DOCUMENT

DE TRAVAIL

N° 542

SPECIALIZATION PATTERNS IN INTERNATIONAL TRADE

Walter Steingress

March 2015

BANQUE DE FRANCE
EUROSYSTÈME

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

SPECIALIZATION PATTERNS IN INTERNATIONAL TRADE

Walter Steingress

March 2015

Specialization Patterns in International Trade*

Walter Steingress

Banque de France

*I thank Kristian Behrens, Andriana Bellou, Rui Castro, Jonathan Eaton, Stefania Garetto, Ulrich Hounyo, Joseph Kaboski, Raja Kali, Baris Kaymak, Michael Siemer, Ari Van Assche, Silvana Tenreyro and Michael Waugh for their useful comments and suggestions. This paper also benefited greatly from comments by seminar participants at Boston University, Carleton University, Georgetown, the University of Laval, HEC Montreal and the University of Montreal. All errors are my own. Contact: Banque de France, 31 Rue Croix des Petites Champs, Paris 75001, France (e-mail: walter.steingress@banque-france.fr).

Abstract

The pattern of specialization is key to understanding how trade affects the production structure of an economy. To measure specialization, I compute concentration indexes for the value of exports and imports and decompose the overall concentration into the extensive product margin (number of products traded) and intensive product margin (value of products traded). Using detailed product-level trade data for 130 countries, I find that exports are more concentrated than imports, specialization occurs mainly in the intensive product margin, and larger economies have more diversified exports and imports because they trade more products. Based on these facts, I assess the ability of the Eaton-Kortum model, the workhorse model of modern Ricardian trade theory, to account for the observed patterns. The results show that specialization through comparative advantage induced by technological differences can explain the qualitative and quantitative facts. The key determinants of specialization are: the degree of absolute and comparative advantage, the elasticity of substitution and geography.

Keywords: Ricardian trade theory, specialization, import concentration, export concentration

JEL odes: F11, F14, F17

Résumé

La spécialisation est centrale pour comprendre comment l'impact du commerce a la structure de production d'une économie. Pour mesurer le niveau de la spécialisation, je calcule les index de concentration pour la valeur des importations et des exportations et décompose la concentration totale dans la marge de produits extensive (nombre de produits commercialisés) et la marge de produits intensive (volume de produits commercialisés). En utilisant des données commerciales détaillées au niveau du produit dans 130 pays, mes résultats montrent que les exportations sont plus concentrées que les importations, que la spécialisation se produit principalement au niveau de la marge intensive du produit, et que les économies plus grandes disposent d'importations et d'exportations plus diversifiées, car elles commercialisent plus de produits. Compte tenu de ces faits, j'évalue la capacité du modèle Eaton-Kortum, le principal modèle de la théorie ricar- dienne du commerce, pour représenter les preuves empiriques. Les résultats montrent que la spécialisation à travers l'avantage comparatif induit par les différences de technologie peut expliquer les faits qualitatifs et quantitatifs. De plus, j'évalue le rôle des déterminants clés de la spécialisation : le degré de l'avantage comparatif, l'élasticité de la substitution et la géographie.

Mots-clés: théorie ricardienne, spécialisation, concentration d'exportation, concentration d'importation

Codes JEL: F11, F14, F17

Non-technical summary

In the absence of international trade, a country can only consume what it produces. However, once the country opens up to international trade, it specializes in the production of certain goods in exchange for foreign products. Specialization can take place along two dimensions: first, in the range of goods (producing and exporting as many goods as possible), and, second, in the volume of goods (exporting a good to as many destinations as possible). Based on detailed product level data from 130 countries, the results show that exports are more concentrated than imports and the concentration in volume dominates concentration in the range of products. This implies that countries export and import a fairly large number of products but trade volume is concentrated in few products. In terms of cross-country differences, the results show that larger economies have more diversified imports and exports. This is mostly along the extensive margin, i.e. large economies export and import a wider product range

International trade theory offers different explanations of how countries specialize in the number and sales volume of goods. This paper focuses on the Ricardian theory of comparative advantage and uses the model's implication to shed light on the underlying determinants of the empirical specialization patterns. The simulation of the model shows that comparative advantage can explain the qualitative and quantitative facts. The key determinants of specialization are: the degree of absolute and comparative advantage, the elasticity of substitution and trade costs. The higher the absolute advantage of a country, the higher the level of its technology relative to the rest of the world and the more products a country will export. The degree of comparative advantage increases relative costs across products and heightens concentration. Trade costs impede trade and increase concentration of both, exports and imports. A higher elasticity of substitution provides for better substitution between goods and allows countries to concentrate their expenditure towards low price products.

By looking through the lenses of export and import concentration, this paper analyses how openness to trade changes the production structure of an economy. Openness to trade has important macroeconomic policy implications. Specialization increases a country's exposure to shocks specific to the sectors in which the economy concentrates. As a result, the likelihood that product specific shocks have aggregate effects in terms of output volatility and/or a negative impact on the terms of trade increases with openness. Diversifying along the extensive margin reduces such risks. At the same time, openness allows countries to diversify on the intensive margin by exporting to many different destinations. This reduces the exposure to country specific shocks and may reduce aggregate volatility.

1 Introduction

The pattern of specialization is at the core of international trade theory. A consequence of international trade is that countries do not need to produce all their goods: - instead they can specialize in the production of certain goods in exchange for others. Trade theory offers different explanations of how countries specialize in the number and sales volume of goods. Assessing the empirical relevance of the underlying theory is of vital interest since it not only allows us to evaluate the gains from trade due to specialization but also informs on how trade affects the structure of an economy.[1]

The contribution of this paper is twofold. First, it documents facts on the pattern of specialization by looking at both export *and* import concentration. It decomposes the overall level of concentration into a measure for the extensive and intensive product margin and documents concentration levels for exports and imports in both margins. The extensive product margin indicates the degree of specialization in the *number* of goods traded. The concentration index for the intensive margin measures specialization in the *value* of goods traded. Secondly, the paper evaluates the Eaton and Kortum (2002) model's ability to account for the observed specialization patterns. Specifically, it assesses the model based on three basic questions about specialization: What explains the level of specialization in exports and imports? What determines the gap between import and export specialization? Does specialization occur in the intensive or extensive product margin?

The starting point of my analysis is an empirical assessment of cross-country specialization patterns using several measures of concentration for exports and imports. Based on product-level trade data, the results show that countries specialize more in exports than imports and that specialization occurs predominately in the intensive product margin. This implies that countries export and import a fairly wide range of products but that trade value is concentrated in a small number of products. The gap between export and import concentration is due to the fact that countries specialize in exporting a few goods and diversify their imports by acquiring a large number of products from abroad. Focusing on cross-country differences, the results show that larger economies have more diversified imports and exports. This diversification is mostly along the extensive margin, i.e. large economies export and import a wider product range.[2]

Having documented the observed specialization pattern, I then simulate the standard Ricardian trade model developed by Eaton and Kortum (2002) to shed light on the underlying determinants

[1]For example, a high degree of specialization increases the likelihood that product-specific shocks will have aggregate effects in terms of output volatility and/or an impact on the terms of trade. Papers that study the link between the number of exporting sectors and volatility are, for example, Koren and Tenreyro (2007), Koren and Tenreyro (2013) and di Giovanni and Levchenko (2012).

[2]These results are robust to different levels of aggregation and product classification schemes (NAICS, HS, SITC).

of specialization. The principal factors are trade costs together with the elasticity of substitution and the degree of absolute and comparative advantage. The simulation also shows that for a given set of trade costs, we can replicate the observed cross-country specialization patterns. Key ingredients are asymmetric trade costs as in Waugh (2010) with small economies facing higher costs to export than large economies. However, once we calibrate trade costs following the standard approach in the literature, the simulated results show that the model produces the observed specialization pattern qualitatively but not quantitatively. The main reason is the underlying productivity distribution. In the simulation countries export their goods to too many countries in comparison with the data.

At this point, it is important to note that the Ricardian model shares with other models of international trade, most notably monopolistic competition models based on Krugman (1980) and Armington models like Anderson and Van Wincoop (2003), the ability to develop quantitative predictions about specialization patterns in the intensive and extensive product margin (see Hummels and Klenow (2005)). However, in these models, tradable goods are differentiated by location of production since each country is the sole producer of a good. Countries specialize completely and demand all product-country combinations. When this definition of the product space is applied to the data, the analysis shows that countries are more specialized in imports than in exports because they import only a small subset of all available products. This result implies that the empirical implications depend on the definition of the product space, i.e. differentiated versus homogenous goods. Consistent with the Ricardian model, the empirical analysis in this paper is based on the assumption that foreign varieties are perfect substitutes for domestic ones and that local producers compete directly with imports on the basis of price. The robustness section discusses the alternative results based on the Armington assumption.

This paper contributes to the extensive literature that analyses cross-country patterns of specialization in production (for example Imbs and Wacziarg (2003), Koren and Tenreyro (2007) and Koren and Tenreyro (2013)), exports (Schott (2004) and Cadot et al. (2011)) and imports (Jaimovich (2012), Cadot et al. (2014)). While the literature focuses mainly on each of the sectors separately, the aim of this paper is to fill the gap by exploring the cross-country patterns of both export and import concentration. To do so, I apply the analysis of Cadot et al. (2011) to exports and imports and compare the measures obtained along the extensive and intensive margins. The results show that, for the majority of countries, exports are indeed more concentrated than imports. However, for some large economies, imports are more concentrated than exports.

The analysis also relates to the international trade literature on the relationship between income and trade patterns in the intensive and extensive margins (Hummels and Klenow (2005) and Cadot et al.

(2011) for the product level and Eaton et al. (2011) for the firm level). Hummels and Klenow (2005) test several models that rely on the Armington assumption on the positive link between country size and the extensive margin of exports. Cadot et al. (2011) explore the decomposition of the measure of concentration into an extensive and an intensive margin for exports and study cross-country differences over time. The analysis in this paper builds on the previous papers and extends the decomposition to imports. The novelty of this paper is that it synthesizes the individual facts on export and import concentration along both margins and to show that the Ricardian model of Eaton-Kortum (EK) replicates the observed patterns.

This paper contributes to the growing literature in quantifying the importance of Ricardian comparative advantage in explaining trade patterns using the EK framework, (for example, Chor (2010), Shikher (2011), Levchenko and Zhang (2011) and Costinot et al. (2012)). These papers specify a multi-sector Ricardian model with both inter- and intra-industry trade in order to derive implications on a sectoral level. In contrast, I abstract from intra-industry trade and attach a sectoral interpretation to the continuum of traded goods within the standard Eaton-Kortum framework. Given this notion, the number of traded sectors arises endogenously and is not fixed as in the previous papers. While the standard model has been primarily used to explain bilateral trade flows and trade volume, (see, for example, Eaton and Kortum (2002), Alvarez and Lucas (2007) and Waugh (2010)), I focus on the implications for the pattern of trade and analyze how geography, tastes and absolute and comparative advantage induce countries to specialize in narrow sectors.

The rest of the paper is organized as follows. Section 2 describes the data and presents the empirical evidence for import and export concentration. Section 3 lays out the theoretical framework. Section 4 describes the calibration that allows the model to replicate the empirical facts. Section 5 estimates trade costs and presents the simulation results based on the estimated trade costs. Section 6 discusses the robustness of the results and section 7 concludes.

2 Empirical evidence and data

The starting point of my analysis is an empirical assessment of the observed specialization patterns in world trade using detailed product-level trade data. Before describing the data and the empirical evidence, I examine the properties of the concentration measurements used, which form the basis for the qualitative and quantitative tests of the model.

2.1 Concentration measurements

I compute two measures of specialization for product level sales, the Gini coefficient and the Theil index. The Theil index has the advantage of being decomposable into an extensive and intensive product margin measure. For concreteness, I focus on exports - concentration measures for imports are entirely analogous. The two measurements are defined as follows. Let k index a product among the N products found in the world economy, let R_k be the corresponding export sales revenue, say, in a given country. The export Gini in this country is defined as :

$$G = \frac{2}{N}\frac{(\sum_{k=1}^{N} kR_k)}{\sum_{k=1}^{N} R_k} - \frac{N+1}{N} \tag{1}$$

where export revenues for product k, R_k, are indexed in increasing order of size, i.e. $R_k < R_{k+1}$, and N denotes the total number of products in the world. A Gini coefficient of zero expresses complete diversification across trade revenues, i.e. (1) a country exports all products and (2) and each prodcut generates the same amount of revenues. An index of one expresses complete specialization in which case export revenues stem from one product only. Alternatively, the Theil index is a weighted average of the log difference in relation to the mean export revenue $(\bar{R})$ and is defined by the following formula

$$T = \frac{1}{N}\sum_{k\in N}\frac{R_k}{\bar{R}}\ln\left(\frac{R_k}{\bar{R}}\right) \tag{2}$$

The index takes the value of zero in the case of complete diversification and $\ln(N)$ in the case of complete specialization. Cadot et al. (2011) decompose the Theil index into a measure for the intensive and extensive product margins, $T = T^{ext} + T^{int}$. The extensive Theil index (T^{ext}) captures the concentration in the number of products (extensive product margin) whereas the intensive Theil (T^{int}) measures the concentration in the sales volume of products (intensive product margin). The intensive Theil index is given by:

$$T^{int} = \frac{1}{N_x}\sum_{k\in N_x}\frac{R_k}{\bar{R}_x}\ln\left(\frac{R_k}{\bar{R}_x}\right) \tag{3}$$

and the extensive Theil index is

$$T^{ext} = \ln\left(\frac{N}{N_x}\right) \tag{4}$$

N_x denotes the number of exported products and $\bar{R}_x$ represents the mean value of exported products.

2.2 Data

To build my empirical evidence, I use the BACI data set provided by CEPII (Gaulier and Zignago (2009)) and choose the 1992 6-digit HS product classification scheme as the preferred level of disaggregation. I follow Hummels and Klenow (2005) and refer to import flows of the same 6-digit product from different trading partners as different varieties of the same product. I assume that the tradable goods sector corresponds to the manufacturing sector.[3] Using a correspondence table provided by Feenstra et al. (1997), I identify a total of 4,529 tradable products. The baseline sample covers 130 countries representing all regions and all levels of development between 1995 and 2011 (17 years). In total, the sample consists of 2210 observations (country-years).

Note that the data contain import and export flows for each 6 digit product category. The model I am assessing is Ricardian and does not feature trade in different varieties of the same product. To establish a mapping between the model and the data, I net out the within-product component by considering net trade flows instead of gross trade flows.[4] To measure the importance of trade in products and trade in varieties of products, I follow Grubel and Lloyd (1975) and calculate the percentage share of trade between products with respect to total trade. I obtain an average value of 81 percent across all countries. The overall share of the total value of net trades flows with respect to gross trade flows is 66 percent. Both findings suggest that the majority of trade flows between countries in this sample is in products.[5]

Based on net trade flows at the product level, I calculate concentration indexes for each country on all margins for each year and then take the average over the whole sample period. Because the concentration indexes used are independent of scale, the calculation on a year-to-year basis avoids the need to deflate the data. Figure 1 plots the mean export concentration against the mean import concentration for each country, together with the 45 degree line. In terms of overall concentration, Figures 1(a) and 1(b), show that the vast majority of observed levels lie above the 45 degree line indicating that exports are more concentrated than imports for almost all countries. With regard to the intensive product margin, Figure 1(c), the specialization level of exports is slightly higher than that of imports. Figure 1(d) plots the results for the extensive product margin with countries exporting fewer products than they import.

[3]This is a simplification, but it is reasonable as a first-order approximation because, for all countries in the sample, this represents on average 76 percent of all merchandise imports; the median is 91 percent.

[4]I compute total net exports at the 6-digit product level and consider a country as an exporter of that product if net exports are positive and an importer otherwise.

[5]In the appendix I present an alternative approach to account for observed intra-industry trade in the data. The basic idea is to develop a measurement device that enables the model to characterize trade within and across products. The suggested procedure converts the product units in the model to product units in the data and allow us to examine specialization patterns based on gross trade flows. In the rest of the paper, I follow the net trade flow approach. I present the estimation and results of the alternative procedure in the appendix.

Table 1: Mean concentration indexes over 2210 country-year pairs

	Gini		Theil Exports (X)			Theil Imports (M)		
	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
Level of concentration	0.98	0.91	2.60	2.13	4.73	1.10	1.61	2.71
% share of overall concentration			55%	45%		40%	60%	

Figures 1(a) and 1(b) also reveal that some countries have higher import than export concentration (the United States, China, Germany, India and Italy). An examination of the extensive margin shows that these countries import the same number of products as they export, so the difference must stems from the intensive margin: in order words these economies have diversified export revenues and spend their import expenditure on relatively few goods. The reason for the diversification of export revenues across products is that export revenues are also diversified across countries, i.e. exporter export each of their products to many different destinations. Indeed, Figure 2 shows a significant negative correlation between export concentration and the average number of export destinations per product, pointing to complementarities in the concentration of export revenues in terms of products and in terms of destinations.

Table 1 summarizes the sample statistics by giving the average year-by-year indexes for the 2210 country-year pairs. As implied by Figure 1, exports are more concentrated than imports on all margins. With respect to overall concentration, the summary statistics reveal high levels of export and import concentration, with a Gini coefficient of 0.98 for exports and 0.91 for imports. In the case of exports, the high level of concentration is due to the fact that countries export few products and hence specialization is primarily driven by the extensive margin. For imports, the decomposition favors an alternative explanation. Countries import a fairly wide range of products but concentrate their trade in the value of few products. With regard to the gap between export and import concentration, Table 1 shows that this can mainly be explained by the extensive margin. The Theil of 1.10 on the import extensive margin implies that, on average, a country net imports 33.3 percent of all products. On the other hand, the extensive Theil for exports indicates that a country net exports 7.4 percent of all products. In terms of the intensive margin, a country receives roughly 50% of its export revenues from 1% of the products it exports and spends 50% of its import expenditure on 2% of the products it imports. Overall, these results are consistent with the idea that openness to trade spurs countries to specialize in a few exporting sectors and to diversify its importing sectors.

Turning to cross-country differences, the empirical evidence shows that larger economies diversify more than smaller economies. Figure 3 plots the log of the mean levels of concentration as a function of market size, including the best linear fit for all margins. Market size is measured by the log of average GDP relative to the United States ($USA = 0$). As Figures 3(a) and 3(b) show, the overall Theil index decreases as relative GDP increases, i.e. smaller economies specialize more. This relationship is more pronounced for exports than for imports, with an R square of 0.58 compared to 0.41. The decomposition reveals that intensive margin specialization does not vary with market size for either exports (Figures 3(e)) or imports (3(f)). The main driver of specialization differences across countries is therefore the extensive margin.The linear relationship is especially robust for the export extensive margin, with an R square of 0.75. Bigger economies are more diversified because they export more products, which is consistent with Koren and Tenreyro (2007)'s observation that larger economies are more diversified because they produce and export more products. The relationship between market size and specialization in the extensive margin of imports follows an L-shaped pattern. As an economy increases in size, it diversifies its imports, until it reaches a certain market size after which concentration is roughly equal across countries.

At this point, the key qualitative and quantitative facts have been established. First, exports are more specialized than imports. Second, concentration is driven by the extensive margin for exports and by the intenisve margin for imports. Third, the target levels of concentration are displayed in Table 1. Fourth, the cross-country patterns imply a negative relationship between market size and specialization caused by the extensive margin, i.e. larger economies export and import more products. The rest of this paper evaluates the Ricardian model's ability to account for these stylized facts. In the next section, I present the relevant parts of the Alvarez and Lucas (2007) extension of the Eaton-Kortum framework.

3 Model

The Eaton–Kortum model is Ricardian, with a continuum of goods produced under a constant-returns technology. In this paper, I focus on the Alvarez and Lucas (2007) model and include capital as in Waugh (2010). Next, I derive the relevant theoretical predictions on the pattern of trade and evaluate the importance of the key model parameters for import and export specialization.

Consider a world economy with I countries, where each country produces tradable intermediate goods as well as non-tradable composite and final goods. Following Alvarez and Lucas (2007), I define $x = (x_1, ..., x_I)$ as a vector of technology draws for any given tradable good and refer to it as "good x" with $x \in \mathbf{R}^I_+$. The production of an intermediate good in country i is defined by:

$$q_i(x_i) = x_i^{-\theta}[k_i^{\alpha} s_i^{1-\alpha}]^{\beta} {q_{mi}}^{1-\beta}.$$

Technology x_i differs between goods and is drawn independently from a common exponential distribution with density ϕ and a country specific technology parameter λ_i, i.e. $x_i \sim \exp(1/\lambda_i)$. I denote the interest rate by r_i, the wage by w_i and the price of the intermediate aggregate good by $p_{m,i}$. The intermediate good sector is perfectly competitive. Producers of the intermediate good minimize input costs and sell the tradable intermediate good at price

$$p_i(x_i) = B x_i^{\theta} [r_i^{\alpha} w_i^{1-\alpha}]^{\beta} p_{mi}^{1-\beta}.$$

where $B = \beta^{-\beta}(1-\beta)^{-(1-\beta)}$. The continuum of intermediate input good x goes into the production of the composite good q_i symmetrically with a constant elasticity of substitution $(\eta > 0)$

$$q_i = \left[\int_0^{\infty} q(x)^{1-1/\eta} \phi(x) dx \right]^{\eta/(1-\eta)}.$$

The aggregate output of intermediate good q_i can then be allocated at no cost to the production of final goods or can be used as an input in the production of other intermediate goods. Similarly, capital and labor can be used either to produce intermediate or final goods. Finally, consumers draw utility linearly from the final good. All markets are perfectly competitive. Since these features are not central to the implications I derived in this paper, I omit them. I refer interested readers to Alvarez and Lucas (2007) for a full description of the model.

3.1 General equilibrium

Once a country opens up to international goods markets, intermediate goods are the only goods traded. Final goods are not traded and capital and labor are immobile between countries. Trading intermediate goods between countries is costly. We define "Iceberg" transportation costs for good x from country i to country j by κ_{ij} where $\kappa_{ij} < 1 \ \forall \ i \neq j$ and $\kappa_{ii} = 1 \ \forall i$. As in Alvarez and Lucas (2007), we also consider tariffs. ω_{ij} is the tariff levied by country i on goods imported from country j. Tariffs distort international trade but do not entail a physical loss of resources. Incorporating the trade costs, composite good producers in country i will buy the intermediate good x from country j that offers the lowest price

$$p_i(x) = B \min_j \left[\frac{[r_j^{\alpha} w_j^{1-\alpha}]^{\beta} p_{mj}^{1-\beta}}{\kappa_{ij}\omega_{ij}} x_j^{\theta} \right]. \tag{5}$$

Equation 5 shows that whether country i specializes in the production of good x depends on the productivity realizations, factor prices and trade costs. If country i does not offer a good at the lowest cost in the local market, the good is imported. Following Alvarez and Lucas, the resulting price index of tradable goods in country i is

$$p_{mi} = (AB) \left(\sum_{j=1}^{I} \left(\frac{w_j^{\beta} p_{mj}^{1-\beta}}{\kappa_{ij}\omega_{ij}} \right)^{-1/\theta} \lambda_j \right)^{-\theta} \tag{6}$$

where $A = \Gamma(1 + \theta(\eta - 1))$ is the Gamma function evaluated at point $(1 + \theta(\eta - 1))$. Next, we calculate the expenditure shares for each country i. Let D_{ij} be the fraction of country i's per capita spending $p_{mi}q_i$ on tradables that is spent on goods from country j. Then, we can write total spending of i on goods from j as

$$p_{mi}q_i D_{ij} = \int_{\mathbf{B}_{ij}} p_i(x) q_i(x) \phi(x) dx$$

where $\mathbf{B}_{ij}$ defines the set of goods country j attains as a minimum in equation 5. Note that D_{ij} is simply the probability that country j is selling good x in country i at the lowest price and calculated to be

$$D_{ij} = (AB)^{-1/\theta} \left(\frac{[r_j^{\alpha} w_j^{1-\alpha}]^{\beta} p_{mj}^{1-\beta}}{p_{mi}\kappa_{ij}\omega_{ij}} \right)^{-1/\theta} \lambda_j. \tag{7}$$

Equation 7 shows that in this model the sensitivity of trade between countries i and j depends on the level of technology λ, trade costs ω, geographic barriers κ and the technological parameter θ (reflecting the heterogeneity of goods in production) and is independent of the elasticity of substitution η. This result is due to the assumption that η is common across countries and does not distort relative good prices across countries. Note also that, by the law of large numbers, the probability that country i imports from country j is identical to the share of goods country i imports from j. In this sense, trade shares respond to costs and geographic barriers on the extensive margin: as a source becomes more expensive or remote it exports/imports a narrower range of goods. It is important to keep in mind that the number of intermediate input industries that enter into the production of the composite good is fixed. Each country uses the whole continuum of intermediate goods to produce composite goods. There are no gains from trade due to an increased number of varieties. Welfare gains are realized through incomplete specialization. Domestic production competes with imports and countries specialize through the reallocation of resources made available by the exit of inefficient domestic producers.

To close the model, we impose that total payments to foreigners (imports) are equal to total receipts from foreigners (exports) for all countries i

$$L_i p_{mi} q_i \sum_{j=1}^{I} D_{ij} \omega_{ij} = \sum_{j=1}^{I} L_j p_{mj} q_j D_{ji} \omega_{ji} \tag{8}$$

The previous equation implies an excess demand system which depends only on wages. Solving this system, describes the equilibrium wage for each country together with the corresponding equilibrium prices and quantities. Next, I describe the predictions for export and import concentration in both margins.

3.2 Concentration of exports and imports

In the model, the pattern of trade is established by domestic producers competing with importers to sell intermediate goods in the local market. Given the equilibrium price, $p(x)$, and quantity, $q(x)$, the total amount that country i spends (c.i.f.) on imported good x, $R_{iM}(x)$, is:

$$R_{iM}(x) = L_i p_i(x) q_i(x) \qquad x \notin \mathbf{B}_{ii}$$

where $\mathbf{B}_{ii} \subset \mathbf{R}_+^I$ is the set of goods for which country i obtains the minimum price at home. Similarly, domestic producers export their good to all foreign markets where they attain the minimum price. The set of exported goods is simply a the sum of goods country i exports to any destination j, $x \in \cup_{j \neq i}^{I} \mathbf{B}_{ji}$. As a result, (f.o.b.) export revenues for good x, $R_{i,X}(x)$, are given by:

$$R_{iX}(x) = \sum_{k \neq i}^{I} L_k p_k(x) q_k(x) \kappa_{ki} \omega_{ki} \qquad x \in \cup_{j \neq i}^{I} \mathbf{B}_{ji}$$

Given the described pattern of trade, the concentration index for imports is identified. To show this, I decompose the overall concentration into a concentration measure for the intensive and extensive product margins. Using equation 3, the Theil index for the concentration of imports in the intensive margin can be written as:

$$T_{iM}^{int} = \int_{x \notin \mathbf{B}_{ii}} \frac{R_{iM}(x)}{\bar{R}_{iM}} \ln\left(\frac{R_{iM}(x)}{\bar{R}_{iM}}\right) \phi(x) dx$$

In the appendix, I show that import expenditure follows a Fréchet distribution with shape parameter $1/\theta(\eta-1)$ and scale parameter s_i. Solving the integral, the intensive Theil index of imports for country i becomes:

$$T_{iM}^{int} = \ln\left(\Gamma(1+\theta(1-\eta))\right) - \int_0^1 \ln\left(u^{(-\theta(1-\eta))}\right) e^{-u} du \tag{9}$$

where $\Gamma(.)$ stands for the Gamma function. Import specialization in the intensive margin is independent of equilibrium prices, trade costs, geography and the level of technology λ. It is determined solely by preferences (i.e the elasticities of substitution) and heterogeneity in production (i.e. the degree of comparative advantage). A higher elasticity of substitution (η) increases specialization by allowing producers in the composite intermediate good sector to better substitute cheap for expensive products and concentrate expenditure in these sectors. Similarly, an increase in the degree of comparative advantage (θ), which corresponds to a higher variance of productivity realizations and therefore an increase in unit price differences across goods, heightens the degree of concentration.

To compute the concentration in the extensive margin of imports, note that the set of goods produced is disjoint form the set of goods imported. Consequently, we can express the share of goods imported as 1 minus the share of goods produced, $(1-D_{ii})$. The Theil index for the extensive margin of imports is equal to :

$$T_{iM}^{ext} = \ln\left(\frac{N}{N_{iM}}\right) = -\ln(1-D_{ii}) \tag{10}$$

where

$$D_{ii} = (AB)^{-1/\theta}\left(\frac{[r_i^{\alpha} w_i^{1-\alpha}]}{p_{mi}}\right)^{-\beta/\theta} \lambda_i$$

and depends on the level of technology and equilibrium prices. To assess the level of specialization in exports, I simulate the model within a discrete product space in the following section. I then calculate the export concentration index for the intensive margin according to equation 3.

Having outlined the pattern of trade and the corresponding implications for the specialization pattern of exports and imports, the next section discusses the simulation of the model. It contains special equilibria cases designed to spell out step-by-step the main implications of the model for export and import concentration.

4 Calibration and simulation

To simulate the theoretical model, which assumes an infinite amount of goods, I "discretize" the Fréchet distribution of total factor productivity and calculate the respective trade value for each product x. Regarding the parameters of the model, we need values for α, β, γ, η and θ. For α, β and γ, I use the same values as Alvarez and Lucas (2007): I set the capital share to $\alpha = 0.3$, the efficient labor share in the tradable goods sector to $\beta = 0.5$ and the labor share in the production of non-tradable final goods to $\alpha = 0.75$.

To calibrate the elasticity of substitution (η) and the variance of the productivity distribution (θ), I use the model's implication on the value and the volume distribution of imports. As shown in the previous section, the distribution of import value follows a Fréchet with shape parameter $1/\theta(\eta - 1)$. Similarly, the distribution of the quantities imported can also be shown to follow a Fréchet with shape parameter $1/(\theta\eta)$. Using the fact that the Theil index for the intensive margin depends solely on the shape parameter, I first calculate the average Theil index for import expenditure, $T_{iM}^{int} = 1.61$, and for imported quantities, $T_{iQ}^{int} = 3.58$.[6] Then, using equation 9, I get the corresponding shape parameters and obtain 2 equations with 2 unknowns. The solution of the system gives an elasticity of substitution ($\eta = 8$) and a degree of comparative advantage ($\theta = 0.10$). The elasticity of substitution is high but still in the parameter range found in the literature (see Broda and Weinstein (2006)). The degree of comparative advantage also lies in the parameter range 0.08 to 0.15 estimated in the literature (see Eaton and Kortum (2002)). However, compared to more recent estimates by Simonovska and Waugh (2011), $\theta = 0.10$ is rather low.

In the following subsections, I analyze import and export concentration in special cases of equilibrium by assuming different trading schemes. Doing this builds an intuition of how taste, technology and geography determine specialization. To illustrate the impact of each factor separately, it is instructive to start the analysis by assuming symmetric countries and introducing heterogeneity across countries at a later stage. Finally, I show that, for a particular configuration of trade costs, the Eaton-Kortum model is able to replicate the specialization patterns observed in the data.

[6] In addition to the dollar value of imports, BACI also reports the volume of goods imported, which are measured in tons. To calculate the Theil index of the volume of imports, we follow the procedure applied to the import revenues. We first calculate the net volume of imports for each good and then apply equation 3. The value $T_{iQ}^{int} = 3.58$ represents the cross-country average over the sample period.

4.1 Symmetric countries

All countries are identical. Trade costs are symmetric and set to $\kappa_{ij} = \kappa \ \forall \ i \neq j$ with $\kappa_{ii} = 1$ and $\omega_{ij} = 1 \ \forall i, j$. Due to symmetry, factor prices equalize across countries. The corresponding trade share matrix D is symmetric and the (i, j) element is given by:

$$D_{i,j} = \frac{(\kappa)^{1/\theta}}{1 + (I-1)(\kappa)^{1/\theta}} \ \forall i \neq j \ \text{ and } \ D_{i,i} = \frac{1}{1 + (I-1)(\kappa)^{1/\theta}}$$

In free trade, $\kappa = 1$, each country's intermediate good producers specialize in a distinct set of goods equal to the relative size of the economy and export all products produced, $D_{ii} = D_{ij} = 1/I$. The corresponding share of imported products is $1 - D_{ii} = (I-1)/I$. In this case, Ricardian specialization forces are strongest and the gap between export and import concentration reaches a maximum.

Extensive Margin Concentration Including trade costs, the concentration index for imports equals the share of goods country i imports from all countries in the world and is given by:

$$T^{ext}_{iM} = -\ln((1 - D_{ii})) = \ln(1 + (I-1)(\kappa)^{1/\theta}) - \ln((I-1)(\kappa)^{1/\theta})$$

Concentration in the extensive margin of imports increases with trade barriers κ and decreases with the number of trading partners $I - 1$ and the degree of comparative advantage θ. Regarding exports, the extensive Theil index is given by the number of products exported to any destination divided by the total number of products in the world. To count the number of products exported, define the set of products exported as the union of the set of products exported to each destination, $U_{ex} = \cup^{I}_{j \neq i} \mathbf{B}_{ji}$. Because the set of products exported to destination j overlaps with the set of products exported to destination k, $\mathbf{B}_{ji} \cap \mathbf{B}_{ki} \neq \emptyset$, I apply the Inclusion Exclusion principle to avoid double counting. As I show in the appendix, under the assumption of symmetry, the extensive Theil index of exports is given by:

$$T^{ext}_{iX} = -\ln\left(\sum_{k=1}^{I-1} (-1)^{k-1} \begin{pmatrix} I \\ k \end{pmatrix} a_k\right) \tag{11}$$

where the share of products exported to k destinations, a_k, is given by:

$$a_k = \frac{(\kappa)^{1/\theta}}{k + (I-k)(\kappa)^{1/\theta}}$$

The concentration of exports increases with geographical barriers, the degree of comparative advantage and the number of trading partners. In general, a larger number of trading partners increases

Table 2: Simulated export and import concentration indexes for benchmark parameters.

Parameters	Gini		Theil Exports (X)			Theil Imports (M)		
	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
$(\kappa = 1)$	0.99	0.72	5.01	1.91	6.92	0.01	1.61	1.62
$(\kappa = 0.7)$	0.99	0.77	5.04	1.18	6.22	0.10	1.61	1.71
$(\kappa = 0.7$, NT=10)	0.98	0.86	2.47	2.45	4.92	1.09	1.61	2.70
Data	0.98	0.91	2.60	2.13	4.73	1.10	1.61	2.71

competition between production and imports in the domestic market, resulting in the production of fewer goods at home and an increase in the number of goods imported. Also, a higher number of trading partners increases competition between exporters in foreign markets, forcing countries to specialize more idn the extensive margin of exports. Impediments to trade, i.e. a reduction in κ, and a higher degree of comparative advantage, θ, reduce import competition and, as a result, fewer goods are exported and imported. Note that, in the special case of free trade, all goods produced are exported and concentration of production equals concentration of exports. Where trade costs are applied, countries export a subset of produced goods, leading to greater concentration of exports relative to production.

Intensive Margin Concentration As noted previously, import expenditure follows a Fréchet distribution and is pinned down by the elasticity of substitution (η) and the degree of comparative advantage (θ). Regarding the distribution of export revenues, the simulation shows that it depends positively on the elasticity of substitution (η), the degree of comparative advantage (θ) and geographical barriers (κ). In the case of free trade, countries export all their goods to all destinations and, given that preferences are identical, export and import concentration in the intensive margin equalize.

The results presented in Table 2 show that the free trade calibration of Alvarez and Lucas (2007) is able to replicate the qualitative fact that, overall, exports are more concentrated than imports. While the simulated overall level of export concentration attains the degree of specialization observed in the data, in the benchmark free trade parametrization countries diversify excessively in imports because they import too many goods.

Next, I introduce 42 percent symmetric trade costs for all trading partners, $\kappa = 0.7$. Row 3 of Table 2 shows the results. Impediments to trade reduce the number of products exported and imported and lead to an increase in extensive margin concentration for both, exports and imports. Note that higher trade costs lower the level of intensive margin concentration for exports. Due to the increase in trade costs, only very efficient producers export and their export revenues are more evenly distributed across products and trade partners. Still, the gap between export and import concentration remains substantial. The reason is that the degree of competition countries face in export and domestic markets is too high. In the symmetric setting the only way to reduce competition is to limit the number of trading partners. Using equation 11, the number of trading partners (NT) corresponding to the empirical Theil index is 10 (see fourth row of Table 2). Limiting the number of trading partners (NT) by introducing infinite trade costs for countries outside of the block reduces competition in all markets. Less competition in the domestic market increases the survival rate of local producers and reduces the amount of goods imported. Note that the revenues of exporting industries are now geographically more concentrated and hence specialization in the intensive margin of exports intensifies.

To summerize, with the introduction of symmetric trade costs, the model can replicate the mean levels of concentration observed in the data. In particular, by creating trade blocks, which amounts to introducing zeros in the bilateral trade matrix, we can calibrate the model to explain the mean pattern of specialization.

4.2 Asymmetric countries

In this section I analyze the effects of cross-country heterogeneity on specialization. The empirical facts imply a negative relationship between specialization and market size. For this reason, I introduce heterogeneity in technology λ_i to reflect the observed GDP differences in the data. To start with, consider equation 8 in a free trade equilibrium:

$$(w_i L_i + r_i K_i) = \sum_{j=1}^{N} \left(w_j L_j + r_j K_j\right) D_{ji}$$

which can be simplified to

$$\lambda_i = C\left(w_i L_i + r_i K_i\right)\left(w_i^{\alpha} r_i^{1-\alpha}\right)^{\frac{\beta}{\theta}} \tag{12}$$

where C is a constant. Using equation 12, I obtain technology as a function of GDP $(w_i L_i + r_i K_i)$ and endowments L_i and K_i assuming that the inputs are chosen optimal. To calibrate λ, I use GDP, capital and population data from the Penn World table. I follow Waugh (2010) and normalize the parameters obtained for λ_i, L_i and K_i relative to the United States.

Extensive Margin Concentration Plugging in the equilibrium wage into equation 7, I get the corresponding trade share matrix D with the (i, j) element given by:

$$D_{ji} = \frac{(w_i L_i + r_i K_i)}{\sum_{k=1}^{I} (w_k L_k + r_k K_k)} \ \forall j \tag{13}$$

Equation 13 shows that under the assumption of free trade country i's share of the number of products exported is equal to its level of GDP relative to world GDP. Hence, larger economies export more and import fewer products compared to smaller economies. This result is at odds with the empirical evidence. In the data, larger economies export and import more products.

4.3 Asymmetric trade costs

To reconcile the cross-country concentration differences for imports, I consider trade costs as a function of either an export cost (ex_j) or an import cost (im_i). While both types of cost can reconcile the fact that larger countries import more goods, the following exposition focuses on the export cost.[7] In this case, each country pays a country specific-cost in order to export, which is independent of the importing country j, $\kappa_{ji} = ex_i \ \forall i \neq j$ and $\kappa_{i,i} = 1 \ \forall j = i$. Due to asymmetric trade costs, wages and composite good prices do not equalize. The trade share matrix is asymmetric and given by:

$$D_{ji} = (AB)^{-1/\theta} \left(\frac{[w_i^{\alpha} r_i^{1-\alpha}]^{\beta} p_{m,i}^{1-\beta}}{p_{mj} ex_i} \right)^{-1/\theta} \lambda_i \ \forall i \neq j \quad \text{and} \quad D_{ii} = (AB)^{-1/\theta} \left(\frac{w_i^{\alpha} r_i^{1-\alpha}}{p_{m,i}} \right)^{-\beta/\theta} \lambda_i$$

Focusing on the expression for the share of goods that country i exports to country j, D_{ji}, shows that a higher export cost reduces the fraction of the good that arrives in destination j $(ex_i \downarrow)$ and decreases the number of goods country i exports to any destination j. Solving for the equilibrium and assuming that composite good prices across countries are approximately equal, one can show that the share of goods imported is approximately:

$$(1 - D_{ii}) \approx \left(1 - C_1 ex_i^{\frac{-1}{\theta}} (w_i L_i + r_i K_i) \right) \tag{14}$$

where C_1 is a constant. Equation 14 shows that the share of goods imported increases as country-specific export costs decerase, $(\partial(1 - D_{ii})/\partial ex_i > 0)$. Lower export costs allow producers to pay higher factor prices at home and still be competitive in export markets. At the same time, higher unit costs of production reduce competitiveness at home and result in a larger share of imported goods.

[7] The same reasoning can be applied to import costs.

Table 3: Simulated export and import concentration indexes for asymmetric countries.

Parameters	Gini		Theil Exports (X)			Theil Imports (M)		
	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
($\kappa = 1$)	0.99	0.73	5.75	1.91	7.66	0.007	1.61	1.62
($\kappa = ex$, NT=10)	0.98	0.85	2.59	2.67	5.26	1.10	1.61	2.71
Data	0.98	0.91	2.60	2.13	4.73	1.10	1.61	2.71

Hence, an exporter-specific cost can reconcile the fact that larger economies import and export more goods.

The main difference between import costs and export costs in terms of import concentration lies in the implication for the price of tradable goods. It is possible to show that the export cost imply a near constant price for tradable goods across countries. As a result, unit cost differences between countries are predominantly driven by factor price differences. In contrast, import costs lead to large cross-country price differences, with smaller economies facing a higher tradable price level. In this case, unit cost differences are driven by factor as well as tradable goods price level differences. Waugh (2010) provides evidence that countries have similar tradable good prices. Hence, in the remaining analysis, I focus only on the case of the exporter-specific cost.

To summerize, the introduction of asymmetric trade costs in the form of country-specific export or import costs allows the model to replicate the import specialization pattern across countries, in particular when larger economies face relatively low costs to either export or import. Waugh (2010) argues that trade costs have to be asymmetric, with poor countries facing higher export costs than rich countries, in order to reconcile bilateral trade volumes and price data. While both our approaches highlight the importance of asymmetric trade costs in explaining trade data, our analysis differs. Waugh uses the Eaton Kortum model to explain bilateral trade volumes and price data whereas I look at the model's implications for export and import specialization patterns. In this respect, the results presented in this paper provide further evidence on the importance of asymmetry in trade costs when studying trade volumes and trade patterns across countries.

Row 1 of Table 3 presents the simulation results in the case of asymmetric countries and free trade. Note that introducing technology differences increases the mean level export concentration and decreases import concentration relative to the symmetric country case. The underlying reason is that the technology distribution is skewed towards less productive countries and these countries export fewer goods and import more goods. Apart from these differences, the results are similar to the symmetric case.

To reconcile the empirical evidence that larger countries import more goods, I introduce country-specific export costs with larger economies facing lower export costs than smaller economies. In particular, I calculate the implied export cost from equation 14 by replacing the share of goods produced at home with the extensive Theil index of imports observed in the data , $D_{ii} = 1 - \exp(-T_M^{Ext})$. Row 2 of Table 3 shows the corresponding mean concentration levels. Regarding the cross-country pattern, Figure 4 plots the simulated (in red) and empirical (in blue) concentration levels against GDP for both margins. The figures show that country-specific export costs in combination with technology and endowment differences can replicate the cross-country evidence on all margins.

In the previous section I analyzed special equilibrium cases to study the different factors that determine specialization in the Eaton Kortum model. The key determinants are the degree of comparative advantage, the elasticity of substitution and asymmetric trade costs. However, I treated trade costs as free parameters and showed that, for a particular configuration of trade costs, the model is able to reproduce concentration levels for the mean as well as the cross-country specialization pattern for both exports and imports. In the next section, I estimate trade costs and technology parameters based on bilateral trade shares using the model's structure and check whether, for given trade shares, the model is able to generate the observed specialization pattern in the data.

5 Estimating trade costs from bilateral trade shares

The starting point for the estimation of technology and trade costs is a structural log-linear "gravity" equation that relates bilateral trade shares to trade costs and model's structural parameters. To derive the relationship, I simply divide each country i's trade share from country j, see equation 7, by country i's home trade share. Taking logs yields $I - 1$ equations for each country i :

$$\log\left(\frac{D_{ij}}{D_{ii}}\right) = S_j - S_i + \frac{1}{\theta}\log(\kappa_{ij}) + \frac{1}{\theta}\log(\omega_{ij}) \tag{15}$$

in which S_i represents the structural parameters and is defined as $S_i = \log([r_i^{\alpha} w_i^{1-\alpha}]^{-\beta/\theta} p_{mi}^{-(1-\beta)/\theta} \lambda_i)$. In order to estimate the trade costs κ and technology λ implied by equation 15 I use data on bilateral trade shares across 130 countries. I follow Bernard et al. (2003) and calculate the corresponding bilateral trade share matrix as the ratio of total gross imports of country i from country j, M_{ij}, divided by absorption Abs_i

$$D_{ij} = \frac{M_{ij}}{Abs_i}.$$

Absorption is defined as total gross manufacturing output plus total imports, M_i, minus total exports, X_i. To compute absorption, I use gross manufacturing output data from UNIDO.[8] Combined with trade data from BACI, I get the expenditure share, D_{ij}, which equals the value of the inputs consumed by country i and imported from country j divided by the total value of inputs in country i. Note that instead of focusing on a particular year, I compute the expenditure share for each year of the period 1995 - 2011 and take the average expenditure share over the sample period.[9]

In total, there are only $I^2 - I$ informative moments and I^2 parameters of interest. Thus, restrictions to the parameter space are necessary. To create them, I follow Eaton and Kortum (2002) and assume the following functional form of trade costs.

$$\log\left(\kappa_{ij}\right) = b_{ij} + d_k + \omega_{ij} + ex_j + \epsilon_{ij}$$

Trade costs are a logarithmic function of distance (d_k) a shared border effect between country i and j (b_{ij}), a tariff charged by country i to country j and an exporter fixed effect (ex_j). Tariff (ω_{ij}) represents the weighted average ad valorem tariff rate applied by country i to country j. The distance function is represented by a step function divided into 6 intervals. Intervals are in miles: [0, 375); [375, 750); [750, 1,500); [1,500, 3,000); [3,000, 6,000); and [6,000, maximum]. ϵ_{ij} reflects barriers to trade arising from all other factors and is orthogonal to the regressors. The distance and common border variables are obtained from the comprehensive geography database compiled by CEPII.

To recover technology, I follow Waugh (2010) and use the estimated trade costs, $\hat{\kappa}$, and structural parameters, $\hat{S}$, to compute the implied tradable good prices, $\hat{p}_m$, by rewriting equation 6 in terms of $\hat{S}$:

$$\hat{p}_{mi} = (AB)\left(\sum_{j=1}^{I} e^{\hat{S}_j}\left(\hat{\kappa}_{ij}\omega_{ij}\right)^{1/\theta}\right)^{-\theta}$$

From the obtained prices and the estimates $\hat{S}_i$, I get the convolution of wages and technology, $\log(w_i^{-\beta/\theta}\lambda_i)$.

[8]Details are provided in the appendix.

[9]The resulting sample consists of 130 times 129 potential observations if each country trades with all other countries. In our sample the total number of observations is 15904 implying a small number of zeros in the bilateral matrix. I conduct a robustness test where I estimate the model with the Poisson estimator proposed by Silva and Tenreyro (2006). The results are similar and are available upon request.

Table 4: Estimation Results

Summary Statistics			
Observations	TSS	SSR	R^2
7952	2,22E+05	4,08E+04	0.82

Geographical barriers			
Barrier	Paremeter estimate	Standard error	% effect on cost
[0,375)	-5,29	0,12	69,8%
[375,740)	-6,13	0,07	84,7%
[750,1500)	-7,16	0,04	104,8%
[1500,3000)	-8,19	0,03	127,1%
[3000,6000)	-9,42	0,02	156,6%
[6000,max)	-10,15	0,03	175,9%
Tariff	-0,43	0,10	6,7%
Shared border	1,21	0,11	-15,1%

Then, given the bilateral trade shares, D_{ij} and the balanced trade condition in equation 8, I follow Alvarez and Lucas (2007) and use the relationship between factor payments and total revenue to calculate equilibrium wages.[10]

$$w_i = \left(\frac{1}{(1 - s_{fi})L_i} \right) \left(\sum_{j=1}^{I} L_j w_j \frac{(1 - s_{fj})}{F_j} D_{ji} \omega_{ji} \right)$$

where s_{fi} is the share of labor in the production of final goods

$$s_{fi} = \frac{\gamma(1 - (1 - \beta)F_i)}{(1 - \gamma)\beta F_i + \gamma(1 - (1 - \beta)F_i)}$$

and F_i is the fraction of country i spending on tradable goods net of tariff expenses.

$$F_i = \sum_{j=1}^{I} D_{ji} \omega_{ji}$$

The obtained equilibrium wages together with tradable good prices, determine the implied technology levels $\hat{\lambda}$ for each country given the structural estimates of the gravity equation.

Table 4 summarizes the regression outcome of the gravity equation. In terms of fitting bilateral trade flows, I obtain an R^2 of 0.82, which is slightly lower than the R^2 of 0.83 reported by Waugh.

[10]Given factor endowments and optimal factor choice, the interest rates equals: $r_i = \alpha/(1-\alpha) w_i (L_i/K_i)$

Table 5: Simulated concentration level with exporter fixed effect

Model	Gini		Theil Exports (X)			Theil Imports (M)		
	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
Simulation	0.99	0.89	4.83	3.32	8.15	0.84	1.61	2.45
			59%	41%		34%	66%	
Data	0.98	0.91	2.60	2.13	4.73	1.10	1.61	2.71
			55%	45%		40%	60%	

The coefficients obtained for trade costs are consistent with the gravity literature, where distance and tariffs are an impediment to trade. The magnitudes of the coefficients reported in Table 4 are similar to those in Eaton and Kortum (2002) and in Waugh (2010), which consider a similar sample of countries without tariffs. The overall size of the trade costs in percentage terms are similar to those reported in Anderson and Van Wincoop (2004).

Next, I feed the model with the estimated trade costs and technology levels.[11] Table 5 shows the mean concentration levels for the simulated countries. The results show that the calibrated model replicates the fact that countries are more specialized in exports than in imports on all margins. However, the levels of export concentration are almost twice as high as the ones observed in the data: mean export (import) concentration on the extensive margin is 4.83 (0.84) compared to 2.60 (1.10) in the data. This implies that, in the simulated model, countries export (import) 0.8 (43.2) percent of the product space compared to 7.4 (33.3) percent in the data.

Figure 5 plots the corresponding cross-country pattern for simulated and empirically observed concentration levels against the log of GDP. The model replicates the empirical pattern with export concentration decreases as market size increases. However, the simulated concentration levels in the extensive margin are too high, particularly for smaller economies. Countries specialize excessively in the number of products exported. On the import side, the calibrated model is unable to replicate the L-shaped relationship between market size and concentration: indeed, the relationship does not follow any particular pattern. However, the simulated countries tend to import more goods than in the data. With regard to the intensive margin, Figures 5(e) and 5(f), the results show that, in line with the data, the model predicts no relationship between concentration and market size. To summerize, the calibrated model is able to replicate the qualitative pattern for exports but produces relatively high levels of concentration compared to the data, particularly for the extensive margin.

[11] See Table 7 at the end of the paper.

5.1 Discussion of results

A potential explanation for the excessive export concentration lies in the underlying productivity distribution. While the model reproduces the bilateral trade volumes, it fails to capture the underlying distribution of trade volumes across products. To shed light on why countries trade in too few products, I follow Haveman and Hummels (2004) and plot the empirical and simulated density of the number of exporters and importers per product.[12] Figure 6 shows the results. The simulated countries export their goods to too many destinations. The assumed productivity distribution generates producers are so efficient that even firms facing high trade costs can sell their products to numerous destinations around the world. As a consequence, the number of exporting countries per product is small. In the data (in blue) more than a third of the products are exported by 25 or more countries. In the simulation (in red) no product is exported by more than 25 countries. With regard to imports, Figure 6(b) shows that, contrary to exports, the simulated distribution of the number of countries importing a product is closely related to the empirical one.

To investigate why the model is not able to reproduce the cross country pattern of the number of products imported, we compare import expenditure and import product shares. The product share, π_i, is defined as the number of products imported, $N_{i,M}$, divided by the total number of potential goods, N, i.e. number of HS codes:

$$\pi_i = \frac{N_{i,M}}{N} \tag{16}$$

and the expenditure share, m_i, equals the total value of imports, M_i, divided by domestic absorption, Abs_i.

$$m_i = \frac{M_i}{Abs_i} \tag{17}$$

Expenditure shares are used to calibrate trade costs while the import product share defines the extensive Theil of imports. Note that the model implies that these two are equal but empirically they are not, see Figure 7.

One potential reason could be that countries differ in the number of intermediate goods used in production. When calculating the share of goods imported, I divide the total number of net products imported by the total number of HS codes, which is common to all countries. If countries do not

[12]To get the empirical distribution of the number of exporters and importers per product, I count for each HS code the number of countries that *net* export or *net* import the product. Similarly, the model implied distribution represents the number of exporters and importers for each simulated product.

make use of all tradable goods (for example they do not have the underlying technology to use a particular intermediate good), then the calculated import product share for these countries is downward biased. Ethier (1982) argues that larger economies use a higher number of intermediate goods because of increasing returns to scale in the production process of the final good. To shed light on the potential impact of market size on the number of intermediate products in the economy, we impose equality between product shares and expenditures shares, ($\pi_i = m_i$). Given this assumption, we can rewrite this equation as:

$$\frac{M_i}{N_{i,M}} = \frac{Abs_i}{N} \tag{18}$$

implying that the average per product import expenditure equals the average per product tradable expenditure. Since the number of tradable goods is the same for all countries, we expect that the elasticity of the average per product import expenditure with respect to absorption is 1. Figure 8 reveals a strong positive correlation with an $R^2 = 0.84$ and an elasticity of 0.6, significantly different from 1. Ethier's argument that larger economies have a higher degree of specialization and use a larger number of intermediate inputs in the production of tradable goods can explain why the elasticity is less than 1. In this case, the number of tradable goods would be country specific and increases with the size of the tradable sector.

Non-homothetic preferences may represent an alternative explanation for the fact that some countries spend, on average, relatively more on few imported goods. Note that, according to equation 18, the ratio of per product import expenditure to per product tradable expenditure should be one. This result relies on the assumption of homothetic preferences. Figure 9 plots the log of the ratio against the log of GDP per capita. The figure shows a negative correlation of -0.67 with an $R^2 = 0.23$. This evidence is consistent with non-homothetic preferences, where poorer countries spend more per imported good than rich ones.

6 Robustness

6.1 Alternative classification schemes

This section addresses concerns about robustness of the observed empirical concentration indexes. In particular, the level of disaggregation as well as the chosen classification scheme may affect the empirical concentration measures and the decomposition of the intensive and extensive margins. For this reason, I re-calculated the concentration indexes for both margins using (1) 5-digit SITC product codes and (2) 6-digit NAICS codes instead of the 6-digit HS codes. The advantage of the NAICS and SITC classification system is that the products are grouped according their economic function as well as their material and physical properties, rather than for tariff purposes as in the HS system. Table 6 shows the calculated concentration indexes based on SITC and NAICS classification and their correlation with respect to HS based concentration indexes. The qualitative estimates for all classification are very similar: exports are more concentrated than imports; concentration is driven by the extensive margin for exports and by the intensive margin for imports; and in terms of cross-country evidence, larger countries import and export more goods. Strikingly, the L pattern of the extensive margin also appears when the SITC and NAICS classifications are used. Differences between the various classification schemes appear in the levels of import concentration. The reason for this is that the SITC and NAICS classifications comprise a much smaller number of codes than the HS system (2,442 and 460 codes repsectively for SITC and NAICS, versus 4,529 for the HS system). Overall, however, the high correlation between the different classification standards shows that the results are robust to the classification system.

6.2 Intra-industry trade

In this section I address the discrepancy of the product space in the data and the model caused by intra-industry trade. In the main part of the paper I established a correspondence between the model and the data by netting out within product trade. This approach leaves out valuable information and may bias the results. In an alternative approach, I deal with intra-industry trade by developing a "measurement device" that enables the model to characterize intra- and inter-industry trade. The basic idea is that, in reality, the true state of the world is indeed Ricardian, i.e. varieties are in fact products, but the data are not sufficiently disaggregated to capture the true number of products. Instead, these "Ricardian products" are aggregated into sectors according to a classification scheme, i.e. HS codes. The suggested procedure converts the measurement of product units in the model to product units in the data and allows us to examine gross trade flows. Because the classification scheme is unobserved, I assume that varieties are randomly assigned an HS code following a Poisson process. Using the structure of the model, I can then estimate the Poisson parameter and characterize the "measurement device". I obtain a value of 0.94 for the Poisson parameter implying that, on average, one "Ricardian product" is equivalent to one HS product category. Based on this result, I group simulated Ricardian products randomly into artificial HS codes and calculate the implied concentration indexes. The results, presented in detail in the appendix, show that this approach produces similar results to the net trade flow approach.

Table 6: Mean concentration indexes for gross trade flows based on the Armington assumption: 130 countries

	Gini		Theil Exports (X)			Theil Imports (M)		
	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
Mean index (HS 6 digit)	0.98	0.9	1.81	2.59	4.40	3.53	2.78	6.31
% share of overall concentration			41%	59%		56%	44%	

6.3 Implications of alternative trade models

Finally, I want to compare my analysis to alternative trade models, in particular, to monopolistic competition models based on Krugman (1980), and Armington models based on Anderson and Van Wincoop (2003). The key difference with respect to the Ricardian model is that in the alternative models, tradable goods are differentiated by location of production. Applying this definition of the product space to the data implies that each country is the sole producer/exporter of an HS code and demands all country-product combinations. Hence, the number of potential goods exported is 4,529 and the number of potential goods imported is 4,529 times 129 trading partners.

Table 6 presents the corresponding concentration indexes. The results show that, contrary to the findings of my model, countries are more specialized in imports than in exports and the extensive margin drives the import concentration. The extensive Theil of imports implies that the average country imports only 17121 products (around 3 percent) out of the 4529 times 129 available products. This suggests that the empirical implications used to evaluate a model depend on the definition of a product, i.e. Armington assumption versus perfect substitutes as in the classical models of comparative advantage. While it is certainly possible to produce the results in Table 6 using a model based on the Armington assumption, the underlying mechanism to generate specialization will be very different.[13] In this paper, the analysis is based on the assumption that foreign varieties are perfect substitutes for domestic ones. One motivating observation is that the Grubel and Lloyd (1975) index of 0.19 indicates that the majority of the trade flows are inter-industry (81 percent) rather than intra-industry. However, I cannot reject these alternative hypotheses for the observed concentration patterns and would like to pursue them in future research.

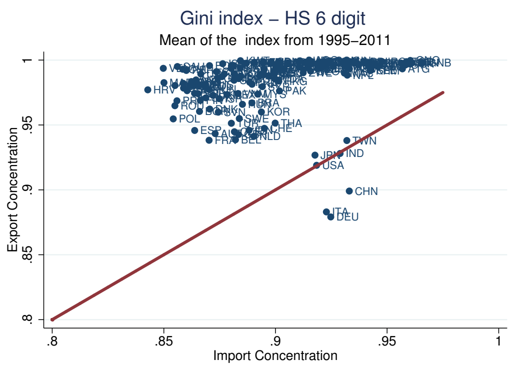
[13] For example by introducing fixed trade costs (see Romer (1994)) or declining marginal utility of varieties (see Ottaviano et al. (2002))

7 Conclusion

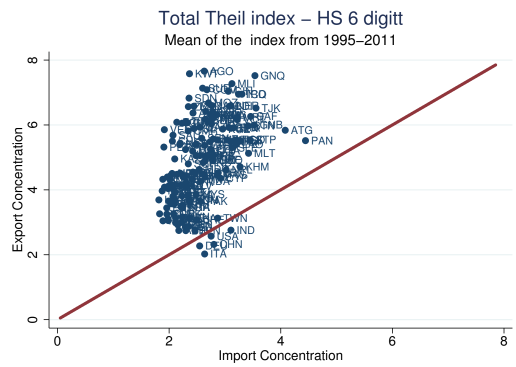
I have argued that examining export and import concentration in combination and decomposing them into a measure of extensive and intensive product margin concentration provides new quantitative and qualitative evidence on specialization patterns in world trade. Based on detailed trade data, the calculations show that exports are more concentrated than imports on all margins and that specialization is driven by the extensive product margin for exports and by the intensive product margin for imports. The extensive product margin explains the gap between export and import concentration and drives specialization differences across countries. Larger economies diversify more because they export and import more products. Furthermore, I show that the Eaton Kortum model is consistent with the observed patterns and partly replicates the stylized facts as well as the cross-country differences, both qualitatively and quantitatively. Overall, my results stress the importance of geography and absolute as well as comparative advantage in determining the pattern of specialization.

By looking through the lense of export and import concentration, this paper analyses how openness to trade changes the production structure of an economy and how these changes relate to income. My results show that the relationship between income and concentration is primarily driven by the extensive margin. This relationship has important macroeconomic policy implications. Specialization increases a country's exposure to shocks specific to the sectors on which its economy concentrates. As a result, the likelihood that product-specific shocks will have aggregate effects in terms of output volatility and/or the terms of trade increases with tarde openess. Analyzing this question is an avenue for future research.

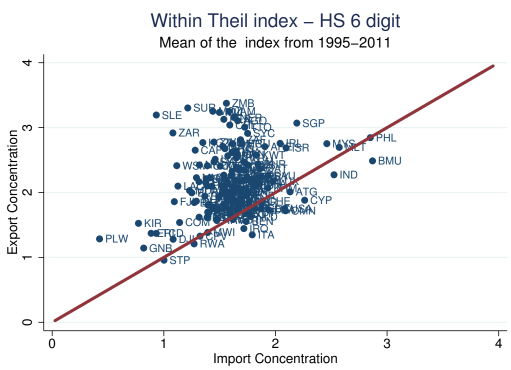
8 Figures



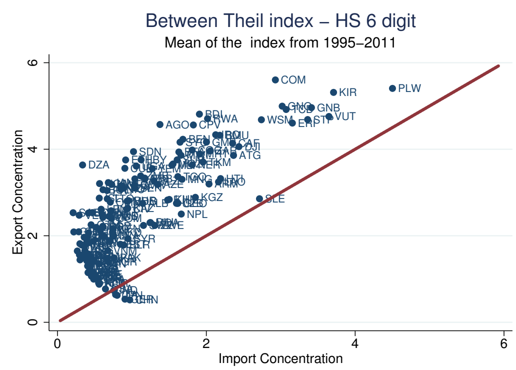
(a) Gini coefficient



(b) Theil index



(c) Intensive margin



(d) Extensive margin

Figure 1: Average export versus import concentration for the period 1995 to 2011 for 130 countries

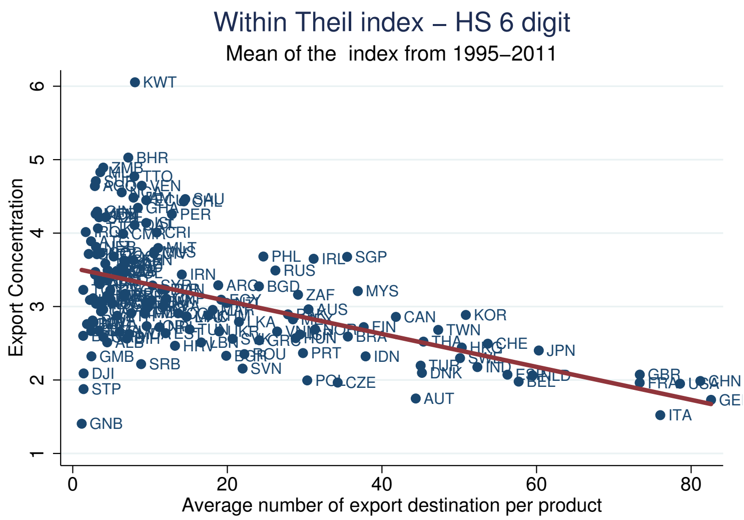
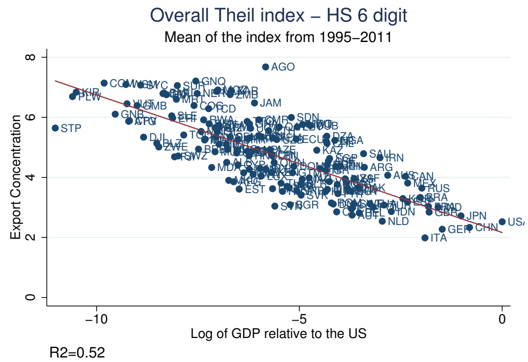
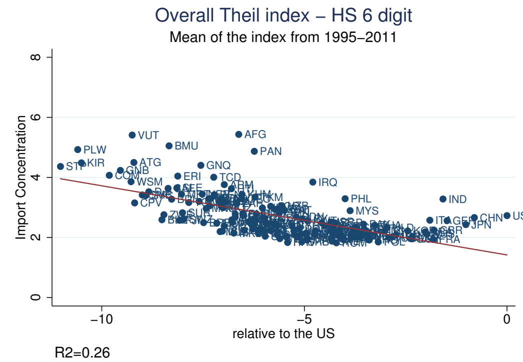


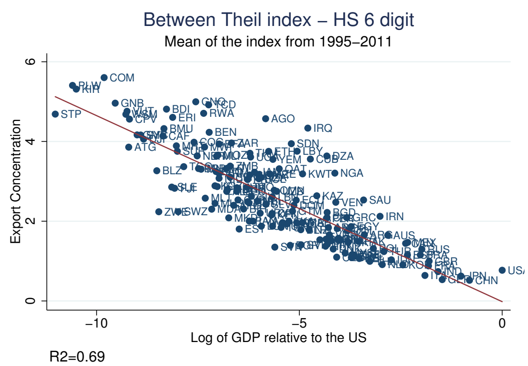
Figure 2: The average export concentration on the intensive margin versus the average number of export destinations per product for the period 1995 to 2011 for 130 countries



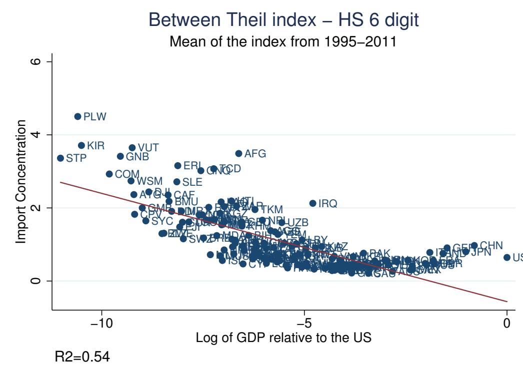
(a) Overall concentration of exports



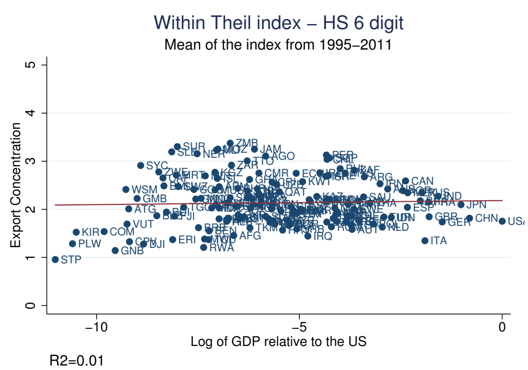
(b) Overall concentration of imports



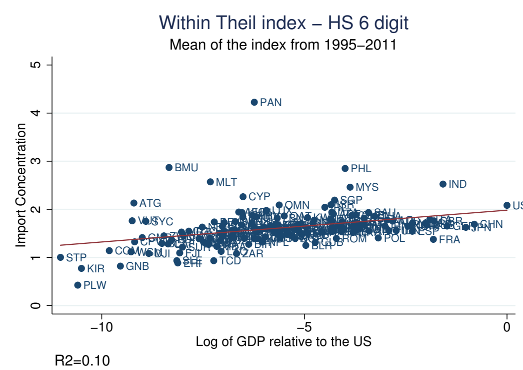
(c) Extensive margin of exports



(d) Extensive margin of imports

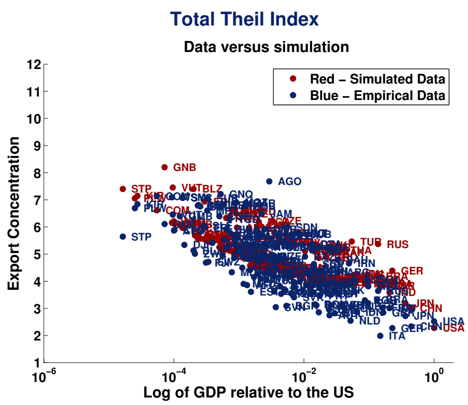


(e) Intensive margin of exports

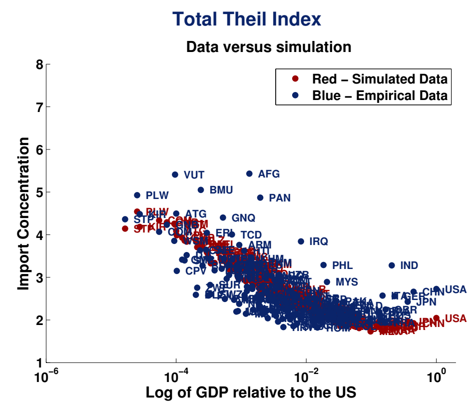


(f) Intensive margin of imports

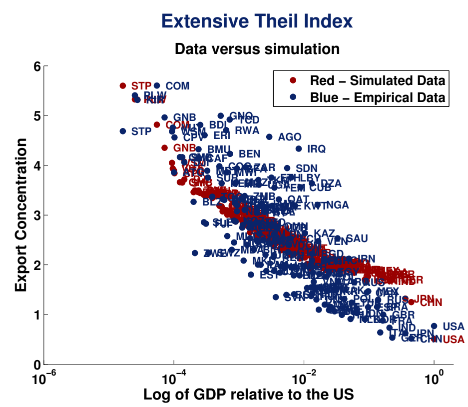
Figure 3: Average export and import concentration versus the log of average relative GDP with respect to the United States for the period 1995 to 2011 for 130 countries.



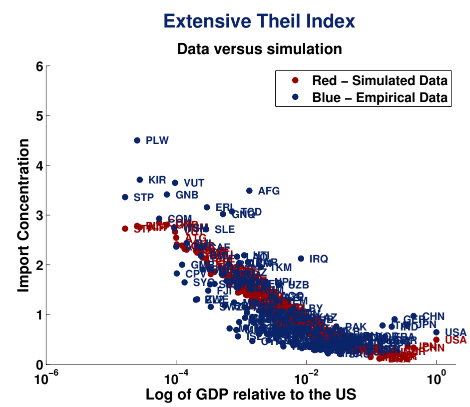
(a) Overall concentration of exports



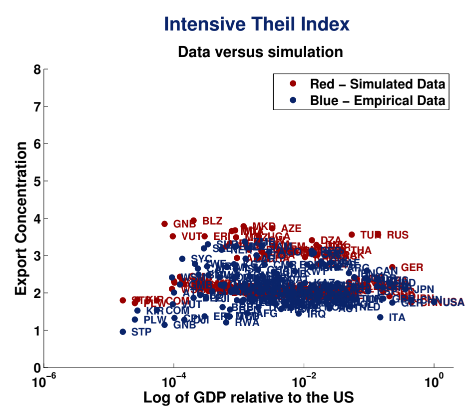
(b) Overall concentration of imports



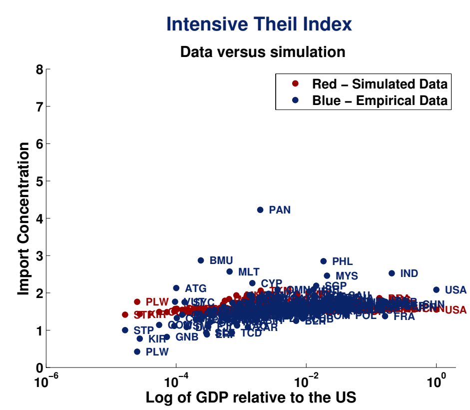
(c) Extensive margin of exports



(d) Extensive margin of imports

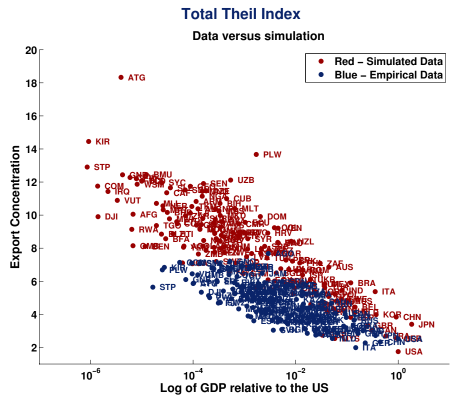


(e) Intensive margin of exports

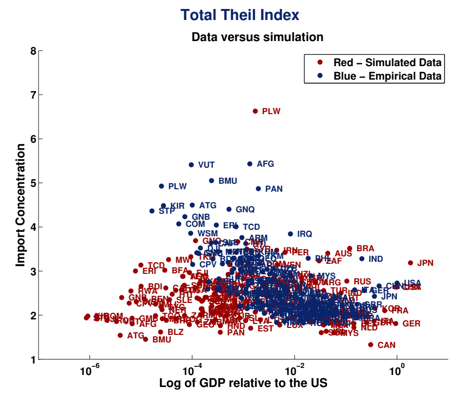


(f) Intensive margin of imports

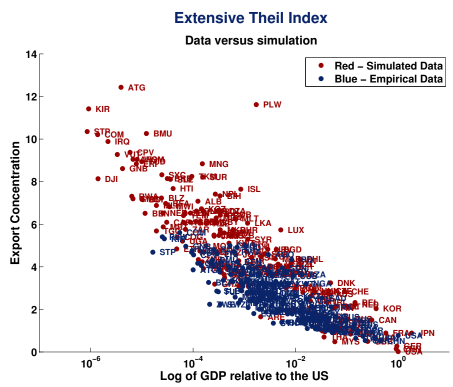
Figure 4: Simulated (in red) and empirical observed (in blue) export and import concentration versus GDP across 130 countries. The simulation uses parameterized trade costs to match the data using a country specific export cost.



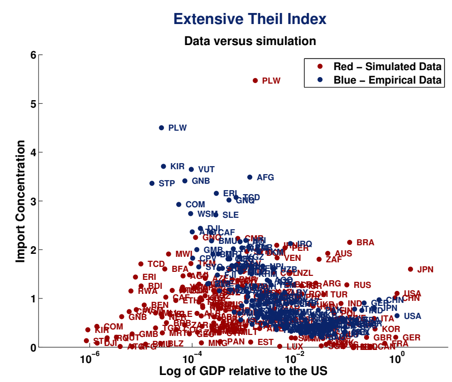
(a) Overall concentration of exports



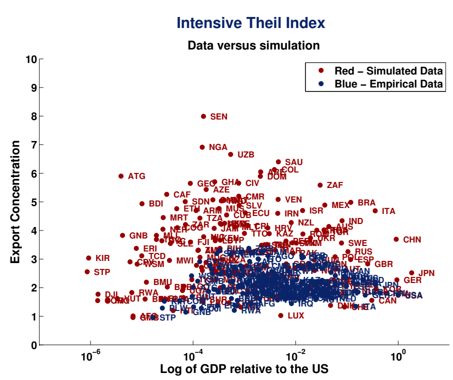
(b) Overall concentration of imports



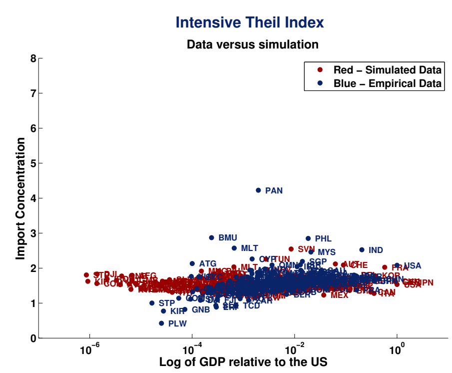
(c) Extensive margin of exports



(d) Extensive margin of imports

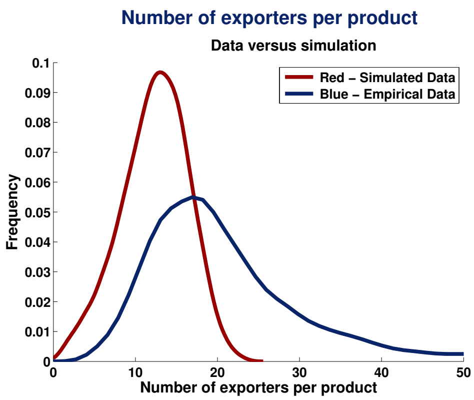


(e) Intensive margin of exports

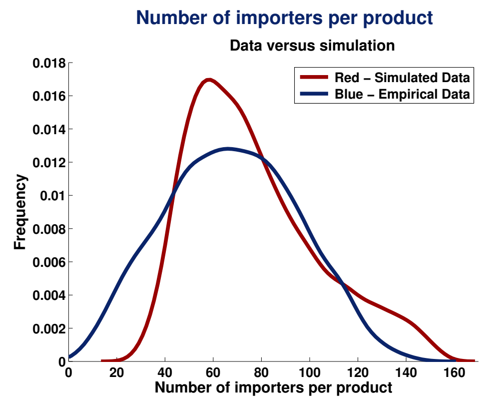


(f) Intensive margin of imports

Figure 5: Simulated (in red) and empirical observed (in blue) export and import concentration versus GDP across 130 countries. The simulation is based on estimated trade costs form bilateral trade shares including an exported fixed effect.



(a) Share of products per exporting country



(b) Share of products per importing country

Figure 6: The simulated (in red) and empirical observed (in blue) share of the number of products traded against the number of trading countries.

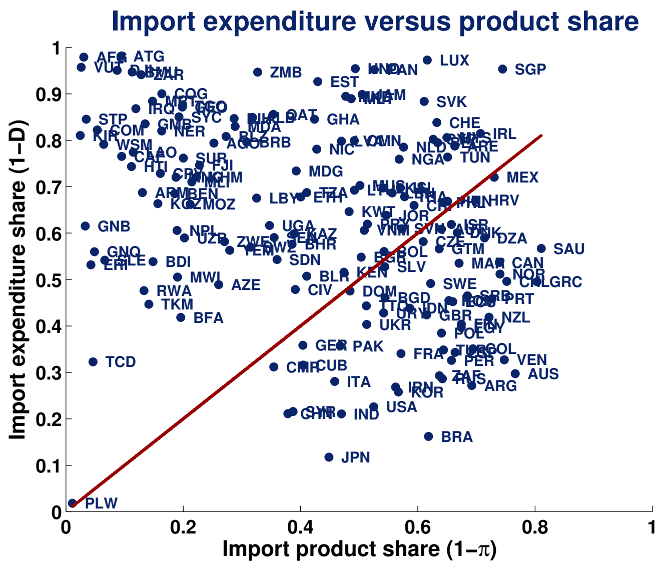


Figure 7: The import expenditure share versus the import product share.

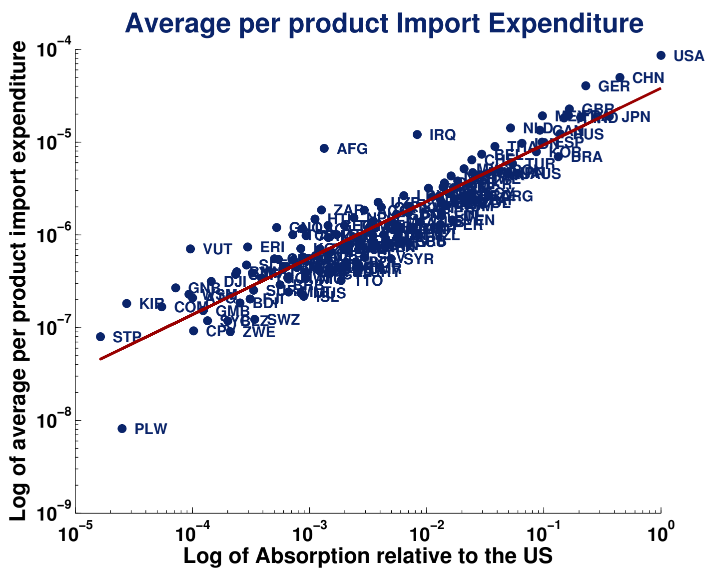


Figure 8: The log of average per product import expenditure against log of Absorption.

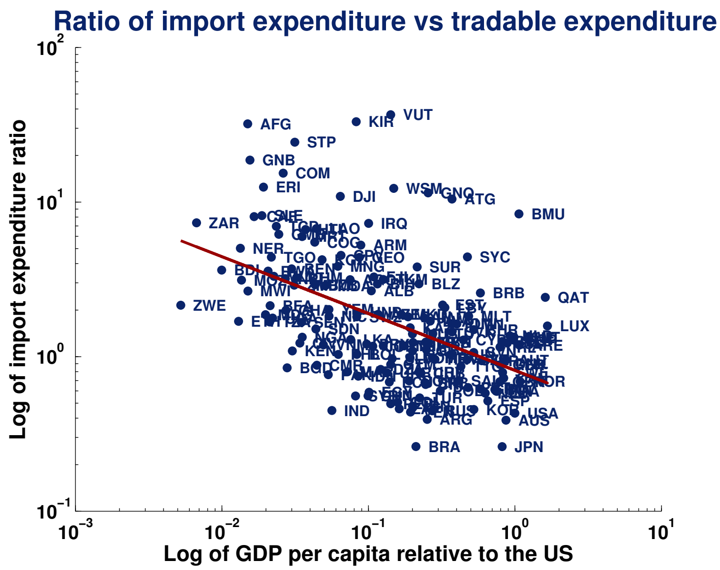


Figure 9: The log of the ratio of average per product import expenditure with respect to average per product tradable expenditure against log of GDP per capita.

9 Tables

Table 7: Country-Specific Technology and Trade Costs estimates

Country	Exporter FE	Standard error	Precent cost	Si	Standard error	$(\lambda_{US}/\lambda_i)^\theta$
USA	6.96	0.20	-50.15	-0.33	0.14	1
AGO	-2.96	0.22	34.51	-0.75	0.15	5.76
ALB	-3.84	0.21	46.86	0.02	0.15	2.7
ARG	2.19	0.20	-19.66	1.16	0.14	1.09
ARM	-3.29	0.21	38.95	-0.16	0.15	2.64
AUS	3.98	0.21	-32.87	-0.15	0.14	1.14
AUT	2.22	0.20	-19.89	0.64	0.14	0.7
AZE	-2.41	0.21	27.23	0.14	0.15	2.9
BDI	-4.38	0.22	55.01	-0.45	0.15	6.09
BEL	5.01	0.20	-39.39	-0.85	0.14	0.71
BEN	-3.62	0.21	43.64	-0.92	0.14	6.7
BFA	-3.99	0.22	48.97	-0.28	0.15	3.93
BGD	0.52	0.20	-5.09	0.18	0.14	2.25
BGR	0.40	0.20	-3.95	0.33	0.14	1.35
BHR	-0.71	0.20	7.38	0.26	0.14	1.02
BIH	-4.31	0.21	53.88	1.29	0.14	1.51
BLR	-1.18	0.21	12.54	1.44	0.15	1.04
BOL	-2.42	0.21	27.32	0.35	0.15	2.43
BRA	3.76	0.20	-31.33	0.94	0.14	1.14
BRN	-5.38	0.22	71.28	1.51	0.16	1.2
CAF	-4.92	0.21	63.48	0.49	0.15	3.52
CAN	4.52	0.20	-36.37	-0.49	0.14	0.87
CHE	3.82	0.20	-31.74	-0.16	0.14	0.63
CHL	2.16	0.20	-19.46	-0.09	0.14	1.3
CHN	5.19	0.20	-40.50	1.22	0.14	0.91
CMR	-2.34	0.21	26.35	0.38	0.14	2.49
COL	0.29	0.20	-2.84	0.55	0.14	1.92
CPV	-4.51	0.22	57.00	-0.6	0.15	3.82
CRI	0.51	0.21	-4.98	-0.5	0.14	1.44
CYP	0.29	0.20	-2.86	-0.75	0.14	1.65
CZE	1.24	0.20	-11.62	0.89	0.14	0.75
DNK	2.82	0.20	-24.58	0.24	0.14	0.71
DOM	-0.96	0.20	10.05	0.03	0.14	1.72
ECU	-0.48	0.21	4.90	0.26	0.14	2.32
EGY	0.96	0.20	-9.16	0.14	0.14	2.3
ESP	3.94	0.20	-32.54	0.12	0.14	0.96
EST	0.65	0.20	-6.30	-0.82	0.14	1.17
ETH	-1.26	0.20	13.45	-1.59	0.14	8.18
FIN	2.16	0.20	-19.41	0.94	0.14	0.59
FJI	-3.17	0.22	37.26	-0.02	0.15	1.72
FRA	4.96	0.20	-39.12	0.09	0.14	0.78
GAB	-1.88	0.21	20.67	-0.75	0.14	2.98
GBR	5.43	0.20	-41.90	-0.53	0.14	0.94
GEO	-1.78	0.21	19.43	0	0.15	2.62
GER	5.00	0.20	-39.33	0.45	0.14	0.64
GHA	0.45	0.20	-4.41	-1.58	0.14	3.66
GMB	-3.20	0.21	37.75	-1.7	0.15	7.03
GNB	-6.57	0.25	92.91	-0.14	0.17	8.85
GRC	1.07	0.20	-10.13	0.31	0.14	1.44
GTM	-0.64	0.21	6.63	0.29	0.14	1.98
HKG	6.04	0.20	-45.34	-2.06	0.14	1.12
HND	-0.93	0.21	9.76	-0.69	0.15	1.98
HRV	-1.06	0.20	11.14	0.76	0.14	1.23
HUN	0.86	0.20	-8.20	0.79	0.14	0.88
IDN	3.88	0.20	-32.18	0.2	0.14	1.41
IND	3.48	0.20	-29.38	0.92	0.14	1.64
IRL	3.19	0.20	-27.29	-0.27	0.14	0.64
IRN	-0.41	0.21	4.21	1.02	0.15	2.27
ISL	-0.87	0.21	9.07	-0.24	0.15	1
ISR	0.14	0.20	-1.41	1.41	0.14	0.79
ITA	4.36	0.20	-35.31	0.43	0.14	0.77

Table 8: Country-Specific Technology and Trade Costs estimates - cont.

Country	Exporter FE	Standard error	Precent cost	Si	Standard error	$(\lambda_{US}/\lambda_i)^\theta$
JAM	-0.73	0.21	7.54	-0.89	0.15	2.06
JOR	-1.04	0.21	10.99	0.05	0.15	1.95
JPN	5.16	0.20	-40.29	1.04	0.14	0.58
KAZ	-0.56	0.21	5.73	0.4	0.14	1.72
KEN	-0.22	0.20	2.27	-0.42	0.14	3.44
KGZ	-4.39	0.22	55.15	0.52	0.16	2.36
KNA	-4.04	0.24	49.75	-1.26	0.16	2.09
KOR	4.78	0.20	-37.99	0.6	0.14	0.62
KWT	-1.01	0.20	10.64	0.41	0.14	0.99
LAO	-3.78	0.24	45.88	0.13	0.18	2.95
LBN	0.52	0.20	-5.09	-1.47	0.14	3.18
LCA	-4.18	0.23	51.92	-1.32	0.16	2.86
LKA	2.20	0.20	-19.71	-1.34	0.15	2.78
LTU	-0.15	0.21	1.48	0.31	0.15	1.13
LVA	-0.86	0.21	9.03	0.62	0.15	1.21
MAR	0.71	0.20	-6.83	-0.03	0.14	1.86
MDA	-2.74	0.21	31.56	0.12	0.14	2.11
MDG	-1.03	0.21	10.90	-1.05	0.14	3.84
MDV	-3.89	0.24	47.55	-0.97	0.17	2.95
MEX	3.07	0.20	-26.45	0.08	0.14	1.29
MLT	0.25	0.20	-2.46	-0.66	0.14	1.07
MRT	-3.53	0.21	42.40	-0.78	0.14	4.92
MUS	0.33	0.20	-3.29	-1.17	0.14	1.75
MWI	-3.92	0.21	47.92	0.5	0.15	3.47
MYS	5.00	0.20	-39.37	-0.79	0.14	1.03
NER	-2.49	0.21	28.26	-1.23	0.15	5.3
NGA	0.20	0.20	-1.97	-1.2	0.14	5.21
NLD	5.17	0.20	-40.37	-0.82	0.14	0.76
NOR	1.90	0.20	-17.30	0.41	0.14	0.81
NPL	-3.94	0.21	48.36	0.52	0.15	3.39
NZL	3.11	0.21	-26.74	-0.19	0.14	1
OMN	0.17	0.21	-1.64	-0.64	0.14	2.11
PAK	1.94	0.20	-17.63	0.11	0.14	2.29
PAN	1.63	0.21	-15.06	-1.52	0.15	2.19
PER	0.10	0.20	-1.02	0.63	0.14	1.56
PHL	2.23	0.20	-20.02	-0.09	0.14	1.57
POL	1.49	0.20	-13.88	0.84	0.14	0.98
PRT	2.23	0.20	-19.99	-0.13	0.14	1.09
PRY	-1.54	0.21	16.70	0.67	0.15	2.05
QAT	0.22	0.21	-2.21	-0.63	0.15	1.41
ROM	0.51	0.20	-4.95	1.01	0.14	1.19
RUS	2.58	0.20	-22.73	0.81	0.14	1.29
RWA	-3.45	0.21	41.21	-1.26	0.15	8.02
SAU	1.98	0.20	-17.94	-0.35	0.14	1.7
SEN	-1.52	0.21	16.37	-0.75	0.14	3.62
SGP	5.06	0.20	-39.73	-1	0.14	0.68
SLV	-1.52	0.21	16.43	-0.21	0.15	1.87
STP	-4.17	0.24	51.75	-1.36	0.17	4.87
SVK	0.75	0.20	-7.20	0.32	0.14	0.93
SVN	0.16	0.20	-1.59	0.22	0.14	0.84
SWE	2.98	0.20	-25.78	0.74	0.14	0.61
SYR	-1.24	0.21	13.22	0.14	0.15	2.85
TCD	-6.43	0.23	90.15	0.75	0.16	3.97
THA	4.10	0.20	-33.63	0.19	0.14	1.08
TJK	-5.25	0.24	69.11	1.56	0.17	2.22
TKM	-6.02	0.24	82.62	2.66	0.17	1.41
TTO	-1.09	0.21	11.53	-0.03	0.14	1.12
TUN	-0.30	0.20	3.07	0.2	0.14	1.41
TUR	2.20	0.20	-19.76	0.88	0.14	1.13
TZA	-1.21	0.20	12.84	-0.78	0.14	4.26
UGA	-2.60	0.21	29.69	-0.22	0.14	5.04
UKR	1.17	0.20	-11.02	1.04	0.14	1.4
URY	0.92	0.21	-8.79	-0.21	0.15	1.46
UZB	-4.30	0.23	53.70	1.95	0.17	1.9
VEN	-0.55	0.21	5.69	1.01	0.14	1.44
VNM	3.10	0.20	-26.65	-0.58	0.14	1.96
YEM	-3.13	0.21	36.79	0.26	0.15	3.77
ZAF	3.86	0.20	-32.01	-0.44	0.14	1.32
ZAR	-3.92	0.22	48.02	0.68	0.16	3.55

References

ALVAREZ, F. AND R. J. LUCAS (2007): "General equilibrium analysis of the Eaton Kortum model of international trade," *Journal of Monetary Economics*, 54(6), 1726–1768.

ANDERSON, J. E. AND E. VAN WINCOOP (2003): "Gravity with gravitas: a solution to the border puzzle," *American Economic Review*, 93, 170–192.

——— (2004): "Trade costs," *Journal of Economic Literature*, 42, 691–751.

BERNARD, A. B., J. EATON, J. B. JENSON, AND S. KORTUM (2003): "Plants and productivity in international trade," *American Economic Review*, 93, 1268–1290.

BRODA, C. AND D. WEINSTEIN (2006): "Globalization and the Gains from Variety," *Quaterly Journal of Economics*, 121, 541–585.

CADOT, O., C. CARRÈÈRE, AND V. STRAUSS-KAHN (2011): "Export Diversification: What's behind the Hump?" *The Review of Economics and Statistics*, 93, 590–605.

——— (2014): "OECD imports: diversification of suppliers and quality search," *Review of World Economics*, 150, 1–24.

CHOR, D. (2010): "Unpacking sources of comparative advantage: A quantitative approach," *Journal of International Economics*, 82, 152–167.

COSTINOT, A., D. DONALDSON, AND I. KOMUNJER (2012): "What goods do countries trade? A quantitative exploration of Ricardo's ideas," *Review of Economic Studies*, 79, 581–608.

DI GIOVANNI, J. AND A. A. LEVCHENKO (2012): "Country Size, International Trade and Aggregate Fluctuations in Granular Economies," *Journal of Political Economy*, 120, 1083–1132.

EATON, J. AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An anatomy of international trade: Evidence from French firms," *Econometrica*, 79, 1453–1498.

ETHIER, W. J. (1982): "National and international returns to scale in the modern theory of international trade," *The American Economic Review*, 389–405.

FEENSTRA, R. C., R. E. LIPSEY, AND H. P. BOWEN (1997): "World Trade Flows, 1970–1992, with Production and Tariff Data," *National Bureau of Economic Research Working Paper*, 5910.

GAULIER, G. AND S. ZIGNAGO (2009): "BACI: International trade database at the product-level," .

GRUBEL, H. G. AND P. J. LLOYD (1975): *Intra-industry trade: The theory and measurement of international trade in differentiated products*, vol. 12, Macmillan London.

HAVEMAN, J. AND D. HUMMELS (2004): "Alternative hypotheses and the volume of trade: the gravity equation and the extent of specialization," *Canadian Journal of Economics/Revue canadienne d'économique*, 37, 199–218.

HUMMELS, D. AND P. KLENOW (2005): "The Variety and Quality of a Nations Exports," *American Economic Review*, 95, 704–23.

IMBS, J. AND R. WACZIARG (2003): "Stages of Diversification," *American Economic Review*, 93, 63–86.

JAIMOVICH, E. (2012): "Import diversification along the growth path," *Economics Letters*, 117, 306–310.

KOREN, M. AND S. TENREYRO (2007): "Volatility and development," *The Quarterly Journal of Economics*, 122, 243–287.

——— (2013): "Technological diversification," *The American Economic Review*, 103, 378–414.

KRUGMAN, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 16, 670–675.

LEVCHENKO, A. A. AND J. ZHANG (2011): "The evolution of comparative advantage: Measurement and welfare implications," Tech. rep., National Bureau of Economic Research.

OTTAVIANO, G. I. P., T. TABUCHI, AND J. F. THISSE (2002): "Agglomeration and Trade Revisited," *International Economic Review*, 43(2), 409–436.

ROMER, P. (1994): "New goods, old theory, and the welfare costs of trade restrictions," *Journal of Development Economics*, 43, 5–38.

SCHOTT, P. K. (2004): "Across-product versus within-product specialization in international trade," *The Quarterly Journal of Economics*, 119, 647–678.

SHIKHER, S. (2011): "Capital, technology, and specialization in the neoclassical model," *Journal of International Economics*, 83, 229–242.

SILVA, J. AND S. TENREYRO (2006): "The log of gravity," *The Review of Economics and Statistics*, 88, 641–658.

SIMONOVSKA, I. AND M. WAUGH (2011): "The Elasticity of Trade: Estimates and Evidence," Tech. rep., National Bureau of Economic Research.

WAUGH, M. E. (2010): "International trade and income differences," *American Economic Review*, 100, 2093–2124.

Appendix Chapter 1

The distribution of import expenditure

Note, that the set of imported goods is defined as the sum of all the fractions of goods imported from all other countries in the world . $D_{n,k}$ describes the fraction of goods country n imports from country k. Assuming that a country imports a good from only 1 source country (so the sets of products imported from different countries are mutually exclusive), we can sum add together all the fractions of goods imported across all trading partners in order to obtain the import probability.

$$\Pr(imp) = \sum_{k \neq n}^{I} D_{nk}$$

The corresponding distribution function for prices (p) is given by:

$$M_n(p) = \frac{\sum_{k\neq n}^{I} \int_0^p \prod_{s\neq n}^{N} (1 - G_{ns}(q)) dG_{nk}(q)}{\sum_{k\neq n}^{I} D_{nk}}$$

$$M_n(p) = \frac{\sum_{k\neq n}^{I} D_{nk} \int_0^p \varphi_n \frac{1}{\theta} q^{\frac{1}{\theta}-1} \left(e^{-\varphi_n q^{\frac{1}{\theta}}} \right) dq}{\sum_{k\neq n}^{I} D_{nk}}$$

Define $u = \sum_{s=1}^{N} \lambda_s (\frac{w_s^{\beta} p_{ms}^{1-\beta}}{\kappa_{ns}\omega_{ns}})^{-\frac{1}{\theta}} q^{\frac{1}{\theta}}$ and $du = \varphi_n \frac{1}{\theta} q^{\frac{1}{\theta}-1} dq = \frac{1}{\theta} \frac{u}{q} dq$, we get:

$$M_n(p) = \int_0^p (e^{-u}) du$$

Hence, the import price distribution is independent of the source country:

$$M_n(p) = 1 - e^{-\varphi_n p^{\frac{1}{\theta}}} = F_n(p)$$

Using the import price distribution, we can derive the distribution of import expenditure by using the following transformation. Note, that the import expenditure of country n on good x in the case of CES preferences is given by:

$$q_n(x) p_n(x) = \left(\left(\min_{i \neq n} p_i(x) \right)^{1-\eta} p_{mn}^{\eta} q_n \right) = \left(\min_{i\neq n} B \frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}} x_i^{\theta}) \right)^{1-\eta} p_{mn}^{\eta} q_n \tag{19}$$

and the probability that it will import at price p is given by $M_n(p)$. Hence, we can write the distribution function for import expenditure at price p as

$$E_n(p) = 1 - e^{-\varphi_n k_n^{\frac{1}{(1-\eta)}} (p)^{\frac{1}{\theta(1-\eta)}}}$$

where $k_n = \left(p_{mn}^{\eta} q_n\right)$ is a constant. The corresponding Fréchet pdf is

$$e(p) = \frac{1}{\theta(1-\eta)} \varphi_n \left(\frac{p}{k_n}\right)^{\frac{1}{\theta(1-\eta)}} \left(\frac{1}{p}\right) e^{-\varphi_n \left(\frac{p}{k_n}\right)^{\frac{1}{\theta(1-\eta)}}}$$

with location parameter $s_n = {k_n}^{-1} \varphi_n^{\theta(1-\eta)}$ and shape parameter $\alpha = \frac{-1}{\theta(1-\eta)}$.[14]

Given that the price distribution is independent of the source country and follows a Fréchet, we can calculate the corresponding concentration indexes analytically. The intensive margin Theil index for country n can be written approximately in terms of the continuous probability density function:

$$T_n^W = \left[\frac{1}{N_a} \sum_{k \in G_a} \left(\frac{R_k}{\bar{R}_a}\right) \ln \left(\frac{R_k}{\bar{R}_a}\right)\right] \approx \int_0^{\infty} \left(\frac{R}{\bar{R}_a}\right) \ln \left(\frac{R}{\bar{R}_a}\right) f(R) dR$$

Plugging the location parameter $s_n = {k_n}^{-1} \varphi_n^{\theta(1-\eta)}$ and shape parameter $\alpha = \frac{-1}{\theta(1-\eta)}$ into the density of the Fréchet distribution, we get:

$$T_n^W = \int_0^{\infty} \left(\frac{R}{\bar{R}_a}\right) \ln \left(\frac{R}{\bar{R}_a}\right) \frac{\alpha}{s_n} \left(\frac{R}{s_n}\right)^{-\alpha-1} e^{-(R/s_n)^{-\alpha}} dR$$

where $\bar{R}_a$ is the mean import expenditure. Solving the integral, gives:

$$T_n^W = \frac{1}{\Gamma(1-\frac{1}{\alpha})} \left(\int_0^{\infty} \ln \left(u^{(-1/\alpha)}\right) u^{(-1/\alpha)} e^{-u} du\right) - \ln \left(\Gamma(1 - \frac{1}{\alpha})\right) \tag{20}$$

The Theil index for the intensive margin of imports does not depend on the scale parameter s_n and is thus identical across countries. The index is completely determined by the shape parameter $\alpha = -1/(\theta(\eta - 1))$ and depends only on the elasticity of substitution η and the degree of comparative advantage θ. The integral in equation 20 cannot be solved analytically. To compute the exact Theil index implied by the shape parameter α, I approximate the integral numerically via the Gauss Laguerre procedure.

[14]The generic form of the Fréchet probability density function is: $f(x) = \frac{\alpha}{s} \left(\frac{x}{s}\right)^{-1-\alpha} e^{-(x/s)^{-\alpha}}$.

The share of products exported

The extensive margin concentration index for exports is given by the number of products exported to any destination divided by the total number of products in the world. The Law of Large Numbers implies that the probability that a country will probability a good is equal to the share of products exported. I define the set of products that country n exports as the union of the set of products exported to each destination j, $U_{ex} = \left| \cup_{j \neq n}^{I} \mathbf{B}_{jn} \right|$. Note that the set of products exported to destination j overlaps with the set of products exported to destination k. The total number of products exported to both countries is the sum of the two sets minus the number of products that are in both, i.e. $\left| \mathbf{B}_{jn} \cup \mathbf{B}_{kn} \right| = \mathbf{B}_{jn} + \mathbf{B}_{kn} - \mathbf{B}_{jn} \cap \mathbf{B}_{kn}$. Generalizing this expression to all possible destinations implies that the share of products country n exports follows the Inclusion and Exclusion Principle and is given by

$$
\begin{aligned}
U_{ex} &= \sum_{i=1}^{I-1} (A_i) - \sum_{i,j:1 \leq i < j \leq I-1} \left(A_i \cap A_j \right) + \\
&+ \sum_{i,j,k:1 \leq i < j \leq I-1} \left(A_i \cap A_j \cap A_k \right) - \cdots + (-1)^{I-2} \left(A_1 \cap \cdots \cap A_{I-1} \right)
\end{aligned}
$$

where A_i defines the set of products exported to destination i, i.e. A_i contains all products for which country i obtains the minimum price in country n, $A_i = p_{n,i}(x) \leq \min_{j \neq n} \left[p_{n,j}(x) \right]$. If we denote the intersection of all A_i's with an index L

$$
A_L := \bigcap_{i \in L} (A_i)
$$

we can rewrite the set of products exported in a more compact form

$$
U_{ex} = \sum_{i=1}^{I-1} (-1)^{k-1} \sum_{L \subset \{1,\ldots,I-1\}, |L| = k} (A_L)
$$

The above equation sums up all subsets L of the indexes $1, ..., I-1$ where k describes the number of destinations a product is exported to.

In the special case of symmetric countries, the number of products exported to k destinations is the same for all destinations and the intersection A_L only depends on the cardinality of L. As a result, we can express the set of products exported to k destinations as the $L = k$ intersections of A,

$$a_k = (A_L)$$

and the set of products exported can be simplified to

$$U_{ex} = \sum_{k=1}^{I-1} (-1)^{k-1} \begin{pmatrix} I \\ k \end{pmatrix} a_k$$

where $I - 1$ is the total number of destinations. The resulting share of products exported can easily be calculated. For example, the share of products country n exports to one particular destination j is given by $a_1 = D_{jn}$ and is equal to

$$a_1 = \frac{(\kappa)^{1/\theta}}{(1 + (I-1)(\kappa)^{1/\theta})}$$

The share of products country n exports to any pair of destinations j and k is given by the probability of obtaining the minimum price in those two destinations.

$$a_2 = \Pr \left\{ \left(p_j(x) = p_{ij}(x) \leq \min_{l \neq i} \left[p_{lj}(x) \right] \right) \wedge \left(p_k(x) = p_{ik}(x) \leq \min_{l \neq i} \left[p_{lk}(x) \right] \right) \right\}$$

where the price of good x offered by country n in destination j is $p_{n,j}(x) = B \left[\kappa x_n^{\theta} \right] \forall j \neq n$ and the price offered at home is $p_{n,n} = B x_n^{\theta}$. Note that, the only factor differentiating between prices is whether the country sells in its home market or in a foreign market. Since $\kappa < 1$ implies that $p_{jj} < p_{jl}$ and $p_{kk} < p_{kl}$ $\forall j, k \neq l$, we can write the set of products exported to destinations j and k as the corresponding probability of obtaining the minimum price in the respective destinations

$$a_2 = B \Pr \left\{ \kappa x_i^{\theta} \leq \min_{l \neq i,j,k,} \left[x_j^{\theta}, x_k^{\theta}, \kappa x_l^{\theta} \right] \right\}$$

If we apply the properties of the exponential distribution, this probability is equal to

$$a_2 = \frac{(\kappa)^{1/\theta}}{2 + (I-2)(\kappa)^{1/\theta}}$$

In similar manner, we obtain the probability to export (and hence the share of products exported) to k destinations

$$a_k = \frac{(\kappa)^{1/\theta}}{k + (I-k)(\kappa)^{1/\theta}}$$

As a result, we can write the extensive Theil for exports as the inverse of the log of the share of products exported to any destination as

$$T^{ext}_{i,X} = -\ln \left(\sum_{i=1}^{I} (-1)^{k-1} \binom{I}{k} a_k \right)$$

where

$$a_k = \frac{(\kappa)^{1/\theta}}{k + (I-k)(\kappa)^{1/\theta}}$$

Trade data

To build my empirical evidence, I use the BACI data set based on the Comtrade data set collected by the United Nations. I choose the 1992 6-digit HS product classification scheme as the preferred level of disaggregation. I assume that the tradable goods sector corresponds to the manufacturing sector. Using a correspondance table provided by Feenstra et al. (1997), I identify 4,529 tradable manufacturing products. I construct trade shares D following Bernard et al. (2003) and Waugh (2010) in the following manner:

$$D_{i,j} = \frac{\text{Imports}_{i,j}}{\text{Gross Mfg. Production}_i \text{ - Exports}_i + \text{Imports}_i}$$

The numerator consists in the aggregate value of manufactured goods that country i imports from country j. These data are obtained directly from BACI. The denominator is gross manufacturing production minus total manufactured exports plus manufactured imports (against all countries in the sample), see Eaton and Kortum (2002). Basically, this simply computes the share of expenditure by dividing the value of inputs consumed by country i and imported from country j divided by the total value of inputs in country i. Gross manufacturing data are from either UNIDO (2012) or imputed from value added data obtained from the UN National accounts.

Alternative classification schemes (5 digit SITC and 6 digit NAICS)

In this paper I analyzed total net trade flows for the 6-digit HS industry classification. This section shows that the results obtained in the main part of the paper are influenced by the industry classification scheme and in fact apply in a more general sense. As a robustness check, I use the 5-digit SITC and 6-digit NAICS classification schemes. The total number of SITC products is 2,442 and the total number of NAICS products is 460.

Table 9: Mean concentration indexes for net trade flows based on the 5-digit SITC and 6-digit NAICS industry classifications: 130 countries

	Gini		Theil Exports (X)			Theil Imports (M)		
	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
Mean index (SITC 5 digit)	0.98	0.88	2.34	2.29	4.64	0.70	1.93	2.63
% share of overall concentration			51%	49%		27%	73%	
Correlation to HS	0.95	0.88	0.94	0.91	0.95	0.85	0.93	0.90
Mean index (NAICS 6 digit)	0.97	0.82	1.98	1.75	3.73	0.52	1.23	1.75
% share of overall concentration			53%	47%		29%	71%	
Correlation to HS	0.94	0.89	0.95	0.86	0.91	0.53	0.79	0.66

Table 9 shows the descriptive statistics for the SITC and the NAICS samples. The qualitative estimates for SITC and NAICS classifications are very similar to those of the HS classification. Exports are more concentrated than imports. Concentration is driven by the extensive margin for exports and the intensive margin for imports. Also, in terms of cross-country evidence, larger countries import and export more goods. Strikingly, the L pattern of the extensive margin also appears when the SITC and NAICS classifications are used.

Poisson parameter approach

The data contain intra-industry trade whereas the model is a pure Ricardian model. In this section I outline an alternative approach that converts the measurement of product units in the model to the product units in the data. Suppose that the true level of disaggregation of Ricardian products, as defined in the Eaton and Kortum (2002) (EK) model, is unobserved and the classification scheme measures only an aggregate of those Ricardian products. For example, when products, in the sense of the EK model, arrive at the border, customs agents aggregate those products into an industry according to the HS classification standard. The number of EK products that customs agents assign to an HS industry classification is unobserved to the researcher. Given this interpretation, I model the classification process as a randomization device following a Poisson process with parameter μ.

The parameter μ informs on how many EK Ricardian products, on average, make up one HS code (the observed product unit in the data).

To estimate the Poisson parameter, I proceed as follows. According to the law of large numbers, the probability of importing a particular EK product equals the share of the number of EK products imported with respect to the total number of EK products. In the model, the probability that an EK product is imported equals $P(imp_{EKprod}) = 1 - D_{ii}$, where D_{ii} is the probability of not importing an EK product. The probability of not importing any EK product within an HS code is D_{ii}^{μ}, where μ is the average number of products that comprise an HS code. As a result, we get the probability of importing an HS code (product unit in data), which corresponds to one minus the probability of not importing any EK products in that industry, $P(imp_{HScode}) = 1 - D_{ii}^{\mu}$. Since the probability of importing a product equals the share of products imported, N_M/N, we can use the definition of the Theil index on the extensive margin,

$$T_i^{bm} = -\ln(N_M/N) = -\ln\left(1 - D_{ii}^{\mu}\right)$$

to obtain μ:

$$\mu_i = \quad \frac{\ln\left(1 - \exp(-T_i^{bm})\right)}{\ln(D_{ii})}\Bigg)$$

We compute the Poisson parameter for each country and take the average value as our estimate of $\hat{\mu} = 1/I \sum_{i=1}^{I} \mu_i$. The results imply that, on average, $\hat{\mu} = 0.94$ EK products correspond to an HS code. code.

Empirical evidence and simulation results

In my simulation the total number of intermediate goods (N) is the product of the 4,529 industries in the data times 0.94, the average number of products in an industry, $N = 4,258$. One advantage of the this approach is that we can make use of the full data sample and do not lose 35 percent of trade flows when converting the data into net trade flows. Next, I present the empirical results for the full sample together with the corresponding simulation results that replicate the data.

Figure **??** shows that, in general, the pattern of export and import concentration in the full sample is similar to the pattern in the net trade flow sample. Exports are more concentrated than imports for

Table 10: Mean concentration indexes for gross trade flows over 2210 country-year pairs

	Gini		Theil Exports (X)			Theil Imports (M)		
	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
Level of concentration	0.96	0.89	1.81	2.59	4.40	0.94	1.76	2.70
% share of total concentration			41%	59%		34%	66%	
Correlation with net trade	0.98	0.87	0.96	0.45	0.87	0.98	0.82	0.90

almost all countries and on all margins. The pattern on the extensive margin is displayed in Figure **??**. Figure **??** shows the patterns on the intensive margin.

Regarding the quantitative differences, we observe that the overall level of concentration decreases with respect to the net trade flow sample for both exports and imports. The decomposition reveals that the effects differ across the margins. In the case of the extensive margin, concentration decreases with respect to the net trade flow sample whereas intensive margin concentration increases, thus reversing the relative importance of each margin in terms of overall export concentration. Intra-industry trade increases the number of products traded and the sales value of those products. As a result, we observe a lower (higher) concentration index on the extensive (intensive) margins. The overall concentration index is primarily driven by the intensive margin with a share of 59% for exports and 66% for imports (see Table 11).

Table 11: Simulated concentration level with Poisson parameter $\mu = 0.94$ and exporter fixed effect

	Gini		Theil Exports (X)			Theil Imports (M)		
Model	Exports	Imports	Extensive Margin	Intensive Margin	Total	Extensive Margin	Intensive Margin	Total
Simulation	0.99	0.89	4.97	3.32	8.29	1.15	1.76	2.91
			60%	40%		39%	61%	
Data (gross trade flows)	0.96	0.89	1.81	2.59	4.40	0.94	1.76	2.70
			41%	59%		34%	66%	
Data (net trade flows)	0.98	0.91	2.60	2.13	4.73	1.10	1.61	2.71

Using data on gross trade flows, I re-estimate trade cost and technology parameters. I then simulate the model, calculate the resulting concentration indexes using the Poisson measurement device and compare the simulated results with the data. Table 11 shows the results. Export concentration is higher than import concentration on all margins. With respect to decomposition, in a similar way to the net trade case, the extensive margin dominates overall concentration for exports and the intensive margin dominates for imports. The concentration levels obtained for imports are close to the one in the data. In the case of exports, simulated concentration levels are too high. In terms of the cross-country concentration pattern, the calibrated model fits the data well, particularly for exports. However, like the net trade results, the model cannot capture the negative relationship between import concentration and GDP.

Documents de Travail

530. L. Kaas, P. A. Pintus and S. Ray, "Land Collateral and Labor Market Dynamics in France," December 2014

531. S. Gabrieli and C.-P. Georg, "A network view on interbank market freezes," December 2014

532. A. Bozio, D. Irac and L. Py, "Impact of research tax credit on R&D and innovation: evidence from the 2008 French reform," December 2014

533. L. Clerc, A. Derviz, C. Mendicino, S. Moyen, K. Nikolov, L. Stracca, J. Suarez, A. P. Vardoulakis, "Capital Regulation in a Macroeconomic Model with Three Layers of Default," December 2014

534. M. Debbich, "Why Financial Advice Cannot Substitute for Financial Literacy?," January 2015

535. E. Challe, J. Matheron, X. Ragot and J. F. Rubio-Ramirez, "Precautionary Saving and Aggregate Demand," January 2015

536. K. Galuščák and G. Kátay, "Labour Force Participation and Tax-Benefit Systems: A Cross-Country Comparative Perspective," February 2015

537. Patrick Fève and Jean-Guillaume Sahuc, "On the Size of the Government Spending Multiplier in the Euro Area," February 2015

538. N. Berardi and F. Marzo, "The Elasticity of Poverty with respect to Sectoral Growth in Africa," February 2015

539. S. Guilloux-Nefussi, "Globalization, Market Structure and the Flattening of the Phillips Curve," March 2015

540. P. Andrade, C. Cahn, H. Fraisse and J-S. Mésonnier, "Can the Provision of Long-Term Liquidity Help to Avoid a Credit Crunch? Evidence from the Eurosystem's LTROs," March 2015

541. V. Bignon, R. Breton and M. Rojas Breu, "Monetary Union with A Single Currency and Imperfect Credit Market Integration," March 2015

542. W. Steingress, "Specialization Patterns in International Trade," March 2015

Pour accéder à la liste complète des Documents de Travail publiés par la Banque de France veuillez consulter le site : www.banque-france.fr

For a complete list of Working Papers published by the Banque de France, please visit the website: www.banque-france.fr

Pour tous commentaires ou demandes sur les Documents de Travail, contacter la bibliothèque de la Direction Générale des Études et des Relations Internationales à l'adresse suivante :

For any comment or enquiries on the Working Papers, contact the library of the Directorate General Economics and International Relations at the following address :

BANQUE DE FRANCE
49- 1404 Labolog
75049 Paris Cedex 01
tél : 0033 (0)1 42 97 77 24 ou 01 42 92 63 40 ou 48 90 ou 69 81
email : 1404-ut@banque-france.fr