference country. This alternative setting obviates the possibility that preferential access of some developing countries to the US market may influence the estimations. As the results show, this change in the empirical setting does not affect the sign and significance level of the coefficient of interest.

Both of these robustness tests - by employing an alternative estimation approach and using a different common trading partner - further validate the initial findings: the coefficients on the interaction of trade costs and trade cost intensity remain negative and statistically significant, which shows trade costs are an important determinant of export composition.

Table 6: First-differenced Estimator and Germany as the Common Trade Partner

Dependent Variable: Export Share

Estimation Approach	First-differenced Estimations (Exports to USA)	Germany as a Common Trading Partner
	(1)	(2)
Trade costs x cost intensity	-1.692***	-1.317***
	(0.156)	(0.211)
Country dummies	Y	Y
Industry dummies	Y	Y
Time dummies	Y	Y
Observations	18,468	28,222

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.The coefficients on the fixed effects are not reported. These coefficients in column (1) were obtained using the OLS estimator using Stata IC 13.0, because the PPML estimator does not work on first differenced-data. The coefficients in column(2), however, were obtained using the PPML estimator. As the PPML estimator accounts for the presence of zeros also, the sample size in the column (2) is larger.

Instrumental Variable Estimations

It might be argued that the persistency of trade flows and the endogeneity of trade costs may drive these results as these variables are not independent of the policy environment. Literature suggests that the sunk costs of entry into export markets generate persistency and hysteresis in trade flows, which may create the problem of inconsistency in estimations. To overcome this problem, we use lagged values of the dependent variable as instruments. We use the “system GMM” estimator on the dynamic panel data model in column (1) of Table 7. This estimator employs lagged first-differenced values of the dependent variable as instruments for the level equation, and lagged values of the dependent variables at levels as instruments for the differenced equation. This technique is recommended in the literature when the dependent variable is highly persistent. As the results show, this alternative approach yields results similar to those of the baseline estimations.

Many studies use geographical distance between trading partners as an indicator of trade costs. Some studies, however, argue that the variation in trade costs between countries may not necessarily depend upon the geographical distance. In column (2) and (3), we instrument trade costs on the Liner Shipping Connectivity Index (LSCI) and on distance to the market of the common trading partner (USA).

Table 7: Robustness: Using Various Instrumental Variables

Dependent Variable: Export Share

Instrumental Variables:	Lagged Values of Dependent Variable	Distance to Market of the USA	Shipping Connectivity Index
	(1)	(2)	(3)
Trade cost x cost intensity	-2.991***	-2.461***	-2.47***
	(0.203)	(0.164)	(0.261)
County dummies	Y	Y	Y
Industry dummies	Y	Y	Y
Year dummies	Y	Y	
Observations	18,468	30,052	9,680
Number of countries	110	110	110

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients on fixed effects (countries, industries, and time) are not reported. These coefficients were obtained using the OLS estimator using Stata IC 13.0

Using distance as an instruments offers an advantage as it is completely exogenous and has a positive correlation of 0.50 with bilateral trade costs. Column (2) presents the estimation results by instrumenting trade costs on geographical distance. As the results illustrate, the coefficient of the interaction variable has the expected sign and is statistically significant.

Comparing the first two columns indicates that the magnitude of the coefficient is very similar in both models, however, the sample size increases in model (2) because the information about distance is available for a larger set of countries.

In Column (3), we use the Liner Shipping Connectivity Index (LSCI) developed by the United Nations Conference on Trade and Development (UNCTAD) as an instrumental variable for trade costs (Fugazza et al., 2013). This instrument is a 'relevant' one because it is strongly correlated with the trade cost indicator, with the correlation coefficient being '-0.625'. Moreover, it is 'exogenous' to trade flows, for shipping line connectivity (frequency, time, competition) impacts on trade flows only through transportation costs. This is because the improvement in shipping line connectivity of a given country may affect transportation charges, which in turn might influence its trade share. Further, the market structure of shipping lines is oligopolistic in nature because a few shipping lines handle the major quantum of world trade. The three biggest shipping lines – the CMA of France, Maersk of Denmark and Mediterranean Shipping line (MSC) of Switzerland – transport more than half of the container trade between Asia and Europe. This group, namely P3, is a large international grouping, with these companies having entered into a space sharing arrangement (The Economist, 21st June, 2014, p-74). Due to this unique market structure and high correlation of the LSCI with the trade costs indicator, this instrument satisfies both the conditions of exogeneity and relevance.

To operationalise this instrument, we use a two stage least square (2SLS) estimation. In the first stage, trade costs are regressed on the connectivity index and the predicted values are generated. In the second stage, we inter-act the predicted values with trade costs intensity to derive the explanatory variable, which is used in the regression framework. Since this index has a negative correlation with trade costs it would be expected to yield a positive coefficient for the interaction variable, because higher values of the LSCI mean lower trade costs. However, for the sake of clarity and consistency, we use the inverse of this index in the estimations. This transformation generates negative and statistically significant results (Column 3, Table 7). This IV approach, therefore, further validates the earlier findings, though the magnitude of effect is larger than that found in the earlier estimations. The instrumental variable estimations generate similar results even if we include both the instruments in the regression and an over identification test confirms that both the instruments are exogenous. Overall, these estimations further substantiate the earlier findings.

Including Controls for Institutional Variables and Factor Endowments

This sub-section investigates whether the trade costs indicator is picking up the effect of some other policy-related variables. In order to test this, we add in to the regressions the interaction of trade cost intensity with policy-related variables, in addition to its interaction with the trade costs indicator. A brief description of these variables and their data sources is contained in Annex 1.

Table 8: Country, Industry, and Policy Interactions

Dependent Variable: Export Share

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Trade cost intensity x							
Trade costs	-0.796***	-1.260***	-1.248***	-1.394***	-2.435***	-1.365**	-2.147***
	(0.198)	(0.190)	(0.239)	(0.197)	(0.425)	(0.667)	(0.349)
Legal institutions	1.258***						
	(0.279)						
Cost of importing		-0.001					
		(0.001)					
Cost of exporting			-0.002**				
			(0.001)				
Freedom to trade				0.134***			
				(0.024)			
Tariffs					-0.029		
					(0.099)		
Non-tariffs						0.018	
						(0.059)	
GDP per capita							1.642***
							(0.340)
Fixed effects	Y	Y	Y	Y	Y	Y	Y
Observations	12,683	10,913	10,913	19,160	25,718	25,718	28,097

Standard errors in parentheses* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients on fixed effects (countries, industries, and time) are not reported. These coefficients were obtained using the PPML estimator using Stata IC 13.0. The first three columns have relatively fewer observations because the data of these control variables is not available for all countries for the whole period.

As the Table 8 shows, the coefficient on the regressor of interest is negative and significant in all of the extended specifications. The effect of the additional covariates is generally as expected, without confounding the effect of trade costs. For instance, the coefficient on legal institutions, freedom to trade, and GDP per capita is positive and significant with increases in these indicators promoting export shares, whereas increase in the cost of exporting reduces export shares. The impact of tariff and non-tariff barriers is insignificant. In the case of tariffs this could be because they have been substantially lowered in the manufacturing sector. Non-tariff barriers may also be concentrated on specific product types. For both import policy measures, however, the effect may be expected to be limited on the imposing country's export composition.

Table 9: Interaction between Country and Industry Characteristics

Dependent Variable: Trade Share

	(1)	(2)	(3)	(4)
Trade cost x Trade cost intensity	-0.888***	-1.447***	-1.039***	-0.742***
	(0.220)	(0.233)	(0.177)	(0.144)
Physical capital x Physical capital intensity	0.758***			
	(0.184)			
Human capital x Human capital intensity		6.135***		
		(1.724)		
Rule of law x Contract intensity			0.086	
			(0.080)	
Financial development x Ext. financial dependence				0.346***
				(0.053)
Fixed effects	Y	Y	Y	Y
Observations	26,330	21,949	19,160	14,702

Standard errors in parentheses* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients on fixed effects (countries, industries, and time) are not reported. These coefficients were obtained using the PPML estimator using Stata IC 13.0. Columns (3) and (4) have fewer observations since the data for the rule of law and financial development is not available for all developing countries.

Romalis (2004), and many other studies, argue that factor endowments are a major source of comparative advantage and an important driver of trade flows. Therefore, we augment the model with traditional factor endowments i.e. physical capital, human capital, rule of law, and financial development in order to examine their effect on the baseline results. Table 9 reports the estimated coefficients adjusted for the separate effect of each factor endowment, while Table A.6 contains those for the incremental effect of these variables. The inclusion of or control for these additional covariates further validates the base results. Although the magnitude of the coefficient on the trade costs term is slightly lower once the policy variables and factor endowments are accounted for, it still has the expected sign and remains statistically significant at a 1% significance level.

6. Conclusions and Implications

Using the recently released trade costs dataset of the Word Bank for the period 1995 to 2010, we investigate the influence of trade costs on the export composition of developing countries. We find that despite substantial reduction, trade costs in developing countries are still very high and vary widely between developing countries (from 30% ad valorem to 600% ad valorem). Most importantly, we argue that these high trade costs influence the comparative advantage of these economies, in that the industries located in high trade cost countries gain a relatively smaller share in the exports of manufactured goods compared to their counterparts situated in low trade cost countries, and that this effect is greater the more trade cost- intense or sensitive is the production of the export. This statistically significant negative relationship holds for sub-samples of countries and industries, different estimation techniques (OLS, PPML, and first-difference estimator), alternative common trading partners, and even to the inclusion of other sources of comparative advantage, such as factor endowments and institutional quality variables. The study generates, therefore, econometric evidence about how international trade costs systematically fashion the composition, as well as volume, of exports of the developing countries.

The evidence that high trade costs reduce industrial exports of the developing world in general and specific types of industrial exports in particular has development policy implications. First, we find heterogeneity in the effect of trade costs across regions and show that the trade restricting effect is relatively large in sub-Saharan Africa (SSA). This suggests that reducing these costs through trade facilitation measures or addressing policy and non-policy barriers can provide a boost to industrial exports from all developing regions, but that this response will be higher for SSA.

Second, the industrial exports of some sectors (such as autos, telecommunication) appear to be relatively more sensitive to trade costs. This is because these industries are more dependent on imported intermediate inputs, and higher trade costs push up the costs of these inputs as well as increasing the direct costs of exporting. This means that reducing trade costs will generate a higher export response in these industries compared to less trade cost sensitive ones.

This heterogeneous response across industries has implications for economic growth in developing countries, because the mix of goods a country exports is likely to have important implications for its development. Literature shows that all goods are not similar in terms of their economic significance; specialising in some products brings higher growth opportunities than specialising in others for both demand and supply-side reasons. Since the response to trade cost reduction by more trade cost sensitive sectors is shown by the current work to be relatively large, the lowering of developing countries' trade costs can be expected to disproportionately promote exports of more trade cost sensitive and technologically sophisticated goods by these countries. This offers an opportunity for increasing the growth-enhancing effects of export expansion.

References

ANDERSON, J. E. and VAN WINCOOP, E. 2004. Trade costs. National Bureau of Economic Research.

ANGRIST, D.J and PASCHKE.J.S. 2008. Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press.

ARVIS, J.-F., DUVAL, Y., SHEPHERD, B. and UTOKTHAM, C. 2013. Trade Costs in the Developing World: 1995 – 2010. Policy Research Working Paper 6309, The World Bank.

BAIER, L. S. and BERGSTRAND, H.J. 2007. Do free trade agreements actually increase members' international trade? *Journal of International Economics,* 71, 72–95.

DORNBUSCH, R., FISCHER, S. and SAMUELSON, P. A. 1977. Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods. *American Economic Review***,** 67, 823-839.

EATON, J. and KORTUM, S. 2002. Technology, geography, and trade. *Econometrica*, 70**,** 1741-1779.

FEENSTRA, R. C., INKLAAR, R. and TIMMER, M. 2013. The Next Generation of the Penn World Table. National Bureau of Economic Research.

FUGAZZA, M., HOFFMANN, J., and RAZAFINOMBANA, R. (2013). Building a Dataset for Bilateral Maritime Connectivity

HAUSMANN, R., HWANG, J. and RODRIK, D. 2007. What you export matters. *Journal of Economic Growth*, 12**,** 1-25.

HELPMAN, E., MELITZ, M. and RUBINSTEIN, Y. 2008. Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, 123**,** 441-487.

LEVCHENKO, A. A. 2007. Institutional quality and international trade. *Review of Economic Studies*, 74**,** 791-819.

LIMAO, N. and VENABLES, A. J. 2001. Infrastructure, geographical disadvantage, transport costs, and trade. *World Bank Economic Review*, 15**,** 451-479.

MAGEE, C. S. 2008. New measures of trade creation and trade diversion. *Journal of International Economics*, 75**,** 349-362.

MARKUSEN, J. and VENABLES, A. J. 2007. Interacting factor endowments and trade costs: A multi-country, multi-good approach to trade theory. *Journal of International Economics*, 73, 333-354.

MCCALLUM, J. 1995. National borders matter: Canada-US regional trade patterns. *American Economic Review***,** 615-623.

MILNER, C. and MCGOWAN, D. 2013. Trade costs and trade composition. *Economic Inquiry*, 51**,** 1886-1902.

NOVY, D. 2013. Gravity redux: measuring international trade costs with panel data. *Economic Inquiry*, 51**,** 101-121.

NUNN, N. 2007. Relationship-specificity, incomplete contracts, and the pattern of trade. *Quarterly Journal of Economics*, 122**,** 569-600.

OBSTFELD, M. and ROGOFF, K. 2001. The six major puzzles in international macroeconomics: is there a common cause? *NBER Macroeconomics Annual 2000, Volume 15.* MIT Press.

PSACHAROPOULOS, G.1994. Returns to investment in education: A global update. *World Development*, 22, 1325–1343.

RAJAN, R. G. and ZINGALES, L. 1998. Financial dependence and growth. *American Economic Review,* 88**,** 559-586.

ROMALIS, J. 2004. Factor proportions and the structure of commodity trade. *American Economic Review*, 94, 67-97.

ROMER, P. 1994. New goods, old theory, and the welfare costs of trade restrictions. *Journal of Development Economics,* 43**,** 5-38.

SILVA, J. S. and TENREYRO, S. 2006. The log of gravity, zero. *Review of Economics and Statistics,* 88**,** 641-658.

TREFLER, D. 1995. The case of the missing trade and other mysteries. *American Economic Review***,** 1029-1046.

Table A.1: List of Developing Countries Included in the Analysis

Afghanistan	Dominica	Lesotho	Senegal
Albania	Dominican Republic	Macedonia, FYR	Seychelles
Algeria	Ecuador	Madagascar	Sierra Leone
Andorra	Egypt, Arab Rep.	Malawi	South Africa
Argentina	El Salvador	Malaysia	Sri Lanka
Armenia	Eritrea	Maldives	St. Lucia
Azerbaijan	Ethiopia	Mali	Suriname
Bangladesh	Fiji	Mauritania	Swaziland
Belarus	Fm Sudan	Mauritius	Syrian Republic
Belize	Gabon	Mexico	Tanzania
Benin	Gambia, The	Moldova	Thailand
Bhutan	Georgia	Mongolia	Togo
Bolivia	Ghana	Montenegro	Tonga
Bosnia Herzegovina	Grenada	Morocco	Tunisia
Botswana	Guatemala	Mozambique	Turkey
Brazil	Guinea	Namibia	Turkmenistan
Bulgaria	Honduras	Nepal	Tuvalu
Burkina Faso	Hungary	Nicaragua	Uganda
Burundi	India	Niger	Ukraine
Cambodia	Indonesia	Nigeria	Vanuatu
Cameroon	Iran, Islamic Rep.	Pakistan	Venezuela
Cape Verde	Israel	Panama	Vietnam
C.African Republic	Jamaica	Papua New Guinea	Yemen
China	Jordan	Paraguay	Zambia
Colombia	Kazakhstan	Peru	Zimbabwe
Congo, Rep.	Kenya	Philippines	
Costa Rica	Kiribati	Romania	
Cote d'Ivoire	Kyrgyz Republic	Rwanda	
Cuba	Lebanon	Samoa	

Table A.2: ISIC Rev.3 Industry Classification

S.No.	Industry Code	Industry	Trade Cost Intensity
1	15	Coke, refined petroleum products and nuclear fuel	0.045
2	16	Food products and beverages	0.054
3	17	Tobacco products	0.054
4	18	Other non-metallic mineral products	0.058
5	19	Paper and paper products	0.060
6	20	Publishing, printing and recorded media	0.060
7	21	Furniture; manufacturing n.e.c.	0.095
8	22	Wood and of products of wood	0.102
9	23	Chemicals and chemical products	0.113
10	24	Textiles	0.115
11	25	Wearing apparel; dressing and dyeing of fur	0.115
12	26	Tanning and dressing of leather	0.115
13	27	Fabricated metal products, except machinery	0.118
14	28	Electrical machinery and apparatus n.e.c.	0.126
15	29	Basic metals	0.129
16	30	Rubber and plastics products	0.138
17	31	Machinery and equipment n.e.c.	0.141
18	32	Medical, precision and optical instruments, watches	0.142
19	33	Other transport equipment	0.144
20	34	Radio, television and communication equipment	0.165
21	35	Office, accounting and computing machinery	0.189
22	36	Motor vehicles, trailers and semi-trailers	0.234

Source: UN, ISIC Rev.3

Table A.3: Summary Statistics

Variable	Min	Mean	Sd	Max	Obs
Trade costs (Ad valorem)	32.49	191.09	95.17	821.39	32,868
Trade cost intensity	0.03	0.11	0.05	0.24	32,868
Cost of importing	317	1616	946	7709	13,706
Cost of exporting	295	1349	766	5491	13,706
Legal institutions index	0.00	5.13	2.32	10.00	15,752
Economic Freedom Index (EFI)	21.40	55.48	8.07	76.30	24,288
Protection of property rights	5.00	40.48	15.43	75.00	24,288
Freedom from corruption index	7.00	30.33	13.59	80.00	24,288
Freedom of business index	10.00	59.62	12.01	87.90	24,288
Freedom to trade index	0.00	60.90	15.49	89.10	24,288
Applied tariffs(simple average)	0.00	3.03	17.25	350.00	26,615
Linear shipping connectivity	0.00	14.98	13.09	88.14	10,934
Gross domestic product (PPP)	319	5370	4485	28589	36,476
Financial development index	0.03	32.33	42.60	433.10	17,864
External financial development	-0.15	0.32	0.30	0.96	39,776
Physical capital stock	301	570,000	2,376,724	40,000,000	32,384
Physical capital intensity	0.04	0.13	0.16	1.04	39,776
Human capital index	1.13	2.20	0.50	3.27	26,752
Human capital intensity	0.90	0.98	0.02	1.00	39,776
Legal institutions index	0.00	5.13	2.32	10.00	15,752
Contract intensity	0.68	0.93	0.09	1.00	39,776

Table A.4: Base Results (Time-invariant Trade Cost Sensitivity: 1995)

(a) OLS Estimates

Dependent Variable: Export Share

	(1)	(2)	(3)	(4)
Trade costs x cost intensity	-4.526***	-1.010***	-2.294***	-2.654***
	(0.039)	(0.055)	(0.144)	(0.112)
Country dummies		Y	Y	Y
Industry dummies			Y	Y
Time dummies				Y
R^2	0.32	0.63	0.69	0.69
N	22,253	22,253	22,253	22,253

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients on the fixed effects are not reported. These coefficients were obtained using the OLS estimator using Stata 13.0 SE.

(b) PPML Estimates

Dependent Variable: Export Share

	(1)	(2)	(3)	(4)
Trade costs x cost intensity	-2.327***	-0.310***	-2.741***	-2.949***
	(0.045)	(0.073)	(0.205)	(0.271)
Country dummies		Y	Y	Y
Industry dummies			Y	Y
Time dummies				Y
R^2_p	0.13	0.31	0.32	0.32
N	32,868	32,868	32,868	32,868

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients on the fixed effects are not reported. These coefficients were obtained using the PPML estimator using Stata 13.0 SE

Table A.5: Heterogeneity across Industries

Dependent Variable: Export Share	
	(1)
Leather products, luggage	-0.337
	(0.219)
Office equipment, computers	-1.926***
	(0.215)
Chemicals, chemical products	-1.305***
	(0.175)
Coke and refined petroleum	-0.466
	(0.325)
Communication equipment	-2.804***
	(0.198)
Electrical machinery	-3.618***
	(0.209)
Fabricated metal products	-2.085***
	(0.188)
Food products and beverages	-1.018***
	(0.200)
Furniture, manufacturing n.e.c	-1.015***
	(0.195)
Machinery and equipment	-2.175***
	(0.198)
Manufacturing of basic metals	-0.571**
	(0.237)
Medical instruments	-2.398***
	(0.182)
Motor vehicles, trailers	-4.149***
	(0.250)
Non-metallic minerals	-1.738***
	(0.193)
Other transport equipment	-1.010***
	(0.244)
Paper and paper products	-1.607***
	(0.236)
Publishing, printing	-2.251***
	(0.193)
Rubber and plastics products	-1.486***
	(0.155)
Textile products	-0.629***
	(0.202)
Wearing apparels, dresses	-0.307
	(0.210)
Wood, cork, straw products	-0.665***
	(0.198)
R2_p	0.336
N	32,868

Robust standard errors are in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The coefficients on the fixed effects are not reported. These coefficients were obtained using the PPML estimator using Stata IC 13.0

Table A.6: Country and Industry Characteristics

Dependent Variable: Export Share

	(1)	(2)	(3)	(4)
Trade cost x Trade cost intensity	-0.888***	-1.043***	-0.879***	-0.665***
	(0.236)	(0.247)	(0.151)	(0.147)
Physical capital x Physical capital intensity	0.758***	0.778***	0.817***	0.897***
	(0.187)	(0.148)	(0.176)	(0.201)
Human capital x Human capital intensity		6.415***	6.153***	3.560
		(1.578)	(1.631)	(2.176)
Rule of law x Contract intensity			-0.168***	-0.228***
			(0.064)	(0.067)
Financial development x Ext. financial dependence				0.171***
				(0.064)
Fixed effects	Y	Y	Y	Y
Observations	26,330	18,506	12,609	8,948

Standard errors in parentheses; * p < 0.10, ** p < 0.05, *** p < 0.01; Coefficients on the fixed effects are not reported. These coefficients were obtained using the PPML estimator using Stata IC 13.0

Annex 1: Data Sources for Country, Industry and Policy Variables

The description of country, industry, and policy variables (used in Table 8) is as follows:

*Cost of importing***:** Taken from the World Bank (ranging from 317 to 7709 USD). It includes all official costs measured in US dollars (except tariffs or trade taxes) incurred for importing a 20-foot container. It indicates administrative burden on new business for imports, and the higher values indicate higher burden.

Cost of exporting: Taken from the World Bank (ranging from 295 to 5491 USD). It includes all official costs measured in US dollars (except tariffs or trade taxes) incurred for exporting a 20-foot container. It is another indicator of administrative burden.

*Freedom to trade***:** Taken from the Heritage Foundation. It is an economic freedom index (EFI), which reflects an economy's openness to the flow of goods and services from around the world and the citizen's ability to interact freely as buyer or seller in the international marketplace. The index ranges from 0 to 100 and the higher values indicate higher level of freedom to trade.

Legal rights index: Taken from the World Bank. It measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders and thus facilitate lending. The index ranges from 0 to 10, with higher scores indicating that these laws are better designed to expand access to credit.

Tariffs: Taken from the World Integrated Trade Solutions (WITS) dataset. It is ad valorem tariff levied by the United States (common trading partner) on the imports from its trading partners.

Non-tariffs measures: Taken from the World Bank trade cost dataset. It is ad-valorem equivalent of all additional costs other than tariff costs involved in trading goods bilaterally rather than domestically. Its values range from 20% to 550%.

GDP per capita: Taken from the World Bank. It measures income level of trading partners based on purchasing power parity (PPP) at current prices in the US dollars.

Annex 2: Data Sources for Country and Industry Characteristics

The description of country and industry characteristics (used in Table A9) is as follows:

Physical capital stock: Taken from the Penn World Tables (PWT 08) generated by Feenstra et al. (2013) which constructed it using the data of initial assets, investment, and depreciation.

Physical capital intensity: Taken from EU Klems database. It is a ratio of gross fixed capital to the number of employees in the industry.

Human resources index: Taken from the Penn World Tables (PWT 08). It is based on the average years of schooling of population aged fifteen and above and assumed returns from education as in Psacharopulos (1994).

Human capital intensity: Taken from EU Klems database. It is the number of hours worked by the high-skilled workers in each industry in the United States relative to the number of hours of worked by medium and low skilled workers.

Rule of Law: Taken from Heritage Foundation (http://www.heritage.org). It is an economic freedom index for the protection of property rights.

Contract intensity: Taken from Nunn (2007). It is a measure of relationship specific investment across industries and is computed as a fraction of inputs neither traded on organised exchanges nor having a reference price.

Financial development: Taken from the World Bank dataset. It is measured as a ratio of stock market capitalization to GDP.

External financial dependence: Taken from Rajan and Zingales (1998). It is defined as a ratio of the difference between capital expenditures and cash flow (from operations) to the capital expenditures.