Université de Montréal

Essays in International Trade

par
Walter Steingress

Département de sciences économiques
Faculté des arts et sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences économiques

Decembre, 2014

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :
Essays in International Trade

présentée par :
Walter Steingress

a été évaluée par un jury composé des personnes suivantes :

Baris Kaymak	président-rapporteur
Rui Castro	directeur de recherche
Andriana Bellou	directrice de recherche
Alessandro Barattieri	examinateur externe
Michael Hubermann	vice-doyen, représentant du doyen de la FAS

Thèse acceptée le 5 decembre 2014

à mes parents, Walter et Gudrun Steingress

Acknowledgements

Je tiens en premier à remercier mes directeurs de thèse Rui Castro et Andriana Bellou pour leur présence continue, leurs amabilité, leurs conseils de tous genres, pour m'avoir encouragé dans toutes les étapes de cette thèse et d'avoir surtout cru en moi. Ils etait sans doute les personnes clés et les plus déterminantes dans l'élaboration de cette thèse. Merci encore Andriana et Rui.

Mes remerciements vont également à Baris Kaymak qui me suit depuis la première présentation orale de cette thèse. J'aimerais tout aussi remercier Stefania Garetto pour tout l'appui qu'elle m'a apportée lors de mes trois dernières années de thèse. Grazie mille Stefania. J'ai également bénéficié des nombreux commentaires et suggestions de plusieurs chercheurs et professeurs qui m'ont permis d'améliorer ma recherche. Merci: Francisco Alvarez-Cuadrado, Costas Arkolakis, Alessandro Barattieri, Samuel Bazzi, Kristian Behrens, Emanuela Cardia, Jonathan Eaton, Michael Huberman, Joseph Kaboski, Raja Kali, Miklos Koren, Anna Maria Mayda, Markus Poschke, Francisco Ruge-Murcia, Michael Siemer, Sebastian Stumpner, Ari Van Assche, Hylke Vandenbussche, Michael Waugh, et Pierre-Yves Yanni.

Je tiens à remercier l'ensemble des étudiants de Doctorat en particulier mes collègues du bureau Ulrich, Modeste, Romuald, Firmé, Selma, Idrissa, Jonathan et aussi mes autres collègues Juste, Max, Vincent De Paul, Guy, Brahim, Catherine, Louis-Philippe, Didier, William, Hector et Ismael avec qui j'ai passé des moments inoubliables, particulièrement ces discussions dans les couloirs du département qui ont permis de raffermir mes connaissances en économie.

Ce parcours fut difficile d'un point de vue intellectuel, mais aussi moral, à des milliers de kilomètres des belles montagnes du Tirol en Autriche; je devais faire face à ce rude hiver de Montréal. J'ai trouvé sur place une grande "famille" qui m'a sans cesse soutenue dans toutes les épreuves et je ne pourrai terminer sans citer quelques noms.

Je remercie tout particulièrement mes très chers amis Rozie, Pierre-Yves, Francisco, Maliha à Montréal, Fernanda et Mariana à Ottawa, Michael, Mina, Jacopo, Emily, Giulia, Myongjin, Alejandro, Karim, Athanu et Marco à Boston avec qui j'ai tout partagé et qui ont été d'un soutien incommensurable. Je tiens à remercier la famille Pajaro, Matias et Anne-Sophie, ainsi que leurs adorables enfants Esteban et Francisco pour toute la chaleur familiale qu'ils m'ont apporté.

Je remercie le Département de sciences économiques de l'Université de Montréal, le Centre Interuniversitaire de Recherche en économie Quantitative (CIREQ) et la Faculté des études supérieures et postdoctorales pour le soutien financier. Mes remerciements vont également à tout le corps professoral, au personnel administratif et particulièrement à Mélanie qui m'a toujours gratifié de sa bonne humeur et de son adorable sourire. Je termine cette section en portant une attention particulière à ma famille: mes parents Walter et Gudrun pour tout leur soutien et surtout cette éducation qu'ils m'ont inculqué et ma petite soeur Ruth qui qui m'a toujours gratifié de ses sages conseils et enfin mon adorable copine Myriam qui a toujours été si indulgente avec moi et m'a toujours soutenu malgré tout le temps passé sur mon ordinateur à travailler sur cette thèse.

Résumé

Dans ma thèse doctorale, j'étudie trois facteurs importants qui caractérisent le commerce international : les différences technologiques entre les pays, les barrières à l'entrée sous la forme de coûts fixes et la migration internationale.

Le premier chapitre analyse si les différences technologiques entre les pays peuvent expliquer la spécialisation dans le commerce international entre les pays. 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 160 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 ricardienne 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.

Une implication de ces résultats est qu'il est important d'évaluer jusqu'à quel point la volatilité de production mesurée par la volatilité du PIB est motivée par la spécialisation des exportations et des importations. Étant donné le compromis entre l'ouverture du commerce et la volatilité de production, les bénéfices tirés du commerce peuvent s'avérer plus faibles que

ceux estimés précédemment. Par conséquent, les politiques commerciales alternatives telles que l'ouverture graduelle au commerce combinée à la diversification de la production pour réduire la concentration de l'exportation peuvent se révéler être une meilleure stratégie que l'approche du laissez-faire.

En utilisant la relation entre la taille du marché et l'entrée de firmes et produits, le deuxième chapitre évalue si les barrières à l'entrée sous la forme de coûts fixes à exporter sont au niveau de la firme ou au niveau du produit. Si les coûts fixes se trouvent au niveau de la firme, la firme multiproduits a un avantage de coût de production par rapport aux autres firmes parce qu'elles peuvent diviser les coûts fixes sur plusieurs produits. Dans ce cas, le commerce international sera caractérisé par peu de firmes qui exportent beaucoup des produits. Si les coûts fixes sont au niveau du produit, l'entrée d'un produit est associée avec l'entrée de plusieurs firmes. La raison est qu'une fois que la première firme entre et paye les coûts fixes du produit, elle crée un effet d'entrainement qui réduit les coûts fixes pour des firmes rivales. Dans ce cas, le commerce international sera caractérisé par plusieurs firmes qui vendent des variétés différentes du même produit. En utilisant des données détaillées provenant de 40 pays exportateurs à travers 180 marchés de destination, mes résultats montrent que les barrières à l'entrée se trouvent principalement au niveau du produit. Un marché plus large favorise l'expansion d'un plus grand nombre d'entreprises au sein d'une catégorie de produit plutôt que de permettre aux entreprises produisant plusieurs produits de croître dans une gamme de produits. En regardant la différence entre le nombre d'exportateurs au sein d'une catégorie de produit dans des destinations données, je trouve que le taux d'entrée de firmes augmente significativement après qu'un produit entre la première fois dans le marché. J'en déduis donc que le premier entrant réduit les coûts fixes pour les firmes suivantes. Mes recherches démontrent également que malgré une plus grande compétition sur le marché du produit, les entreprises disposent de revenus d'exportation supérieurs et sont plus susceptibles de rester sur les marchés internationaux. Ces résultats sont cohérents avec l'hypothèse que l'effet d'entrainement incite l'entrée de firmes rivales et permettent aux entreprises de produire à plus grande échelle.

Cette recherche dévoile un nombre de conclusions importantes. D'abord, les politiques commerciales encouragent l'entrée de nouveaux produits, par exemple, en promouvant des produits dans les marchés de destination entraînant ainsi des retombées qui se traduiront par un taux de participation plus élevé de l'entreprise et une croissance de l'exportation. Deuxièmement, les consommateurs du pays importateur peuvent bénéficier de prix plus bas pour le produit en réduisant les barrières techniques du commerce. Troisièmement, lorsque l'on effectue des expérimentations politiques sous la forme de réduction des coûts commerciaux, il est de coutume de considérer uniquement une baisse des coûts marginaux et d'évaluer les répercussions sur le bien-être du consommateur. Cependant, un élément important des accords commerciaux est la réduction des barrières techniques au commerce grâce à la négociation de normes communes pour un produit. Négliger l'existence des barrières à l'entrée et les conséquences des réaffectations de l'industrie affaiblit l'impact des réformes commerciales.

Le troisième chapitre prend en compte le rôle de l'information dans la facilitation du commerce international. Les immigrants réduisent les coûts de transaction dans le commerce international en fournissant des informations sur les possibilités d'échange avec leur pays d'origine. En utilisant des données géographiques détaillées sur l'immigration et les importations aux États-Unis entre 1970 et 2005, je quantifie l'incidence qu'ont les nouveaux immigrants sur la demande pour les importations de biens intermédiaires aux États-Unis. Pour établir le lien cause à effet entre le commerce et la migration, j'exploite l'important afflux d'immigrants d'Amérique centrale après l'ouragan Mitch. Les résultats montrent que l'augmentation de dix pour cent d'immigrants a fait croître la demande pour les importations de biens intermédiaires de 1,5 pour cent. Mes résultats sont robustes aux problèmes de la causalité inverse ou la décision d'émigrer est causée par des opportunités de faire du commerce.

Mots-clés : coûts fixes, effets d'entrainement, taille du marché, théorie ricardienne, concentration d'exportation, concentration d'importation, migration, réfugié politique, ouragan Mitch

Abstract

In my PhD thesis I study three factors that shape international trade patterns: technological differences across countries, entry barriers in the form of fixed costs and international migration.

The first chapter analyses whether technology differences across countries can explain specialization patterns in international trade. To measure specialization, I compute concentration indexes for the value of exports and imports and decomposes the overall concentration into the extensive product margin (number of products traded) and intensive product margin (volume of products traded). The results show that exports are more concentrated than imports, specialization occurs mainly in the volume of trade and larger economies have more diversified exports and imports because they trade more products. I then evaluate the ability of the Eaton-Kortum model, the workhorse model of Ricardian trade theory, to account for the observed facts. The results show that technology-induced specialization through comparative advantage can explain the qualitative and quantitative facts. The key determinants of specialization are the degree of comparative advantage, the elasticity of substitution and geography.

Based on the relationship between entry and market size, the second chapter evaluates whether fixed costs are at the firm or at the product level. Within an empirical framework, I argue that fixed costs at the firm level induce a cost advantage for multi-product firms and international trade will be characterized by few firms selling many products. On the other hand, if fixed costs are at the product level, product entry is accompanied with lots of firms entry and international trade will be characterized by many firms selling different varieties of the same product. Using detailed product level data from 40 exporting countries to 180 destination markets, the results indicate

that entry barriers operate at the product level. The key implication of the product fixed cost is that the firm that pays the fixed cost creates a positive externality that lowers entry costs for rival exporters and increases firm entry. Looking at firm entry within products across time and destinations, I find evidence consistent with these spillover effects. Combined these results have important implications for the effects of trade policy on market structure and on the resulting gains from trade in the exporting as well as in the importing country.

The third chapter considers the role of migration in facilitating international trade. Immigrants can increase international trade by shifting preferences towards the goods of their country of origin and by reducing bilateral transaction costs. Using geographical variations across US states for the period 1970 to 2005, we quantify the impact of immigrants on intermediate goods imports. We address endogeneity and reverse causality - which arises if migration from a country of origin to a US state is driven by trade opportunities between the two locations - by exploiting the exogenous allocation of refugees within the US refugee resettlement program. Our results are robust to an alternative identification strategy, based on the large influx of Central American immigrants to the United States after hurricane Mitch. We find that a 10 percent increase in recent immigrants to a given US state raises intermediate imports from those immigrants' country of origin by 1.5 percent.

Keywords : Fixed costs, spillovers, market size, Ricardian trade theory, export concentration, import concentration, migration, political refugees, hurricane Mitch

Contents

List of Figures

List of Tables

Introduction

We live in a more and more globalized world where countries increasingly exchange their economic resources. Deeper international economic integration expands the global goods, capital and labor markets. The costs and benefits of international exchange continue to be widely discussed and are among the most pressing issues in a globalized world. This thesis contains three chapters that study several relationships that are important for understanding the full impact of economic globalization around the world: i) determinants of the pattern of trade; ii) entry barriers to international trade; and iii) the effects of migration on international trade.

In the first chapter, titled "Specialization patterns in International Trade", I assess the ability of a standard workhorse model of international trade to explain the pattern of trade. To start with, I document new facts on the pattern of international specialization by looking at export and import concentration. 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 the intensive product margin (volume of products traded). Using detailed product-level trade data for 160 countries, I find that exports are more concentrated than imports, specialization occurs mainly on the intensive product margin, and larger economies have more diversified exports and imports because they trade more products. Based on these facts, I evaluate 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 technology differences can explain the qualitative and quantitative facts. Also, I evaluate the role of the key determinants of specialization: the degree of comparative advantage, the elasticity of substitution and geography.

These results have important policy implications. Excessive specialization results in a high degree of export concentration implying that only few products are responsible for the majority of export revenue and import expenditure. Given that openness to trade induces specialization, globalisation will increase output volatility and therefore the likelihood that product specific shocks have aggregate effects on output volatility and/or negative effects on the terms of trade. The implied trade-off between openness to trade and output volatility suggests that benefits from trade may actually be lower as previously estimated. As a result, alternative trade policies such as gradual openness to trade combined with diversification of production to lower export concentration may represent a better strategy than a laissez-faire approach.

The second chapter, titled “Entry barriers to International Trade: product versus firm fixed costs“, focuses on the importance of fixed costs to export in shaping the international trade pattern. Fixed costs to export create entry barriers that restrict trading opportunities between countries. In the presence of these costs, market size will matter because larger markets ease the relevance of fixed costs by allowing firms to slide down the average cost curve and produce at a more efficient scale. I use the relationship between the extensive margins of exports and destination market size to evaluate whether fixed costs operate at the firm or at the product level. If fixed costs are at the firm level, multi-product firms have a cost advantage and dominate international trade. If fixed costs are at the product level, many firms export different varieties of the same product. Using detailed product level data from 40 exporting countries to 180 destination markets, the results indicate that entry barriers operate at the product level. The key implication of the product fixed cost is that once a firm or a group of firms decides to pay the fixed costs, subsequent exporters of those products do not need to pay this fixed cost. Looking at firm entry within products across time and destinations, I find evidence consistent with the spillover effect. Once a product enters a destination market for the first time, the firm entry rate increases significantly the following two years before gradually slowing down. Further evidence shows that the efficiency gains in production through lower product fixed costs outweigh the competition effects from more firm entry.

A number of important conclusions emerge from this research. First, trade theory based on increasing returns to scale explains entry behaviour of firms and products into international markets. Second, trade policies encouraging new product entry, such as advertising products in destination markets through export promotion agencies, rather than firm entry could potentially lead to spillover effects that translate into higher level of firm participation and export growth. Third, when conducting policy experiments in the form of a reduction in trade costs, it is standard in the literature to consider only reductions in marginal costs and evaluate the resulting impact on consumer welfare. However, an important aspect of free trade negotiations is the reduction of technical barriers to trade by establishing common product standards, see the current EU-US free trade negotiations. Neglecting the existence of entry barriers and the resulting industry reallocations underestimates the impact of trade reforms.

The third chapter, titled “The causal impact of migration on U.S. intermediate goods’ trade: Evidence from a natural experiment“, focuses on the role of migrants in international trade.International migrants can increase international trade flows via two distinct mechanisms. First, migrants shift preferences towards their origin country’s products, thus creating demand for imports of those goods by their host country. Second, migrants reduce transaction costs between countries, either by carrying information about relevant market characteristics or by attenuating frictions due to imperfect contract enforcement. Using geographical variation across U.S. states for the period 1970 to 2005, we quantify the impact of immigrants on intermediate goods’ imports. In particular, we follow a gravity approach and regress the log of intermediate imports on the log of recent immigrants. We address reverse causality, and more broadly endogeneity, and find evidence of a significantly positive effect of immigrants on U.S. intermediate imports. To estimate a causal relationship, we exploit the large influx of Central American immigrants to the U.S. after hurricane Mitch. Our results are robust to using an alternative identification strategy, based on the exogenous allocation of refugees within the U.S. refugee resettlement program. We find that a 10 percent increase of recent immigrants to a given U.S. state increases imports from those immigrants’ origin country by 1.5 percent. Overall, our results suggest that immigrant networks play an important role in promoting trade across countries.

Chapter 1

Specialization patterns in International Trade

1.1 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 evaluating the gains from trade through specialization but also informs how the trade pattern affects the structure of an economy. For example, a high degree of specialization increases the likelihood that product specific shocks have aggregate effects in terms of output volatility and/or an impact on the terms of trade.

The contribution of the paper is twofold. Firstly, it uncovers new facts on the pattern of specialization by looking at 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 on all margins. The extensive product margin indicates the degree of specialization in the *number* of goods traded. The concentration index on the intensive margin measures specialization in the *volume* of goods traded. Secondly, the paper evaluates the Eaton and Kortum [2002] model's ability to account for the observed specialization patterns. Specifically, it tests the model based on three basic questions about specialization: What explains the level of specialization in exports and imports? What determines the gap between specialization in exports and imports? Does specialization occur on the intensive or extensive product margin?

Based on detailed product-level trade data for 160 countries, the results show that, on average, countries specialize more in exports relative to imports, with Gini coefficients of 0.98 and 0.91 respectively. The decomposition reveals that specialization of exports occurs predominately on the extensive margin. Countries receive their export revenues from few products. At the same time, countries import a wide range of products but concentrate their expenditure towards a small number of products. Hence specialization of imports is driven by the intensive margin. The difference between the concentration levels of exports and imports is due to the extensive margin. Countries specialize in exporting few goods and diversify on imports by acquiring various products from abroad. Focusing on cross-country differences, I find 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.

Having documented the observed specialization pattern, I employ a standard Ricardian trade model developed by Eaton and Kortum [2002] to evaluate its ability to reproduce the stylized facts. To incorporate the fact that goods cross borders multiple times, my analysis relies on Alvarez and Lucas [2007] version of the Eaton and Kortum model. A key implication of this model is that it uncovers how comparative advantage due to technology differences determines specialization endogenously on both the extensive and the intensive product margins. Furthermore, it identifies geography together with the elasticity of substitution and the degree of absolute and comparative advantage as the main determinants of specialization. A higher level of technology increases a country's absolute advantage and diversifies the extensive margin of exports by broadening the product range it exports. The degree of comparative advantage heightens the sensitivity of concentration to changes in unit costs, thereby dictating specialization on both margins. Trade costs decrease comparative advantage and increase specialization on the extensive and intensive margin. A higher elasticity of substitution provides for better substitution between intermediate goods and allows countries to concentrate their expenditure in low price sectors. As a consequence, concentration on the intensive margin increases. Note that the model characterizes import concentration on all margins.

To calibrate the model, I follow Waugh [2010] and use data and the structure of the model to infer trade costs, technology and the elasticity of substitution. Not surprisingly, the simulated results show that the model produces the observed specialization pattern qualitatively with countries being specialized in exports and diversified in imports on all margins. More importantly, the simulated model also reproduces the degree of concentration on the extensive versus the intensive margin for both, exports and imports. However, the obtained levels for exports are too high in comparison to the data. Focusing on the variation across countries, the simulated model replicates the fact that larger economies are more diversified in exports but fails to account for the observed cross-country pattern of imports.

This paper contributes to the international trade literature that analyses the relationship between the pattern of trade and specialization in commodities. Leamer [1984], who tested whether the structure of trade can be explained by the availability of resources, started an empirical literature on specialization by relating the pattern of trade to factor endowments motivated by the Hechscker-Ohlin theorem, see, for example, Bowen et al. [1987], Trefler [1995] and Schott [2003]. On the other hand, MacDougall [1951], Balassa [1963], Golub and Hsieh [2000] and Costinot et al. [2012] use trade data to test the Ricardian prediction that countries export relatively more of the commodities they are relatively more productive in. Unlike these papers, my analysis does not intend to explain why countries specialize in a certain commodity or group of commodities. Instead, it uses the level of concentration in trade data to shed light on the factors that drive specialization in the number and the volume of goods traded. The levels of concentration in each trade direction contain information on the pattern of trade and as such they provide a new quantitative test of the extent of specialization observed in the data.

The analysis presented in this paper is also related to a growing literature in quantifying the importance of Ricardian comparative advantage in explaining trade patterns using the Eaton-Kortum framework, see, 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 the sectorial 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 assumed to be 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 on the pattern of trade and analyze how geography, tastes and absolute and comparative advantage induce countries to specialize in narrow sectors. In particular, I characterize the models predictions on export and import concentration on the intensive and extensive product margin and highlight the implications on the specialization pattern.

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 on the intensive and extensive product margin. However, the underlying mechanisms of generating the specialization pattern differ. In monopolistic competition and Armington models, tradable goods are differentiated by location of production since each country is the sole producer of a good. Thus, countries specialize completely and import all other goods. In the Ricardian model of Eaton and Kortum [2002] each homogeneous good is produced by more than one country and producers directly compete with imports for the lowest price. Because multiple countries export and produce the same goods, the Eaton and Kortum model generates incomplete specialization.

Finally, my investigation adds to the empirical growth literature that analyzes the relationship between income and trade patterns on the intensive and extensive product margins, see Hummels and Klenow [2005] and Cadot et al. [2011]. Like the previous papers, I study cross-country differences by decomposing the trade pattern into extensive and intensive margins. Similar to their analysis I quantify the contribution of each export margin and relate the outcome to income differences across countries. In addition, I apply the decomposition to imports and use the resulting empirical evidence to test the ability of the Eaton-Kortum model to explain the relationship of income differences and the concentration of exports and imports along both margins. While Hummels and Klenow [2005] stress that models with Krugman firm-level product differentiation can explain why larger economies export a larger number of goods, my analysis shows that the Ricardian model of Eaton-Kortum based on constant returns to scale offers an alternative framework to describe the observed patterns. The novel approach of linking cross-country variation of export and import concentration to test the Eaton-Kortum model sheds light on how the interaction between preferences, technology and geography establishes trade patterns on the intensive and extensive product margin. As such, the Eaton-Kortum framework can provide theoretical guidance for future work.

The rest of the paper is organized as follows. Section 2 describes the data and presents the empirical evidence of 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 while section 7 concludes.

1.2 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, we examine the properties of the concentration measurements used, which form the basis of the qualitative and quantitative tests of the model.

1.2.1 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 in operation 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 :

$$G = \frac{2}{N}\frac{\left(\sum_{k=1}^{N} kR_k\right)}{\sum_{k=1}^{N} R_k} - \frac{N+1}{N} \tag{1.1}$$

where export revenues of product k, R_k, are indexed in increasing order, 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) the revenues are the same across them. 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 from the mean export revenue ($\bar{R}$) and defined by the following formula

$$T = \frac{1}{N}\sum_{k\in N}\frac{R_k}{\bar{R}}\ln\left(\frac{R_k}{\bar{R}}\right) \tag{1.2}$$

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 margin, $T = T^{ext} + T^{int}$. 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:

$$T^{int} = \frac{1}{N_x} \sum_{k \in N_x} \frac{R_k}{\bar{R}_x} \ln \left(\frac{R_k}{\bar{R}_x} \right) \tag{1.3}$$

and the extensive Theil index is

$$T^{ext} = \ln \left(\frac{N}{N_x} \right) \tag{1.4}$$

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

1.2.2 Data

To build my empirical evidence, I use the Comtrade data set collected by the United Nations and choose the 6 digit HS 1992 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 to different varieties of the same product. I assume that the tradable goods sector corresponds to manufactures defined to be the aggregate across all 34 BEA manufacturing industries, see Feenstra et al. [1997].[1] Using a correspondence table provided by Feenstra et al. [1997], I identify 4529 tradable manufacturing products. The baseline sample covers 160 countries representing all regions and all levels of development between 1992 and 2009 (18 years). It includes 129 developing countries, defined by the World Bank as countries with per capita GDP under $16,000 in constant 2005 PPP international dollars. In total, the sample consists of 2880 observations (country-years).

1. 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.

Note the data contains import and export flows within 6 digit product categories. The model I am testing is Ricardian and does not feature trade between 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. [2] To measure the importance of trade between products and trade between varieties, I follow Grubel and Lloyd [1975] and calculate the percentage share of trade between products with respect to total trade. I find an average value of 81 percent across countries. The overall share of total net trades flows with respect to total gross trade flows is 65 percent. Both findings suggest that the majority of trade flows between countries in this sample is across products. [3]

Based on net trade flows at the product level, I first calculate the corresponding concentration indexes on all margins for each year and then take the average over the whole sample period. Because the concentration indices employed are independent of scale, the calculation on a year-to-year basis avoids the need to deflate the data. Figure 1.1 plots the mean export against the mean import concentration for each country together with the 45 degree line. In terms of overall concentration, Figures 1.1(a) and 1.1(b), the vast majority of observed levels lie above the 45 degree line highlighting the fact that exports are more concentrated than imports for almost all countries. On the intensive product margin, Figure 1.1(c), the specialization level of exports is similar to imports with slightly higher levels of concentration for exports. Figure 1.1(d) plots the results for the extensive product margin with countries exporting fewer products than they import.

Table 1.1 summarizes the sample statistics with the average year-by-year indices over the 2880 country-year pairs. As implied by Figure 1.1, exports are more concentrated than imports on all margins. With respect to overall

2. 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.

3. 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 allows examining 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 1.1 – Mean concentration indexes over 2880 country-year pairs

	Gini		Theil Exports (X)			Theil Imports (M)		
	Exp	Imp	Ext. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	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%	

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. The share of the intensive margin with respect to overall concentration is 60 percent. Focusing on the gap between export and import concentration, Table 1.1 shows that differences between exports and imports are mainly explained by the extensive margin with a Theil of 2.60 for exports versus 1.10 for imports. The Theil of 1.10 on the extensive margin implies that, on average, a country net imports a 33.3 percent of all products. On the other hand, the extensive Theil of 2.60 indicates that a country net exports 7.4 percent of the product space. 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 few exporting sectors and diversify the importing sectors.

Turning to cross country differences, the empirical evidence shows that larger economies diversify more than smaller economies. Figure 1.2 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 the average relative GDP to the United States ($USA = 0$) over the periods 1992 to 2009. As Figures 1.2(a) and 1.2(b) show, the overall Theil index decreases with respect to relative GDP, 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 specialization on the intensive margin does not vary with market size for both, exports (Figures 1.2(e)) and imports (1.2(f)). The main driver of specialization differences across countries is the extensive margin. Particularly robust is the linear relationship on the extensive margin for exports 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 on the extensive margin of imports follows a L shape pattern. As the size of an economy increases, countries diversify on imports until reaching 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 on all margins. Second, the extensive margin drives concentration of exports and the intensive margin for imports. Third, the target levels of concentration are displayed in Table 1.1. 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 the paper evaluates the Ricardian model's ability to account for these stylized facts. Next, I present the relevant parts of the Alvarez and Lucas [2007] extension of the Eaton-Kortum framework.

1.3 Model

The Eaton–Kortum model is Ricardian, with a continuum of goods produced under a constant-returns technology. In this paper, we focus on the Alvarez and Lucas [2007] model, which builds on the Eaton Kortum framework. I derive the relevant theoretical predictions on the pattern of trade and evaluate the importance of the key model parameters for specialization of imports and exports.

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], 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:

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

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. $x_i \sim \exp(1/\lambda_i)$. Denote the wage by w_i and the price of the intermediate aggregate good by $p_{m,i}$. The intermediate good sector is perfectly competitive. Intermediate good producers minimize input costs and sell the tradable intermediate good at price

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

where $B = \beta^{-\beta}(1-\beta)^{-(1-\beta)}$. The continuum of intermediate input goods x enters 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 produced aggregate intermediate good q_i can then be allocated costless towards the production of final goods or being used as an input in the production of intermediate goods. Similarly, labor, as the only primary factor input, can be used either to produce intermediate or final goods. Finally, consumers draw their 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. The interested reader is refereed to Alvarez and Lucas [2007] for the full description of the model.

1.3.1 General equilibrium

Once a country opens to international goods markets, the intermediate goods are the only goods traded. Final goods are not traded and labor is immobile between countries. Due to trade costs, factor prices do not equalize across countries. The intermediate goods needed to produce the composite good are acquired from the producer of good x in the country that operates at lowest unit costs.

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 charged 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

$$p_i(x) = B \min_j \left[\frac{w_j^{\beta} p_{mj}^{1-\beta}}{\kappa_{ij}\omega_{ij}} x_j^{\theta} \right]. \tag{1.5}$$

Equation 1.5 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 lowest costs in the local market, the good is imported. Following Alvarez and Lucas, the resulting price index of tradable goods in country i is

$$p_{mi} = (AB) \left(\sum_{j=1}^{I} \left(\frac{w_j^{\beta} p_{mj}^{1-\beta}}{\kappa_{ij}\omega_{ij}} \right)^{-1/\theta} \lambda_j \right)^{-\theta} \tag{1.6}$$

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 a minimum in equation 1.5. 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

$$D_{ij} = (AB)^{-1/\theta} \left(\frac{w_j^{\beta} p_{mj}^{1-\beta}}{p_{mi}\kappa_{ij}\omega_{ij}} \right)^{-1/\theta} \lambda_j. \tag{1.7}$$

Equation 1.7 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 at 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 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 of 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.

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

$$L_i p_{mi} q_i \sum_{j=1}^{I} D_{ij} \omega_{ij} = \sum_{j=1}^{I} L_j p_{mj} q_j D_{ji} \omega_{ji} \tag{1.8}$$

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 on export and import concentration on both margins.

1.3.2 Concentration of exports and imports

In the model, the pattern of trade is established by domestic producers competing with importers for selling intermediate goods in the local market. If foreign producers sell a particular good at a lower price than domestic producers, the good will be imported from the cheapest source. Given the equilibrium price, $p(x)$, and quantity, $q(x)$, the total expenditure 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 where country i obtains the minimum price at home. Equivalently, domestic producers export their good to all foreign markets where they offer the minimum price. The set of exporting goods is simply a collection of the set of goods country i exports to any destination j, $x \in \cup^I_{j \neq i} \mathbf{B}_{ji}$. As a result, (f.o.b.) export revenue sales of good x, $R_{i,X}(x)$, are given by:

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

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 margin. Using equation 1.3, the Theil index for the concentration of imports on the intensive production margin can be written as:

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

In the appendix I show that the distribution of import expenditures 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:

$$T^{int}_{iM} = \ln \left(\Gamma(1 + \theta(1 - \eta)) \right) - \int_0^1 \ln \left(u^{(-\theta(1-\eta))} \right) e^{-u} du \qquad (1.9)$$

where $\Gamma(.)$ stands for the Gamma function. Import specialization on the intensive margin is independent of equilibrium prices, trade costs, geography and the level of technology λ. It is solely determined 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 towards these sectors. Similar, 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 of imports on the extensive margin, 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 :

$$T_{iM}^{ext} = \ln\left(\frac{N}{N_{iM}}\right) = -\ln(1 - D_{ii}) \tag{1.10}$$

where

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

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 on the intensive margin according to equation 1.3. The extensive Theil index on the extensive margin is given by the inverse share of the number of goods exported, N_{iX}, with respect to the total number of simulated goods, N.

$$T_{iX}^{ext} = \ln\left(\frac{N}{N_{iX}}\right)$$

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

1.4 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. Concerning the remaining parameters of the model, I use the same values as Alvarez and

Lucas [2007]. I assume a variance of individual productivities $\theta = 0.12$, an elasticity of substitution of intermediate goods $\eta = 2$, an efficient labor share in the production of non-tradable final goods $\alpha = 0.75$ and an efficient labor share in the tradable goods sector of $\beta = 0.5$. The simulation contains a total of $I = 160$ countries. In the following subsections, I analyze import and export concentration in special cases of the equilibrium by assuming different trading schemes. Doing so builds 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 introduce heterogeneity across countries later on. 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.

1.4.1 Symmetric countries

All countries are identical with respect to their size $L_i = L$ and technology parameter $\lambda_i = \lambda$. Trade costs are symmetric and set to $\kappa_{ij} = \kappa \; \forall \; i \neq j$ with $\kappa_{ii} = 1$ and $\omega_{ij} = 1 \; \forall i, j$. I solve for the equilibrium wage in each country, all goods prices and the value of imports and exports. Due to symmetry, wages and composite good prices equalize across countries. Trade costs distort international trade. In comparison to free trade, firms will not be able to buy good x from the cheapest producer world wide and rely more on home production. The corresponding trade share matrix D is symmetric and the (i, j) element is given by:

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

where D_{ij} is the set of goods country j exports to country i and D_{ii} the set of goods country i produces at home. 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 $1 - D_{ii} = (I-1)/I$. Hence, the more countries participate in international trade, the more countries specialize in exports and diversify in imports. In this case, Ricardian specialization forces are strongest and the gap between export and import concentration reaches a maximum.

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

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

Concentration at 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. Note that the randomness of the productivity distribution implies that in this model there is no fixed hierarchy of export destinations as in Melitz [2003], i.e. goods that are exported to the $k+1$ "most popular" destinations are not necessarily exported to the k most popular destinations.[4] 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, $U_{ex} = \cup_{j \neq i}^{I} \mathbf{B}_{ji}$. 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:

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

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 at the extensive margin increases with geographical barriers, the degree of comparative advantage and the number of trading partners. In general, a larger number of trading partners increases competition between production and imports in the domestic market resulting in the production of fewer goods at home and an increase in the number

4. In the basic version of the Melitz [2003] model, exported goods obey a hierarchy, see Eaton et al. [2011]. Any good sold to the k + 1st most popular destination is necessarily sold to the kth most popular destination as well. In that model the total number of exported goods would simply be all the products exported to the most popular destination, i.e. the destination with the lowest trade costs.

of goods imported. Also, more trading partners increase competition among exporters in foreign markets forcing countries to specialize more on 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. Notice that in the special case of free trade all goods produced are exported and concentration of production equals concentration of exports. With trade costs, countries export a subset of produced goods leading to more concentration of exports relative to production.

Concentration on the Intensive Margin As noted previously the distribution of import expenditure follows a Fréchet distribution where the concentration indexes depend on the elasticity of substitution (η) and the degree of comparative advantage (θ). Consequently, given θ and the concentration of imports observed in the data, I can pin down the elasticity of substitution. Concerning the distribution of export revenues across products, the simulation shows that it depends positively on the elasticity of substitution (η), the degree of comparative advantage (θ) and geographical barriers (κ). The number of trading partners has non-monotone effects on the concentration of exports at the intensive margin. Few trading partners increase concentration because high revenue generating exports sell in more markets. However, as the number of trading partners increases, the degree of competition in the export markets also increases and low revenue generating products do not sell in foreign markets anymore. Thus, after a threshold level, concentration among export revenues reduces with the number of trading partners. In the case of free trade, countries export all their goods to all destinations and, given that preferences are identical, export and import concentration on the intensive margin equalize. [5]

5. The intuition behind this result is that preferences are such that the import expenditure distribution is the same for each trading partner. In the appendix, I show in detail that the expenditure distribution of bilateral trade, $E_{ij}(x)$, between importer i and exporter j is the same for each source country j, i.e. $E_{ij}(x) = E_i(x)$, $\forall j \in I$. Furthermore, the bilateral import expenditure distribution, $E_i(x)$, is Fréchet with common shape parameter $1/(\theta(\eta-1))$ and country specific scale parameter s_i. The shape parameters are identical because preferences are common across countries. Note that the bilateral import expenditure distribution of country i, equals the export revenue distribution of country j. In free trade, the exporting country ships the exact same goods to all countries. As a consequence, overall export revenue distribution of country j is equal to the import expenditure distribution of each country scaled up by the number of trading partners. Since the

Table 1.2 – Simulated export and import concentration indexes for benchmark parameters.

Parameters	Gini		Theil Exports (X)			Theil Imports (M)		
	Exp.	Im.	Ext. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	Total
($\eta = 2$,$\kappa = 1$)	0.99	0.09	5.01	0.01	5.02	0.01	0.01	0.02
($\eta = 7.1$,$\kappa = 1$)	0.99	0.72	5.01	1.91	6.92	0.01	1.61	1.62
($\eta = 7.1$, $\kappa = 0.7$)	0.99	0.77	5.04	1.18	6.22	0.10	1.61	1.71
($\eta = 7.1$, $\kappa = 0.7$, 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

The results presented in Table 1.2 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. Focusing on the decomposition reveals the underlying reason: countries import too many goods and the value of those goods is too evenly distributed. Using the fact that for a given value of θ, import concentration on the intensive margin can be determined by the elasticity of substitution, I calibrate $\eta = 7.1$ to match the level observed in the data[6]. As row 2 of Table 1.2 shows, this allows composite good producers to better substitute between intermediate inputs and alters the level of import concentration.

To reduce the gap between export and import concentration caused by the extensive product margin, I follow Alvarez and Lucas [2007] and introduce 42 percent symmetric trade costs to all trading partners, $\kappa = 0.7$. Row 3 of Table 1.2 shows the results. Impediments to trade reduce the number of products exported and imported and concentration on the extensive margin increases

concentration indexes are independent of scale, the concentration of exports and imports on the intensive margin equalize.

6. This value is consistent with previous ones found in the literature, see Broda and Weinstein [2006].

for both. Note that higher trade costs lower the level of concentration on the intensive margin of exports. Due to the increase in trade costs, only efficient producers remain exporters allowing them to distribute their export revenues more evenly across products and trade partners. Although the gap between export and import concentration narrows slightly, the difference is still substantial. The reason for this is that the degree of competition countries face in export and domestic markets is too high. Creating trading blocks by introducing infinite trade costs with countries outside of the block limits the number of trading partners (NT) and reduces competition in all markets. The fit of the model improves on all dimensions, fourth row of Table 1.2. Less competition in the domestic market increases the survival rate of domestic producers and reduces the amount of goods imported. Infinite trade costs reduce the number of countries competing in a particular market and increases the probability to export to any of them. As a result, the gap between export and import concentration diminishes. Note that revenues of exporting industries are now geographically more concentrated and hence specialization on the intensive margin of exports intensifies.

In sum, with the introduction of symmetric trade costs, the model can replicate the mean levels of concentration observed in the data. The key parameters are the elasticities of substitution η and the trade cost function κ. In particular, by creating trade blocks, which amounts to introduce zeros in the bilateral trade matrix, we can calibrate the model to explain the pattern of specialization at the mean.

1.4.2 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 and size L_i to reflect the observed GDP differences in the data. To start with, consider the model's free trade equilibrium relationship between wages, size and technology:

$$w_i = \left(\frac{\lambda_i}{L_i}\right)^{\theta/(\beta+\theta)} \tag{1.11}$$

Using equation 1.11, I back out the level of technology, $\lambda_i = (w_i L_i)^{(\beta+\theta)/\theta} L_i^{-\beta/\theta}$, as a function of GDP ($w_i L_i$) and labor endowment ($L_i$). To calibrate λ, I use GDP and population data from the Penn World table. I follow Waugh [2010] and proxy labor endowment with data on a country's population and normalize the obtained parameters for λ_i and L_i relative to the United States.

Concentration on the Extensive Margin Plugging in the equilibrium wage into equation 1.7, I get the corresponding trade share matrix D with the (i, j) element given by:

$$D_{ji} = \frac{w_i L_i}{\sum_{k=1}^{I} w_k L_k} \ \forall j \tag{1.12}$$

Equation 1.12 shows that under the assumption of free trade country i's share of number of products exported is equal to its relative level of GDP with respect to world GDP. Hence, larger economies export more and import less products compared to smaller economies. This result is at odds with the empirical evidence. In the data, larger economies export and import more products. In the next section, I introduce trade costs and argue that they have to be asymmetric in order to replicate the empirical observations.

Concentration on the Intensive Margin On the intensive margin, under the assumptions of homogenous tastes across countries and free trade, export and import concentration equalize. In this case, export and import concentration on the intensive margin depend only on θ and η and are unaffected by the introduction of heterogeneity in technology and country size.

1.4.3 Asymmetric trade costs

To reconcile the cross-country concentration differences for imports, I introduce asymmetric trade costs. In particular, I consider trade costs as a function of a fixed export cost (ex_j) or a fixed import cost (im_i). The next paragraphs show the different implications of each effect.

1.4.3.1 Importer fixed effect

Under the assumption of a fixed import cost, country i faces the same cost of importing independent of the origin country j. The trade cost matrix becomes $\kappa_{i,j} = im_i \ \forall j \neq i$ 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:

$$D_{ji} = (AB)^{-1/\theta} \left(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{p_{mj} im_j} \right)^{-1/\theta} \lambda_i \ \forall i \neq j \ \text{ and } \ D_{ii} = (AB)^{-1/\theta} \left(\frac{w_i}{p_{mi}} \right)^{-\beta/\theta} \lambda_i$$

Focusing on the expression for the share of goods that country j imports from country i, $D_{j,i}$, shows that higher import costs reduce the fraction of goods that arrive in destination j ($im_j \downarrow$) and decrease the number of goods country j imports from i. Solving for the equilibrium and assuming that price differences across countries are approximately equal to the import cost differences, the corresponding share of goods imported is approximately equal to:

$$(1 - D_{ii}) \approx \left(1 - C_1 im_i^{-\frac{1}{\theta}} w_i L_i \right) \tag{1.13}$$

where C_1 is a constant independent of country i. Equation 1.13 shows that an importer fixed effect can counterbalance the fact that larger economies import less under the assumption that they face lower costs to import. Lower import costs increase the share of goods imported, $(\partial(1 - D_{ii})/\partial im_i > 0)$, and lead to a reduction in the unit cost of production through a lower price index of tradable goods, which in turn increase the competitiveness abroad.

1.4.3.2 Exporter fixed effect

In the case of an exporter fixed effect, each country pays a country specific cost to export, which is independent of the importing country i, $\kappa_{i,j} = ex_j$ $\forall j \neq i$ and $\kappa_{i,i} = 1 \ \forall j = i$. The trade share matrix is asymmetric and given by:

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

Here the expression for D_{ji} implies 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 given by:

$$(1 - D_{ii}) \approx \left(1 - C_2 ex_i^{\frac{-1}{\theta}} w_i L_i\right) \tag{1.14}$$

where C_2 represents a constant independent of country i. Equation 1.14 shows that the share of goods imported is decreasing in the country specific exporting costs, $(\partial(1 - D_{ii})/\partial ex_i > 0)$. Lower exporting costs lead to higher domestic wages, increase unit costs of production and result in a larger share of imported goods. Hence, an exporter fixed effect can reconcile the fact that larger economies import more by assuming that (1) larger economies face lower costs to export and (2) the relationship between GDP and the export cost is more pronounced than relationship between GDP and the share of goods imported, $(\partial(1 - D_{ii})/\partial w_i L_i > 0) \Rightarrow ((\partial(1 - D_{ii})/\partial ex_i)(\partial ex_i/\partial w_i L_i) > \partial(1 - D_{ii})/\partial w_i L_i)$.

The main difference between the import cost and the export cost in terms of import concentration lies in the implication on the price level of tradable goods. One can show that the export cost implies a nearly constant price level of tradable goods across countries. As a result, unit cost differences between countries are predominantly driven by wage differences. On the contrary, the import cost leads to large cross-country price level differences with smaller economies facing a higher tradable price level. In this case, unit cost differences are driven by wage as well as price level differences. Based on Waugh [2010]'s results that countries have similar price levels of tradable goods, I focus only on the case of the exporter fixed effect for the rest of my analysis.

In sum, the introduction of asymmetric trade costs in form of a country specific cost to export or import allows the model to replicate the import specialization pattern across countries, in particular when larger economies face relative low costs to either export or import. Waugh [2010] argues that trade costs have to be asymmetric, with poor countries facing higher costs to export relative to 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 on the models implications on the specialization pattern

Table 1.3 – Simulated export and import concentration indexes for asymmetric countries.

Parameters	Gini		Theil Exports (X)			Theil Imports (M)		
	Exp.	Imp.	Ext. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	Total
($\eta = 7.1$, $\kappa = 1$)	0.99	0.73	5.75	1.91	7.66	0.007	1.61	1.62
($\eta = 7.1$, $\kappa = ex$)	0.99	0.85	8.08	1.68	9.76	1.09	1.61	2.70
($\eta = 7.1$, $\kappa = ex$, 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

of exports and imports. 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 1.3 presents simulations results in the case of asymmetric countries and free trade. Note that in relation to the symmetric country case introducing technology differences increases the mean level of concentration for exports and decreases the level of concentration for imports. The underlying reason is that the technology distribution is skewed towards less productive countries and these countries export fewer and import more goods. Beside these changes, the results are similar to the symmetric case. While exports are more concentrated than imports, the simulated level of concentration for exports (imports) is too high (low) compared to the data. The reason is excessive specialization (diversification) on the extensive margin of exports (imports).

In terms of cross country differences, calibrated GDP differences in combination with zero or symmetric trade costs lead to the false prediction that larger economies import less goods. To reconcile the empirical evidence, I introduce country specific costs to export with larger economies facing relatively lower export costs. In particular, I calculate the implied export cost from equation 1.14 by replacing the share of goods produced at home by the extensive Theil index of imports observed in the data , $D_{ii} = 1 - \exp(-T_M^{Ext})$. Row 2 of table 1.3 shows the results of the corresponding mean concentration levels. While the model matches the cross country concentration pattern, the mean level of export concentration is twice as high as in the data. Excessive competition in the export markets leads to high levels of concentration on the extensive margin of exports. To counter the competition effect, I create trading blocks between countries (i.e. limiting the number of trading partners $NT = 10$) . I assume that countries within a block trade with countries whose market size is similar.[7] Trade blocks conditional on market size improve the fit of the model. The obtained concentration levels match the data on all dimension. Row 3 of Table 1.3 presents the results. Countries are more concentrated in exports than in imports on all margins, the intensive margin dominates in terms of overall concentration and the simulated concentration levels are close to the target levels. In terms of the cross country pattern, Figure 1.3 plots the simulated (in red) and the empirical (in blue) concentration levels against GDP for both margins. The figure shows that the country specific export cost in combination with technology and endowment differences can replicate the across country evidence on all margins. Bigger economies are more diversified because they import/export more products and concentration patterns on the intensive margin are insensitive to market size.

In the previous section I analyzed special cases of the equilibrium 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 at 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.

1.5 Estimating trade costs from bilateral trade shares

The starting point of the estimation of technology and trade costs is a structural log-linear "gravity" equation that relates bilateral trade shares with

7. The precise trade cost configuration is given in the appendix.

trade costs and structural parameters of the model. To derive the relationship, simply divide each country i's trade share from country j, see equation 1.7, by country i's home trade share. Taking logs yields $I-1$ equations for each country i :

$$\log\left(\frac{D_{ij}}{D_{ii}}\right) = S_j - S_i + \frac{1}{\theta}\log(\kappa_{ij}) + \frac{1}{\theta}\log(\omega_{ij}) \tag{1.15}$$

in which $S_i = \log(w_i^{-\beta/\theta} p_{mi}^{-(1-\beta)/\theta} \lambda_i)$ presents the structural parametersof the model. In order to estimate trade costs κ and technology λ implied by equation 1.15 I use data on bilateral trade shares across 160 countries. I calculate the corresponding bilateral trade share matrix by the ratio of total gross imports of country i form 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, we use gross manufacturing output data from de Sousa et al. [2012]. Combined with trade data from Comtrade, we get the expenditure share, D_{ij}, which equals the value of inputs consumed by country i imported from country j divided through 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 1992 - 2009 and then take the average expenditure share over the sample period. [8]

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

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

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

8. The resulting sample consists of 160 times 159 potential observations if each country trades with all other countries. In our sample the total number of observations is 9649 implying a large number of zeros in the bilateral matrix. For this reason, I conduct a robustness test where I estimate the model with the Poisson estimator proposed by Silva and Tenreyro [2006]. The appendix presents the results.

Table 1.4 – Estimation Results

Summary Statistics			
Observations	TSS	SSR	R^2
9649	2,60E+05	4,67E+04	0.82

Geographical barriers

Barrier	Parameter estimate	Std. error	% effect on cost
[0,375)	-4,89	0,10	79,93%
[375,740)	-5,76	0,06	99,60%
[750,1500)	-6,78	0,04	125,62%
[1500,3000)	-7,98	0,03	160,66%
[3000,6000)	-9,05	0,02	196,42%
[6000,max)	-9,81	0,03	224,64%
Tariff	-0,23	0,10	5,47%
Shared border	1,37	0,09	-15,19%

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 1.6 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, $\log(w_i^{-\beta/\theta}\lambda_i)$. Then, given the bilateral trade shares, $D_{i,j}$, and the balanced trade condition in equation 1.8, I follow Alvarez and Lucas [2007] and calculate equilibrium wages according to the following equation.

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

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

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

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 1.4 summarizes the regression outcome of the gravity equation. In terms of fitting bilateral trade flows, I obtain an R^2 of 0.82 slightly lower than the R^2 of 0.83 reported by Waugh. The obtained coefficients on 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 1.4 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 terms of percentage are similar to those reported in Anderson and Van Wincoop [2004].

Having identified trade costs and technology, see Table 1.6 for the estimated technology parameters, I simulate the Eaton and Kortum model to test whether the calibrated version can replicate the concentration levels observed in the data. Table 1.5 presents 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. Focusing on the obtained levels reveals that countries concentrate excessively on exports with respect to the data. The simulated concentration levels 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

Table 1.5 – Simulated concentration level with exporter fixed effect

Model	Gini		Theil Exports (X)			Theil Imports (M)		
	Exp.	Imp.	Int. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	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%	

countries export (import) 0.8 (43.2) percent of the product space compared to 7.4 (33.3) precent in the data.

Figure 1.4 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 decreasing in market size. However, the simulated concentration levels on the extensive margin are too high, particularly for smaller economies. Countries specialize excessively on the number of products exported. On the importing side, the calibrated model is unable to replicate the L shape relationship between market size and concentration. The relationship does not reveal any particular pattern. However, simulated countries tend to import more goods than in the data. Turning the attention to the intensive margin, Figures 1.4(e) and 1.4(f), the results show that, consistent with the data, the model predicts no relationship between concentration and market size. In sum, the calibrated model is able to replicate the qualitative pattern for exports but produces relatively high levels of concentration compared to the data, particularly on the extensive margin.

A potential reason for the excessive concentration in exports lies in the underlying productivity distribution. While the model reproduces the bilateral trade volumes, it fails to capture the underlying distribution of trade volume across products. To shed light on why countries trade too few products, I follow Haveman and Hummels [2004] and plot the empirical and the simulated density of the number of exporters and importers per product.[9] Figure 1.5 shows the results. In the case of exports, simulated countries export their goods to too many destinations. The assumed productivity distribution generates such efficient producers that even firms facing high trade costs can sell their products to many destinations in 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. Turning the attention to imports, Figure 1.5(b) shows that, contrary to exports, the simulated distribution of the number of countries importing a product is closely related to the empirical one. The distributions are similar at the mean, however, the empirical distribution is more dispersed. The average number of importing countries per product is 70 in the data and 75 in the model.

1.5.1 Discussion of results

There are several potential reasons why the model is not able to reproduce the cross country pattern of import concentration on the extensive margin. Note that the model implies that expenditure shares equal to product shares in the tradable sector. The product share, π_i, is defined as the number of products imported, $N_{i,M}$, divided by the total number of HS codes, N:

$$\pi_i = \frac{N_{i,M}}{N} \tag{1.16}$$

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

$$m_i = \frac{M_i}{Abs_i} \tag{1.17}$$

Figure 1.6 plots the empirical relationship between the two shares. The red line marks the 45 degree line where the two are equal. Notice that countries below the 45 degree import a lot of goods and spend relative little on those

9. 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.

goods, whereas countries above the 45 degree line import few goods and spend a lot on them.

One potential reason why expenditure shares do not equal product shares in the tradable sector could be that not all manufacturing products are tradable. When calculating expenditure shares, then ideally one wants to use absorption of the tradable sector instead of absorption of the manufacturing sector. If indeed the size of non-tradables within the manufacturing sector varies between countries, then size differences between countries can explain why countries do not align along the 45 degree line. For example, countries below the 45 degree have a higher product share than expenditure share implying that these countries spend relatively little per good. If these countries have a relatively large non-tradable sector in manufacturing, then the resulting downward bias in the measurement of import expenditure shares can explain why those countries spend relatively little on imported goods.

In addition to potential size differences of the non-tradable sector, relative prices of non-tradables differ across countries. For example, suppose that trade increases productivity but that productivity gains, in accordance with the Balassa-Samuelson hypothesis, are greater in the tradable than in the non-tradable sector. Then, relatively higher productivity in the tradable sector leads to a rise in the relative price of non-tradables. As a result, import expenditure shares with respect to tradables are in fact lower than computed import shares with respect to expenditure on tradables and non-trabables. Alcalá and Ciccone [2004] argue that computing expenditure shares with respect to GDP using real GDP instead of nominal GDP eliminates distortions due to cross-country differences in the relative price of non-tradable goods. For this reason, I experiment with computing absorption with respect to gross manufacturing production by deflating gross manufacturing production by the Purchasing Power Parity index from the Penn World table. The obtain results show that indeed countries that lie below the 45 degree line have a higher PPP index. However, the resulting concentration pattern of imports does only change slightly.

A third potential reason of why countries are not aligned along the 45 degree line can be that not all countries make use of all intermediate goods. 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 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 countries may differ in the number of intermediate tradable goods used for production due to increasing specialization in the production process. He supposes that the production of intermediate goods features increasing returns to scale external to the firm and these returns depend upon the level of technology and the size of the market.

To shed light on the role of market size on the total number of 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:

$$\frac{M_i}{N_{i,M}} = \frac{Abs_i}{N} \tag{1.18}$$

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 market size is 1. Figure 1.7 plots the relationship. Note that the figure reveals a strong positive correlation with a $R^2 = 0.84$ 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 below 1. In this case, the number of tradable goods would be country specific and increases with market size.

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 1.18, the ratio of the per product import expenditure with respect to per product tradable expenditure should be one. This result relies on the assumption of homothetic preferences. Figure 1.8 plots the log of the ratio against the log of GDP per capita. The figure shows a negative correlation of -0.67 with a $R^2 = 0.23$. Hence, richer economies tend to spend on average less per imported product relative to the average tradable product. This evidence is consistent with non-homothetic preferences, where poorer countries spend relatively more per imported good compared to rich ones.

1.6 Robustness checks

The first part addresses concerns on the robustness of the empirical observed concentration indexes. In particular, the level of disaggregation as well as the classification scheme chosen may affect the empirical concentration measures and the decomposition of the intensive and extensive margin. For this reason, I re-calculated the concentration indexes on all margins by defining a product as the equivalent of a 4 digit SITC code instead of a 6 digit HS code. The implied product space changes significantly as it comprises only 642 products compared to 4529 products using 6 digit HS codes. Also, the SITC classification scheme differs from the HS classification scheme by grouping products based on economic functions rather than their material and physical properties for tariff purposes as in the HS system. Table A.1 reports the calculated concentration indexes based on SITC together with the implied correlation between HS based and SITC based concentration indexes. Note that the empirical estimates of the SITC industry classification are very similar to the 6 digit HS code sample. The correlation coefficient between the SITC and HS Gini indexes is 0.95 for exports and 0.88 for imports. In terms of the level of the index, the obtained SITC Gini is slightly lower that the HS Gini because of the higher level of aggregation. However, the qualitative results remain the same. Exports are more concentrated than imports on all margins and the intensive margin dominates concentration for imports and the extensive margin for exports. The obtained shares of the intensive margin in terms of overall concentration are almost identical to the standard sample with 46 percent for exports and 66 percent for imports. Also, the cross country concentration patterns feature a negative log linear relationship between concentration on the extensive margin and market size for both exports and imports. In sum, the obtained results on the 4 digit SITC level support the empirical evidence based on the 6 digit HS classification and highlight the level of generality the results apply.

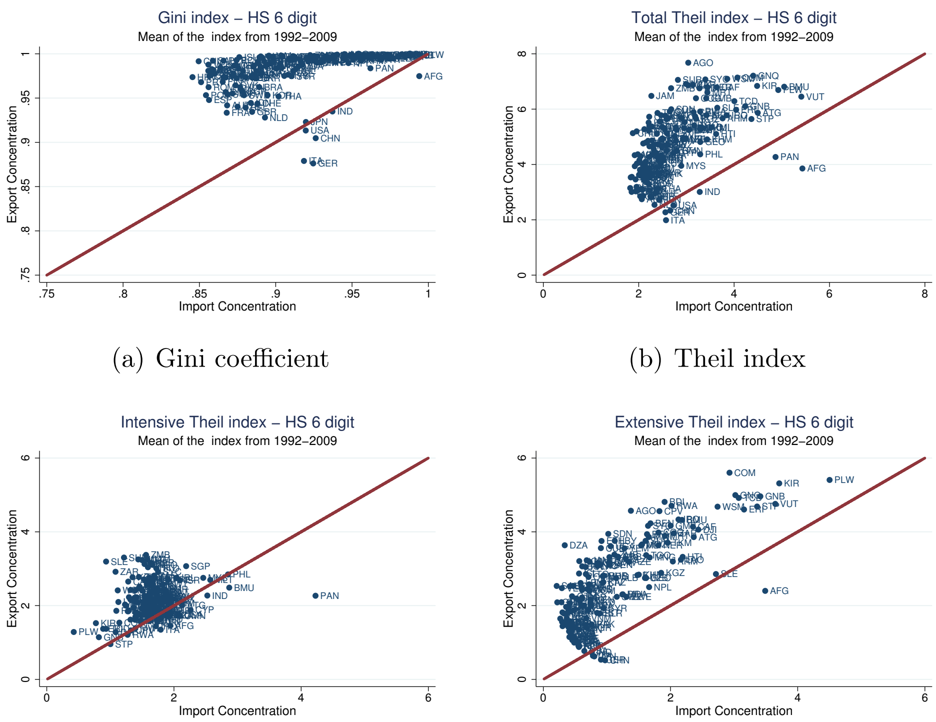
Finally I want to address the discrepancy of the product space between the data and the model caused by intra-industry trade. In the main part of the paper I establish correspondence between the model and the data by netting out within product trade and considering only trade across products. This approach leaves valuable information unused and may bias the results because intra-industry trade flows occur predominantly between OECD countries and to a lesser extend between developing countries. In an alternative approach, I deal with intra-industry trade by developing a "measurement device" that enables the model to characterize trade within and across products. 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 product level. 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 to examine gross trade flows. Because the classification scheme is unobserved, I assume that varieties are randomly assigned to 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" comprises an HS product category. Based on this result, I apply the Poisson process to group simulated Ricardian products randomly into artificial HS codes for which I calculate the implied concentration indexes. The results, presented in detail in the appendix, show that this approach leads to similar results as the net trade flow approach. In particular, it implies a similar value for the elasticity of substitution, $\eta = 7.4$ (compared to $\eta = 7.1$ in the net trade sample), and an exporter fixed effect to reconcile the cross-country concentration pattern on the extensive margin.

1.7 Conclusion

I have argued that export and import concentration in combination with a decomposition into an extensive and intensive product margin concentration measure provide new quantitative and qualitative evidence on specialization patterns in world trade. Based on detailed trade data, my calculations show that exports are more concentrated than imports on all margins and specialization is dominated 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 qualitatively and quantitatively. Overall, my results stress the importance of the role that geography and absolute as well as comparative advantage play in determining the pattern of specialization.

By looking through the lenses 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 differences and concentration differences across countries is primarily driven at the extensive margin. This relationship 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 whereas specialization along the intensive margin by exporting industries that have already a comparative advantage to more destinations intensifies the risk. Analyzing this question is an avenue for future research.

1.8 Figures



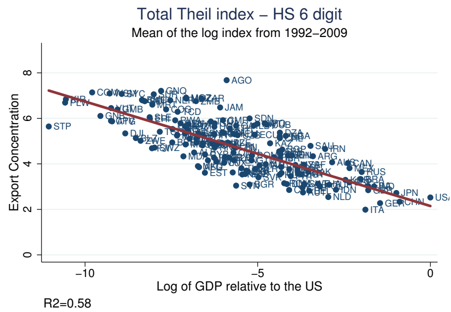
(a) Gini coefficient

(b) Theil index

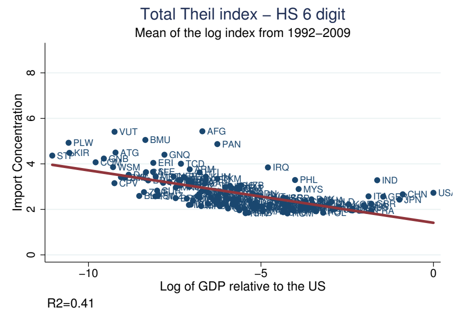
(c) Intensive margin

(d) Extensive margin

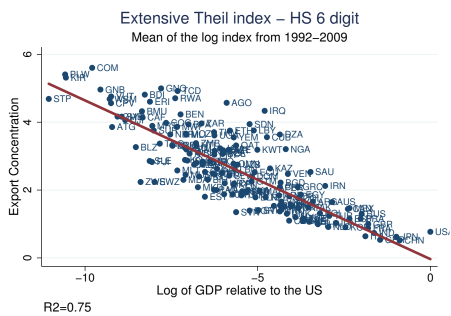
Figure 1.1 – Average export versus import concentration for the period 1992 to 2009 for 160 countries



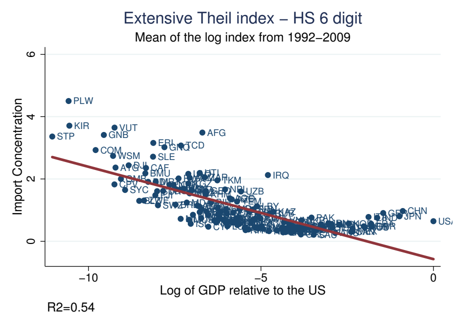
(a) Overall concentration of exports



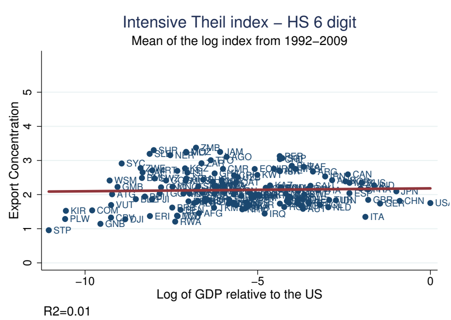
(b) Overall concentration of imports



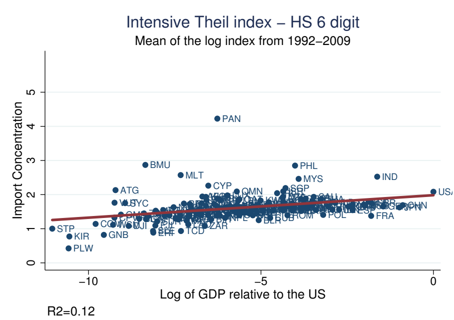
(c) Extensive margin of exports



(d) Extensive margin of imports

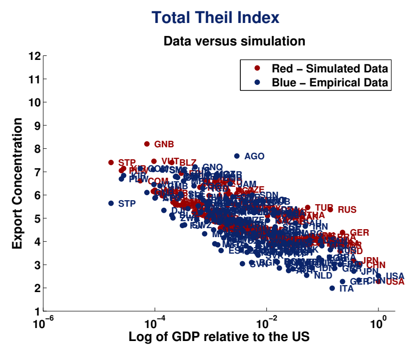


(e) Intensive margin of exports

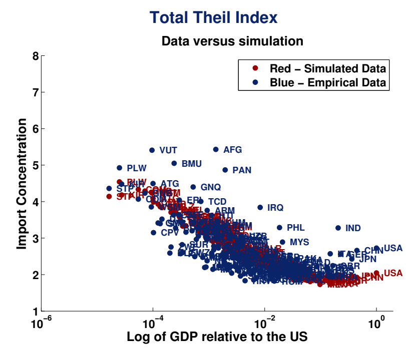


(f) Intensive margin of imports

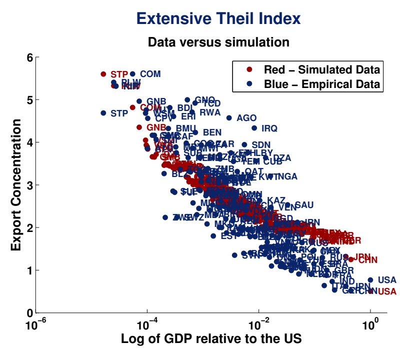
Figure 1.2 – Average export and import concentration versus the log of average relative GDP with respect to the United States for the period 1992 to 2009 for 160 countries.



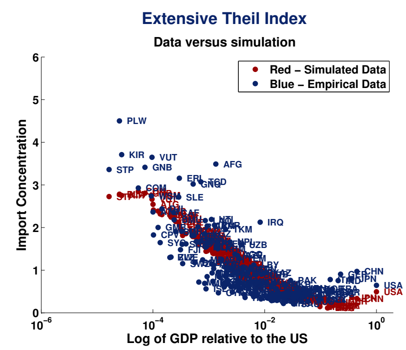
(a) Overall concentration of exports



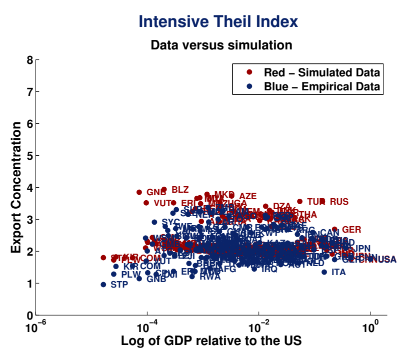
(b) Overall concentration of imports



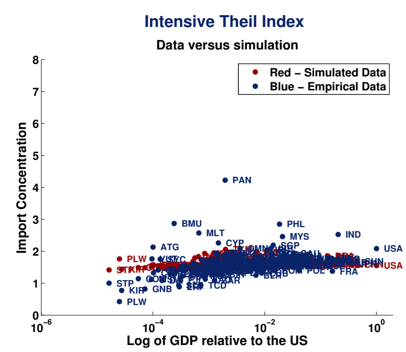
(c) Extensive margin of exports



(d) Extensive margin of imports

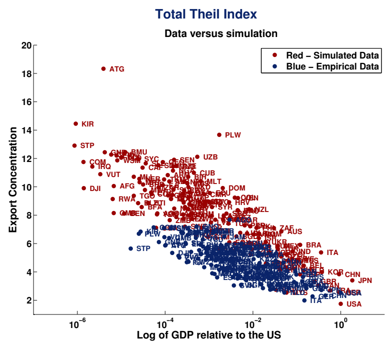


(e) Intensive margin of exports

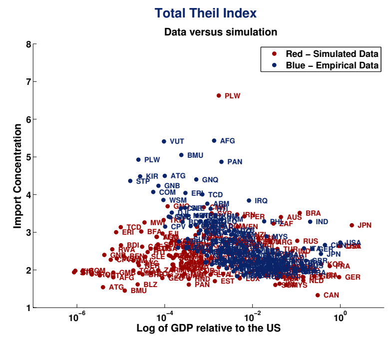


(f) Intensive margin of imports

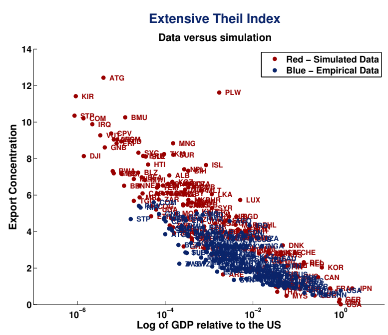
Figure 1.3 – Simulated (in red) and empirical observed (in blue) export and import concentration versus GDP across 160 countries. The simulation uses parameterized trade costs to match the data using a country specific export cost.



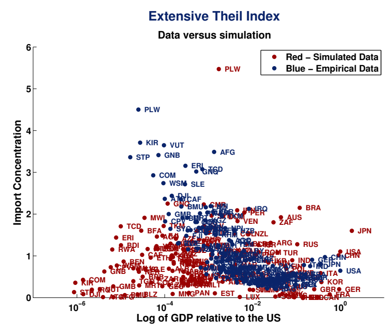
(a) Overall concentration of exports



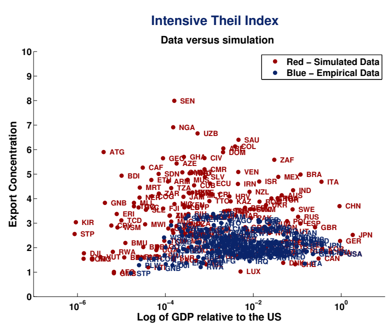
(b) Overall concentration of imports



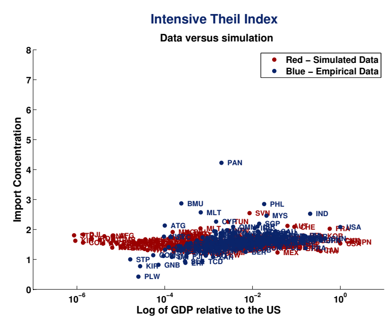
(c) Extensive margin of exports



(d) Extensive margin of imports

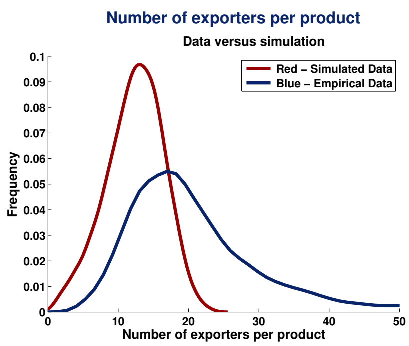


(e) Intensive margin of exports

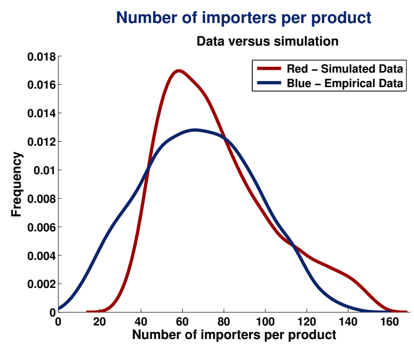


(f) Intensive margin of imports

Figure 1.4 – Simulated (in red) and empirical observed (in blue) export and import concentration versus GDP across 160 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 1.5 – The simulated (in red) and empirical observed (in blue) share of the number of products traded against the number of trading countries.

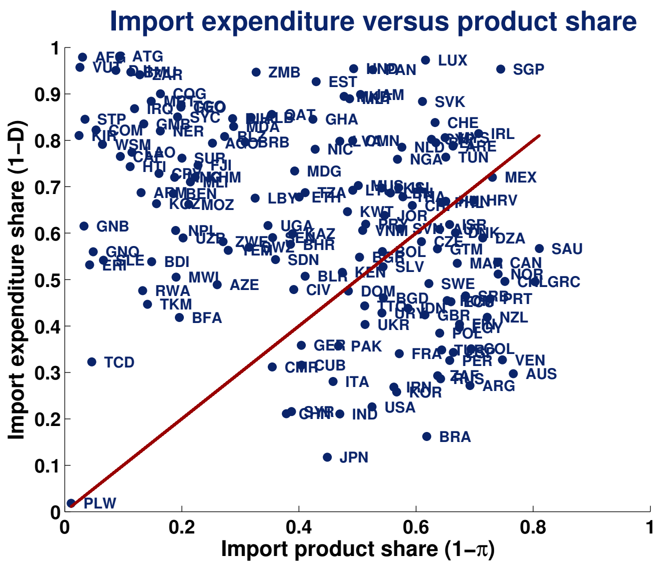


Figure 1.6 – The import expenditure share versus the import product share.

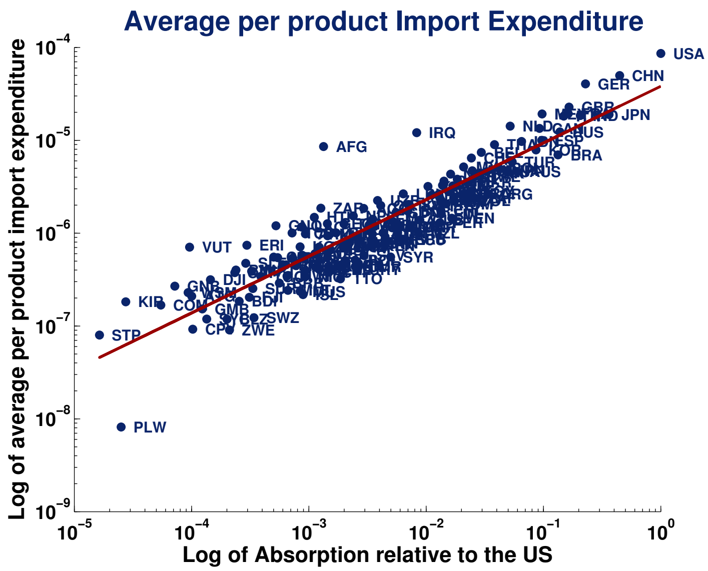


Figure 1.7 – The log of average per product import expenditure against log of Absorption.

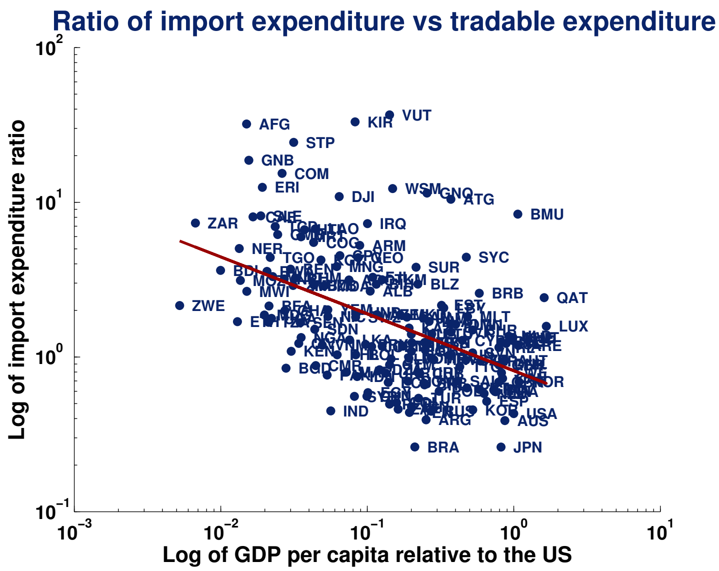


Figure 1.8 – The log of the ratio of average per product import expenditure with respect to average per product tradable expenditure against log of GDP per capita.

1.9 Tables

Table 1.6 – Country-Specific Technology and Trade Costs estimates

Country	Exporter FE	Standard error	Precent cost	Si	Standard error	$(\lambda_{US}/\lambda_i)^\theta$
USA	6,36	0,18	-53,47	0,84	0,13	1
AFG	-0,46	0,25	5,53	-3,06	0,19	193,42
AGO	-1,96	0,23	26,79	-0,97	0,16	23,5
ALB	-3,31	0,23	48,53	-0,12	0,16	9,63
ARE	2,98	0,19	-30,12	-0,71	0,14	2,6
ARG	2,19	0,19	-23,01	1,54	0,14	2,24
ARM	-3,14	0,22	45,55	0,2	0,16	9,61
ATG	1,12	0,45	-12,28	-3,72	0,3	12,93
AUS	3,29	0,19	-32,75	0,98	0,13	1,2
AUT	2,03	0,19	-21,83	1,24	0,13	0,77
AZE	-3,41	0,22	50,25	1,12	0,16	12,2
BDI	-3,45	0,24	51,47	-0,45	0,16	49,84
BEL	5,53	0,18	-48,63	-0,89	0,13	0,85
BEN	-3,11	0,23	45,41	-0,38	0,15	45,53
BFA	-4,45	0,23	70,73	0,6	0,15	36,85
BGD	0,96	0,2	-10,19	0,27	0,14	10,69
BGR	0,05	0,19	-0,59	1,01	0,14	2,79
BHR	-0,83	0,32	10,52	0,26	0,23	1,87
BIH	-3,57	0,24	53,34	1,1	0,17	6,66
BLR	-1,40	0,21	18,12	2,1	0,15	2,17
BLZ	-0,26	0,26	3,53	-1,77	0,18	8,17
BMU	-1,26	0,41	16,48	-1,91	0,28	5,66
BOL	-1,84	0,22	24,88	0,39	0,15	10,01
BRA	3,17	0,19	-31,64	1,71	0,13	2,22
BRB	-1,49	0,23	20,33	-0,91	0,16	5,95
CAF	-2,05	0,26	28,12	-1,11	0,19	21,31
CAN	4,10	0,18	-38,98	0,43	0,13	0,99
CHE	4,79	0,19	-43,90	-0,76	0,13	0,9
CHL	2,13	0,2	-22,51	0,48	0,14	2,34
CHN	5,11	0,18	-45,86	1,57	0,13	2,22
CIV	-0,12	0,2	1,13	0,06	0,14	8,73
CMR	-2,10	0,2	28,96	0,78	0,14	8,12
COG	0,87	0,23	-9,83	-2,63	0,17	16,64
COL	-0,04	0,19	0,63	1,13	0,13	5,89
COM	-3,06	0,29	44,10	-1,78	0,19	42,22
CPV	-3,16	0,32	46,15	-0,66	0,2	16,29
CRI	0,32	0,21	-3,62	0,06	0,15	3,43
CUB	-1,47	0,2	19,41	0,86	0,14	9,57
CYP	0,61	0,19	-6,89	-0,44	0,14	3,51
CZE	1,13	0,19	-12,91	1,37	0,13	1,04
DJI	-1,23	0,28	16,32	-2,99	0,2	50,23
DNK	2,57	0,18	-26,76	0,97	0,13	0,8
DOM	-1,12	0,2	14,60	0,65	0,14	3,72
DZA	-2,29	0,2	31,68	0,61	0,13	17,61
ECU	-0,18	0,2	2,25	0,57	0,14	6,51
EGY	0,42	0,19	-4,75	0,83	0,13	9,14
ERI	-4,87	0,26	79,16	0,12	0,19	43,83
ESP	3,76	0,18	-36,45	0,81	0,13	1,19

Country	Exporter FE	Standard error	Precent cost	Si	Standard error	$(\lambda_{US}/\lambda_i)^\theta$
EST	1,75	0,21	-19,23	-1,36	0,14	2,27
ETH	-1,73	0,2	23,10	-0,6	0,13	70,73
FIN	1,77	0,19	-19,42	1,73	0,13	0,59
FJI	-1,88	0,25	25,42	-0,36	0,17	5,57
FRA	4,56	0,18	-42,30	1,05	0,13	0,8
GBR	4,86	0,18	-44,32	0,57	0,13	1,06
GEO	-0,54	0,21	6,46	-1,25	0,15	15,46
GER	4,74	0,18	-43,54	1,17	0,13	0,65
GHA	1,14	0,2	-12,49	-1,78	0,14	17,47
GMB	-1,69	0,24	23,14	-1,99	0,17	30,89
GNB	-3,13	0,38	45,90	-0,89	0,27	34,18
GNQ	-3,99	0,28	61,56	0,39	0,19	3,92
GRC	0,73	0,19	-8,59	0,93	0,13	2,5
GTM	-1,41	0,21	18,61	0,41	0,14	6,43
HND	1,26	0,24	-13,99	-2,49	0,17	9,19
HRV	-0,60	0,19	7,28	0,92	0,13	2,53
HTI	-3,14	0,32	45,77	-0,5	0,23	26,46
HUN	0,43	0,19	-5,15	1,49	0,13	1,26
IDN	4,30	0,19	-40,26	0,21	0,13	4,69
IND	3,76	0,18	-36,12	1,03	0,13	6,78
IRL	3,90	0,18	-37,55	-0,47	0,13	0,78
IRN	-1,18	0,2	15,38	1,94	0,15	7,13
IRQ	-3,12	0,3	44,88	-1,13	0,21	224,32
ISL	0,08	0,21	-1,09	-0,18	0,15	1,15
ISR	1,26	0,19	-14,17	1,11	0,14	1,26
ITA	3,96	0,18	-38,00	1,27	0,13	0,8
JAM	0,76	0,21	-8,32	-1,7	0,15	6,45
JOR	-0,60	0,2	7,39	0,24	0,14	5,19
JPN	4,91	0,18	-44,65	1,95	0,13	0,48
KAZ	-0,28	0,21	3,08	1,08	0,15	4,17
KEN	-0,24	0,2	3,22	-0,06	0,14	20,53
KGZ	-3,04	0,24	43,58	0,39	0,16	10,95
KHM	-2,22	0,29	30,65	0,71	0,21	10,91
KIR	-2,77	0,39	38,98	-1,68	0,29	20,97
KOR	4,42	0,18	-41,23	1,4	0,13	0,73
KWT	-1,70	0,2	23,09	0,84	0,14	3,69
LAO	-3,15	0,29	46,04	0,54	0,23	11,92
LBN	-0,31	0,19	3,77	-0,23	0,14	7,97
LBY	-1,81	0,24	24,19	0,27	0,17	8,88
LKA	0,98	0,2	-11,03	-0,37	0,14	7,75
LTU	-0,24	0,2	2,69	0,6	0,14	2,6
LUX	1,44	0,25	-16,07	-0,65	0,2	0,86
LVA	-0,64	0,21	7,78	0,3	0,15	2,97
MAR	0,73	0,19	-8,11	0,39	0,14	5,1
MDA	-1,11	0,23	14,13	-0,33	0,16	8,17
MDG	-0,95	0,22	12,10	-0,93	0,15	20,18
MEX	3,42	0,19	-33,56	-0,1	0,13	3,27
MKD	-1,04	0,23	13,26	-0,73	0,15	5,07
MLI	-2,42	0,25	33,82	-0,45	0,17	43,51
MLT	0,30	0,22	-3,45	-0,68	0,16	1,64
MNG	-2,60	0,27	36,44	-0,51	0,19	10,12
MOZ	-1,13	0,21	14,84	-0,55	0,14	18,96
MRT	-0,58	0,24	7,23	-2,13	0,17	21,41
MUS	0,95	0,2	-10,48	-0,98	0,14	3,68
MWI	-3,87	0,23	59,69	0,29	0,15	34,15
MYS	5,40	0,19	-47,75	-0,74	0,14	1,64
NER	-1,89	0,23	25,67	-1,35	0,16	39,78
NGA	0,15	0,21	-1,37	-1,19	0,14	57,57
NIC	-1,13	0,22	14,67	-0,78	0,15	10,6

Country	Exporter FE	Standard error	Precent cost	Si	Standard error	$(\lambda_{US}/\lambda_i)^\theta$
NLD	5,66	0,18	-49,44	-0,88	0,13	1,02
NOR	1,83	0,19	-19,87	1,02	0,13	0,9
NPL	-3,03	0,23	44,04	0,37	0,16	18,27
NZL	2,54	0,19	-26,41	0,58	0,14	1,27
OMN	0,39	0,21	-4,70	-0,59	0,15	6,51
PAK	1,59	0,19	-17,08	0,9	0,14	8,28
PAN	2,82	0,23	-28,76	-2,16	0,17	6,67
PER	0,47	0,2	-5,45	1,17	0,14	3,64
PHL	2,33	0,2	-24,40	0,07	0,14	4,89
PLW	-9,10	0,4	197,33	4,52	0,32	0,19
POL	0,87	0,19	-9,93	1,78	0,13	1,57
PRT	1,76	0,19	-19,23	0,7	0,13	1,47
PRY	-1,36	0,22	17,79	0,6	0,17	7,95
QAT	0,60	0,21	-7,14	-0,62	0,15	2,69
ROM	0,18	0,19	-2,06	1,73	0,13	2,39
RUS	1,98	0,19	-21,33	1,89	0,13	2,15
RWA	-3,73	0,23	57,16	-0,15	0,15	67,74
SAU	1,34	0,19	-14,98	0,36	0,13	4,12
SDN	-2,46	0,2	34,35	-0,12	0,13	41,79
SEN	-0,69	0,21	8,70	-0,57	0,14	14,97
SGP	6,66	0,19	-55,17	-2,19	0,13	0,98
SLE	-0,49	0,49	5,65	-0,94	0,33	22,08
SLV	-1,74	0,21	23,31	0,42	0,15	4,95
SRB	-1,84	0,21	24,59	1,36	0,15	3,15
STP	-2,21	0,33	29,84	-1,89	0,24	30,19
SUR	-1,59	0,26	21,12	-0,75	0,18	3,21
SVK	1,67	0,2	-18,39	-0,18	0,14	1,73
SVN	-0,38	0,19	4,59	1,09	0,13	1,01
SWE	2,74	0,18	-28,26	1,41	0,13	0,66
SWZ	-0,81	0,24	10,26	0,15	0,19	3,87
SYC	-1,17	0,28	15,16	-1,46	0,2	3,51
SYR	-3,28	0,21	48,21	2,26	0,15	7
TCD	-5,68	0,26	98,20	0,93	0,18	52,38
TGO	-1,07	0,23	14,05	-1,56	0,15	32,81
THA	5,42	0,18	-47,82	-0,68	0,13	2,93
TKM	-4,02	0,26	61,53	1,08	0,19	12,48
TTO	-1,00	0,22	12,96	0,46	0,15	2,04
TUN	0,44	0,19	-4,70	0,01	0,14	3,76
TUR	1,92	0,18	-20,57	1,38	0,13	2,55
TZA	-0,25	0,2	3,46	-0,88	0,13	30,26
UGA	-1,79	0,21	24,15	-0,32	0,14	37,54
UKR	0,91	0,19	-10,55	1,75	0,14	3,01
URY	0,76	0,21	-8,63	0,44	0,15	2,72
UZB	-2,14	0,24	29,20	0,68	0,18	12,54
VEN	-0,46	0,2	5,82	1,35	0,14	4,18
VNM	2,46	0,19	-25,57	0,24	0,13	6,16
VUT	-0,93	0,34	11,75	-2,46	0,23	16,57
WSM	-2,40	0,3	32,82	-1,26	0,22	9,02
YEM	-2,67	0,22	37,58	0,39	0,15	31,04
ZAF	3,49	0,19	-34,26	0,48	0,13	2,25
ZAR	1,02	0,27	-11,18	-2,97	0,2	58,54
ZMB	1,85	0,26	-19,62	-2,55	0,17	18,9
ZWE	-1,06	0,22	13,97	0,16	0,16	9,86

Chapter 2

Entry barriers to International Trade: product versus firm fixed costs

2.1 Introduction

Fixed costs to export create entry barriers that restrict trading opportunities. Larger markets ease the relevance of fixed costs by allowing firms to slide down the average cost curve and produce at a more efficient scale. As a result more firms enter larger markets. Fixed costs can be of two types: either at firm or at the product level. The current view of the literature [1] is that fixed costs to export are mainly at the firm level, for example advertising a firm brand or setting up a distribution network. Given this cost structure, multi-product firms have a cost advantage and dominate international trade. Anecdotal evidence [2] suggests an alternative view where fixed costs operate at the product level, for example acquiring export/import licenses or technical barriers to trade. In this case, product entry is accompanied with lots of firm entry and international trade is characterized by many firms selling different varieties of the same product.

This paper develops an empirical framework to infer the dominant nature of fixed costs form the different effects they have on international trade patterns. I argue that the elasticities of the number of exporting firms and exported products with respect to destination market size informs on whether fixed costs operate at the firm or at the product level. Using detailed trade data from 40 exporting countries across 180 destination markets, I assess the relevance of fixed costs by testing for differences in entry elasticities. Within this framework, I then present empirical evidence consistent with the view that product fixed costs create spillovers effects that lead to higher firm entry and give grounds for trade policies promoting exports.

The starting point is to test whether the number of exporters and the number of exported products vary with market size in a significantly different way. In the case of fixed costs operating at the firm level, a firm pays a common fixed cost to advertise the firm brand or to create a distribution network in the export destination for all the products it exports. By spreading the fixed cost over more products, multi-product firms have a cost advantage over single-product firms, as in Feenstra and Ma [2007] and Eckel and Neary [2010]. The presence of spillovers effects through lower per product costs allows the firm

1. see for example Arkolakis and Muendler [2010], Eaton et al. [2011] and Bernard et al. [2011]
2. see for example Hausmann and Rodrik [2003] and Artopoulos et al. [2013]

to expand its product range with market size, i.e. economies of scope. As a result, few firms with many products enter large markets leading to a higher average number of products per firm. The testable implication is that the elasticity of products with respect to market size should be higher than the elasticity of firms.

Fixed costs at the product level are instead costs that firms have to incur in order to introduce a product into a destination market, i.e. acquiring an import license, meeting a product standard or advertise a product in the destination. The key point is that the firm that pays the product fixed cost creates a spillover to rival firms by lowering the fixed cost for subsequent exporters. Hausmann and Rodrik [2003] argue that export pioneers create an externality/spillover by making a considerable investment in attempts opening up foreign markets, cultivating contacts and establishing legal standards. The investment in these costly activities can then be used by domestic rival firms operating in the same product category. The spillover reduces the rival firms' fixed costs to export and facilitates entry into export markets. Due to the higher demand in larger markets, we expect that the export pioneer is more likely to create an externality for subsequent exporters in larger markets because of higher expected export revenues. Given this reasoning, there is relative more firm than product entry once market size increases. The testable implication is that the entry elasticity of firms with respect to market size should be higher than for products.

Using bilateral data for 40 exporting countries in 180 destinations, I find that the entry elasticity of the number of firms with respect to market size is significantly higher than the entry elasticity of the number of products. This holds for a broad set of countries at different levels of development. Two potential explanations for a higher firm than product elasticity are: either the average number of firms per product increases with market size or the average number of products per firm decreases with market size. The first effect points to more product varieties in larger markets and is consistent with product fixed costs. The second effect suggests that multi-product firms enter in small and large markets. However, in larger markets multi-product firms export less products compared to the small market because of more competition from single product firms, see Mayer et al. [2011]. This finding would be consistent with firm fixed costs. My results show that larger markets have on average more firms per exported product and that the number of products per firm does not vary with market size. This finding supports the claim that entry barriers operate on the product level.

Next, I build on the previous framework and present supportive evidence for spillover effects consistent with product fixed costs. Once the export pioneer pays the product fixed costs and introduces a new product into a destination market, rival firms benefit from lower fixed costs. To test this implication, I investigate how firm entry changes over time after a new product is introduced. Also, the willingness of an export pioneer to introduce a new product depends on the number of product market rivals because of business stealing effects, i.e. more firm entry increases competitive pressure and reduces prices. When regressing the number of exporters within a product category in a given destination on export prices and quantities, we expect that the number of exporters should be negatively correlated with prices and positively with quantity. The lower price indicates competitive pressure from the entry of product market rivals. The larger quantity captures the efficiency gains in production through economies of scale, i.e. firms slide down the average cost curve due to lower fixed costs. Furthermore, firm entry should depend on the product type. Exporters of differentiated products face less competitive pressure from product market rivals.

Using detailed product level data (4912 product categories), the results indicate that the first two years after a product is exported for the first time to a destination the firm entry rate within the product category increases significantly. The following years the firm entry rate is lower than the average entry rate and declines steadily over time. Within a product category, countries with more firms per export destination have significant lower export prices and sell a larger quantity. The quantity effect dominates the price effect indicating that lower fixed costs increase average export revenues of the firm. Taking the degree of product differentiation into account, I estimate the model for differentiated, less differentiated and homogenous products separately and test for significant differences in the estimated coefficients. I find that prices of differentiated products are less sensitive to firm entry, which points to lower competitive pressure in differentiated products. Overall, the results are consistent with fixed costs operating at the product level and suggest that rival firms benefit from spillovers that lower fixed costs and increase average export revenues per firm.

Understanding the nature of fixed costs is an important guide for effective trade policy among countries. This is because different set of policies can reduce product as opposed to firm fixed costs to encourage exports. For example, the exporting country can stimulate new product entry by advertising new products in destination markets through export promotion agencies.[3] This may lead to spillover effects that translate into higher level of firm participation and export growth. By subsidizing part of the product fixed costs, the government also increases incentives for firms to explore new export destinations and offsets part of the free riding from rival firms.

More generally, the existence of fixed costs to export implies that trade policy can affect market structure. When conducting policy experiments in the form of a reduction in trade costs, it is standard in the international trade literature to consider only a fall in marginal costs and evaluate the resulting impact on the patterns of trade and consumer welfare. My results suggest that fixed costs at the product level are an important entry barrier to international trade. In addition, a central aspect of free trade negotiations is the reduction of these costs by alleviating technical barriers to trade and establishing common product standards, see the current EU-US free trade negotiations. Kehoe and Ruhl [2013] provide empirical evidence that trade liberalization increases product entry and leads to significant growth in export revenues from these products. Thus, neglecting changes in these barriers underestimates the impact of trade reforms.

My work contributes to the empirical international trade literature that analyses the relationship between market characteristics, fixed costs and product entry. Hummels and Klenow [2005] focus on characteristics of the exporting country when studying the extensive product margin. They suggest that trade models featuring product fixed costs can reconcile the fact that conditional on exporting a country exports a given product only to a subset of destination markets. Contrary to them, my argument for market specific product fixed costs is based on the fact that the number of products increases significantly in destination market size. In addition, the results in this paper show that the presence of product fixed costs entail a spillover that facilitates firm entry. Hummels and Klenow [2005] also show that a larger the

3. see Görg et al. [2008] and Lederman et al. [2010] for empirical evidence on export promotion agencies

economy exports each product to more destinations. Consistent with their argument, I find higher entry elasticities for larger exporting economies, i.e. as destination market size increases, larger economies start to export relative more products than smaller ones. Combined these results suggests that market specific product fixed costs characterize international trade patterns and that home market size matters in overcoming fixed costs.

Hausmann and Rodrik [2003] argue that export pioneers create spillovers by making investments in attempts to open foreign markets, cultivating contacts, establish distribution chains and other costly activities that can be used by rival firms within the same product category. Rivals may also acquire knowledge about the potential demand of their own products in the foreign market once they observe the success of the pioneer, see Eaton et al. [2012]. By analyzing firm entry within products over time and across countries, I find supportive evidence of these spillovers. In line with lower fixed costs, firm entry increases significantly the year after an export pioneer introduces a product into a destination. The lower fixed cost allows rival firms to exploit scale, increase their export revenue and the survival probability in international markets. These results are consistent with micro evidence of spillovers among exporters as found in the case of France (Koenig [2009] and Koenig et al. [2010]) and Argentina (Artopoulos et al. [2013]).

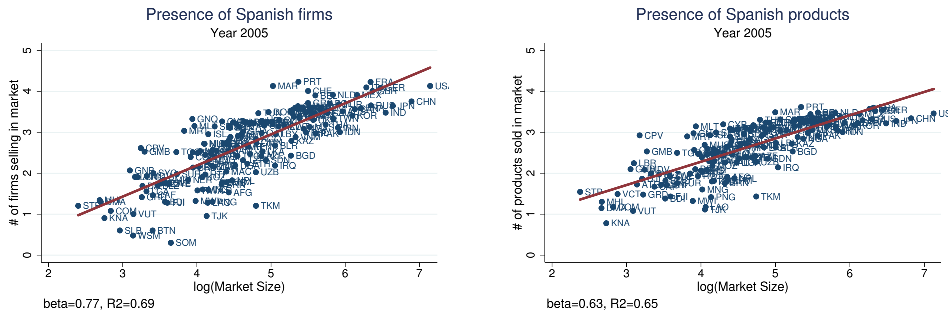
The analysis also contributes to the international trade literature analyzing the empirical relationship between market size and firm entry, for single product firms, see Helpman et al. [2008], Melitz and Ottaviano [2008], Arkolakis [2010], Eaton et al. [2011], and multi-product firms, see Arkolakis and Muendler [2010] and Bernard et al. [2011]. The paper most closely related to this one is Eaton et al. [2011]. Using a monopolistic competition model of heterogeneous firms with CES preferences and fixed costs, Eaton et al. [2011] argue that the variation in the number of French exporters with respect to destination market size informs on fixed costs of exporting at the firm level. This paper builds on their basic empirical insight, looking at the elasticity of firm penetration, and questions whether fixed costs operate at the firm or product level. My results suggest that once we depart from Eaton et al. [2011]'s assumption of one firm produces one product, product fixed costs offer an alternative view consistent their empirical result.

Arkolakis and Muendler [2010] and Bernard et al. [2011] focus on multi-product firms and the determinants of their product scope with respect to destination characteristics. They generalize the cost structure of Eaton et al. [2011] by introducing firm specific product fixed costs. In comparison to these papers, this paper considers product fixed costs that are independent of the firm. Once an export pioneer pays the product fixed cost and introduces the product in an export destination, she creates a spillover that lowers fixed costs for rival firms within the same product category. Note that if fixed costs are parameterized accordingly, the framework of Arkolakis and Muendler [2010] and Bernard et al. [2011] can account for differences in the firm and product elasticity. The key difference with respect to this paper is that their analysis does not allow for spillover effects across firms once new products enter export markets. Empirical evidence suggests that these spillover effects are quantitatively important in explaining the entry behavior of firms.

The rest of the paper is organized as follows. Section 2 describes the conceptual framework. Section 3 presents the methodology together with the testable implications. Section 4 presents the data with the relevant summary statistics and the empirical results. Section 5 illustrates an empirical framework to shed further light on the presence of spillovers induced by product fixed costs. Section 6 concludes.

2.2 Empirical framework

I start my investigation with an assessment of the destinations that exporting firms reach and the characteristics of the destinations that attract many exporters. First, take the perspective of the largest exporting country in my sample, Spain, and its firms. Following Eaton et al. [2011], Figure 2.1(a) plots the log of the number of Spanish firms selling to a particular market d against the log of destination market size proxied by GDP. The number of firms selling to a market tends to increases with the size of the market. A regression line establishes a slope of 0.77 and an $R^2 = 0.69$. Eaton et al. [2011] interpret the positive relationship between firm penetration and market size as evidence of market specific fixed cost. Larger markets offer more demand and thus it is easier for firms to recover fixed costs. As Bernard et al. [2003] show, other trade models based on variable trade costs without fixed costs can also account for the fact that firms select themselves into exporting. However, the authors also note that these models are not able to generate



(a) Firm entry by market size for Spain in 2005

(b) Product entry by market size for Spain in 2005

Figure 2.1 – Number of Spanish exporting firms and number of Spanish products exported versus market size in destination d.

Sources: Exporter Dynamics Database World Bank. Figure (1a) number of firms, Figure (1b) number of 6 digit HS products per destination. Market size is absorption of a country's manufacturing sector. The slopes of the fitted lines are 0.77 (standard error 0.038) for exporting firms from Spain and 0.63 (0.034) for exported products from Spain.

both the observed relative size of total revenues of exporters, compared to non-exporters, and the strong selection with respect to destination market size into exporting. To account for these empirical regularities, international trade models assume additional exporting costs in the form of a market specific fixed costs to export.

An alternative view of the relationship in Figure 2.1(a) is that fixed costs operate at the product rather than on the firm level. To investigate this idea further, Figure 2.1(b) repeats the previous graph but instead of the log number of firms it plots the log number of products that Spain exports to a destination against the market size of the destination along the horizontal axis. The number of products exported to a destination increases systematically with market size, $R^2 = 0.65$, and an elasticity of 0.63. Following the argument of Eaton et al. [2011], an explanation that reconciles the relationship in Figure 2.1(b) is the presence of market specific fixed costs at the product level. Exporting products is only possible at a huge expense in fixed costs and the demand for most of the products in the destination is not sufficient to export all of them profitable.

While both figures display a positive relationship between entry and market size, the slope of the log number of products with respect to market size is significantly lower than in the case of firms. The difference in the elasticities implies that international trade models based on the assumption of one firm produces one product are inadequate in addressing the number of exporters and exported products in destination markets. Departing from Eaton et al. [2011], this paper considers a framework where firms can produce multiple products and a product can be produced by multiple firms. Within this framework, I then ask the following questions: To what extent prevent fixed costs the entry of firms and products in international trade? Are they more prevalent on the firm or on the product margin? To answer these questions, I evaluate how the number of firms and products varies with destination market size controlling for origin, time and bilateral characteristics. I attribute differences in the entry behavior across markets due to fixed cost operating either at the firm or product level. Before describing the empirical model, I define the cost structure and derive testable implications.

2.2.1 Fixed cost at the firm level

Under entry barriers to export at the firm level, I consider market specific fixed costs that the firm needs to pay in order to export its products to a destination market. Such costs can take the form of information costs to acquire knowledge about the market in a destination (Chaney [2011]), advertising costs to establish the firm brand (Arkolakis [2010]) or adaptation costs in the form of building a distribution network. Additional sources of adaptation costs can be cost in order to accommodate to business practices in the export destination (Artopoulos et al. [2013]), or legal fees to establish eligibility of an exporting firm/company by the importing country (Khanna et al. [2009]). The key characteristic of the firm fixed cost is that incurring the cost only benefits the firm.

The presence of fixed costs at the firm level implies cost advantages for multi-product firms because they can spread these costs across more products, as in Feenstra and Ma [2007] and Eckel and Neary [2010]. This lowers the firm's average costs per product and increases its competitiveness relative to single product firms. Multi-product firms benefit from economies of scope. Larger markets offer more demand, increase the cost advantage and attract relative more multi-product firms than in smaller markets. The presence of economies of scope may be one explanation of why multi-product firms are dominant in international trade.[4]

To summarize, in larger markets the expected firm revenue is higher allowing the firm to run down their average costs curve. The lower costs spurs firm entry and each firm produces at a larger scale. Given that some of the firms are multi-product firms, the larger market gives them an additional cost advantage. As market size increases, relative more products enter because multi-product firms either expand their product range or participate more in international trade relative to single product firms. Given this reasoning, *one should observe more product entry in comparison to firm entry resulting in more products per firm in larger markets.*

2.2.2 Fixed cost at the product level

Fixed costs at the product level can take the form of technical barriers to trade (in the form of products standards or certification procedures to ensure the quality) or product advertising. Firms have to pay a fixed costs to introduce the product in the destination because consumers are not aware of the product. Technical barriers to trade imply modifications to the offered product in order to customize it to particular local tastes or legal requirements imposed by national consumer protection laws. Costs also arise from the translation of foreign regulations and/or hiring of technical experts to explain foreign regulations. Note that the use of technical barriers to trade is subject to Agreement on Technical Barriers to Trade administered by the WTO.

The key implication of the product fixed costs is that once a product is established in an export market, many firms start to export differentiated varieties of that product. Incurring the fixed cost to introduce a new product induces a spillover that lowers fixed costs for all firms within the product category. One reason is that ex ante consumers are unaware of the existence of the

4. Based on U.S. trade data in 2000, Bernard et al. [2011]) show that firms exporting more than five products at the HS 10-digit level make up 30 percent of exporting firms and account for 97 percent of all exports. Looking at Brazilian exporter data in the year 2000, Arkolakis and Muendler [2010] find that 25 percent of all manufacturing exporters ship more than ten products at the internationally comparable HS 6-digit level and account for 75 percent of total exports.

product. Once a firm introduces the product successfully in the destination market, consumers demand the product allowing firms to export differentiated varieties of that product. To access the export market, firms can share the fixed product costs in the form of setting up common distribution networks to promote their products jointly (for example, a US car dealerships that sells different brands of German cars). Another form of cooperations are trade associations formed to foster collaboration between companies within a specific product category in order to define common product standards, to advertise their products to foreign consumers or to lobby governments for favorable trade policy, i.e. through export promotion policies, see Lederman et al. [2010] and Görg et al. [2008] for the empirical evidence.

Instead of many firms sharing the product fixed cost, also a single firm can introduce a product into an export market. By doing so, this export pioneer creates a spillover/externality for other firms producing the same product. Product market rivals benefit from lower fixed costs because the export pioneer opened up a foreign market, established contacts and/or distribution chains and invested in other costly activities which they can use. Rival firms may also acquire knowledge about the potential demand of their own products in the foreign market once they observe the success of the pioneer. Khanna et al. [2009] study the concrete example of Metro Group a German retail company that fought years to have access to the Indian market. Once the Foreign Direct Investment permit was granted, rival retail firms like Wal-Mart and Tesco entered immediately by benefiting from the created legal framework and the observed business opportunities in the Indian retail market.

To summarize, under the presence of fixed costs at the product level, the entry of a product is associated with many firms. Due to the higher demand in larger markets, we expect more cooperation among firms in paying the fixed. The first entrant is more likely to create a positive externality for rival firms in larger markets because the expected firm revenue is higher despite the following entry of rival firms. The product fixed cost implies that multi-product firms do not have a cost advantage in larger markets. Under all scenarios, we expect that there is substantially more firm entry than product entry once the market size increases. The testable implication is that *the entry elasticity of firms with respect to market size should be greater than for products implying that the number of firms per product is higher in larger* *markets.*

In the next section, I explain how I distinguish empirically the nature of fixed costs, namely whether they operate at the firm or at the product level. The key element for this distinction relies on the comparison of firm entry and product entry elasticities as market size increases.

2.3 Methodology

To analyze the nature of fixed costs, proceed as follows. First, decompose export revenues from country c to destination d in year t, $X_{d,c,t}$, into the following firm components. Note that the same decomposition also holds for products.

$$X_{d,c,t} = \pi_{d,c,t} X_{d,t} = N_{d,c,t} \bar{x}_{d,c,t} \tag{2.1}$$

$X_{d,t}$ is the market size measured by GDP of destination d in year t, $\pi_{d,c,t} = X_{d,c,t}/X_{d,t}$ is the import expenditure of destination d on goods from country c, $N_{d,c,t}$ is the number of firms (or the number of products) that export from c to d and $\bar{x}_{d,c,t}$ is the average export revenue per firm (or per product) from c to d in t.

To investigate the relationship with exports and market size on the different margins, rewrite equation 2.1 as:

$$X_{d,t} = \left(\frac{N_{d,c,t}}{\pi^{b}_{d,c,t}} \right) \left(\frac{\bar{x}_{d,c,t}}{\pi^{1-b}_{d,c,t}} \right)$$

and taking logs, we get

$$\log X_{d,t} = \log N_{d,c,t} - b \log \pi_{d,c,t} + \log \bar{x}_{d,c,t} - (1-b) \log \pi_{d,c,t} \tag{2.2}$$

Split equation 2.2 into two expressions and evaluate how the extensive margins (the number of exporters)

$$\log N_{d,c,t} = b \log \pi_{d,c,t} + \gamma \log X_{d,t} \tag{2.3}$$

and the intensive margin (the average export revenue per firm)

$$\log \bar{x}_{d,c,t} = (1-b)\log \pi_{d,c,t} + (1-\gamma)\log X_{d,t} \tag{2.4}$$

change with market size.

The parameter of interest is γ. Consider the following two possibilities:

1. If $\gamma = 0$: In the absence of fixed costs and given positive demand for a product or a brand, any firm will find it worthwhile to enter. In this case, we expect that the number of firms and products does not change with market size. Models in international trade that feature this setting are of the Armington type, i.e. Anderson and Van Wincoop [2003]. In these models only the intensive margin matters since the number of exporters per market is assumed to be fixed.
2. If $\gamma > 0$: Under the presence of such costs, firms operate under increasing returns to scale. Firms enter the market until the expected profit is zero, i.e. expected export revenue equals fixed costs. This free entry condition determines the number of firms per market. In larger markets, firms can take advantage of the higher demand by sliding down the average cost curve and sell at lower prices. Thus, the number of firms will be increasing in market size .

To assess differences in firm level and product level entry barriers, we test for significant differences in the estimated elasticities with respect to market size using the following regression specification

$$\log N_{d,c,t} = \alpha + b\log \pi_{d,c,t} + \gamma \log X_{d,t} + d_{c,t} + \epsilon_{d,c,t} \tag{2.5}$$

which restates equation 2.3 and includes origin country-year dummies. The import expenditure variable $\pi_{d,c,t}$ captures the taste that a particular destination d may have for goods from country c. We expect that the higher the expenditure share, the higher the propensity to export in a market. In this basic specification, it proxies also for all other factors, like distance, that determine market entry other than market size. In the robustness section I enrich the model and include further control variables that maybe correlated with market size and the entry decision. Equation 2.5 is estimated separately for products and firms as dependent variables in order to obtain separate entry elasticities (gamma parameter) with respect to market size for each of these two components.

— If product entry (P) is larger than firm entry (F), i.e. ($\hat{\gamma}^P > \hat{\gamma}^F$), then I interpret this as evidence suggestive of fixed costs operating at the firm level
— If firm entry (F) is larger than product entry (P), i.e ($\hat{\gamma}^F > \hat{\gamma}^P$), then I interpret this as evidence suggestive of fixed costs operating at the product level

If the entry elasticities are not significantly different from each other, then within this framework we cannot distinguish whether fixed costs operate on the firm or on the product level.

2.4 Data and descriptive statistics

To build the empirical evidence, I use the Exporter Dynamics Database from the World Bank, see Cebeci et al. [2012]. It contains firm characteristics per destination and per product for 40 exporting countries for the period 1997 to 2010. [5] Following the literature, see Broda and Weinstein [2006], I consider a 6 digit HS code per country as a product category and refer to individual firm products within the product category as varieties of the same product. Given this perspective, a product can be exported by multiple firms and a firm can potentially export multiple products. Firms can be viewed as providing their brand and the brand in turn provides the platform for specific products to be launched. The Exporter Dynamic Database does not contain information on the "Oil and Fuels" sector, HS code 27, leaving a total of 4912 tradable products for each country.

To examine product and firm entry into export markets, I include distance, common border, market size, income per capita and total import expenditure as destination characteristics. Distance and border measures come from Centre d'Etudes Prospectives et d'Informations Internationals (see Mayer and Zignago [2011]) and are in kilometers from capital city in country i to capital city in country j, calculated by the great circle method. Openness, market size and income measurements, defined as GDP and GDP per capita, are taken from the Penn World Table, see Heston et al. [2009]. Data on total

5. I exclude Botswana, Brazil, Egypt, New Zealand and Kuwait because of missing firm characteristics by export destinations. The appendix contains a complete list of the countries used.

c.i.f. import expenditure spend by destination on exporters goods is taken from the Comtrade data set collected by the United Nations.[6] In total the baseline sample covers 40 exporting countries and 180 destination markets.

Table 2.3 in the appendix describes the summary statistics of the combined dataset. The average number of exporters in a destination across all 40 exporting countries is 344 and the average number of exported products per destination is 298. Since firms can export multiple products and a product can be exported by multiple firms we can decompose the extensive margin of exports further. Line 3 in Table 2.3 shows the average number of products per firm is 2.5 suggesting that the majority of firms are multi-product firms. The average number of firms per products is 2.1 implying that strategic interactions between exporters from the same origin country are important. An export pioneer has to take into account the effect of a potential spillover/externality on product market rivals when opening up an export market. Overall, under the assumption of each firm exports a unique product we neglect important interaction between products and firms. In the majority of destinations, a firm sells more than one product and a product is exported by more than one firm.

Table 2.4 displays the results from the estimation of specification 2.3. Focusing on columns (1) and (2), we see that both, the number of firms and products, are increasing in destination market size and import expenditure share. In comparison to the literature, the firm entry elasticity of 0.40 wrt to destination market size is significantly lower than values found in other papers. Bernard et al. [2011] report a value of 0.70 for the United States in

6. To construct import expenditure shares, I use data from the Penn World Table and the Comtrade database. To avoid any potential measurement errors in the exchange rate when combining nominal values from the 2 dataset, I compute the import expenditure share of destination d on goods from country c, $\pi_{d,c}$, as follows. Using the Comtrade data set, I first compute the share of imports with respect to total trade flows. More precisely, I divide bilateral cif imports, $X_{d,c}$, by the sum of total fob exports plus total cif imports for each country, $(Im_p + Ex_p)$. From the Penn World table, I then take openness defined as total exports plus total imports divided by GDP. Hence, I can calculate the share of total cif imports expenditure with respect to GDP as:

$$\pi_{d,c} = \left(\frac{X_{d,c}}{Imp_d + Exp_d} \right) \left(\frac{Imp_d + Exp_d}{X_d} \right)$$

the year 2002 and Eaton et al. [2011] report an elasticity of 0.66 for France in the year 1992 and 0.68 for Denmark and Uruguay in 1993.[7] The results are more comparable to Bernard et al. [2011] as I also use total GDP as a measure of market size whereas Eaton et al. [2011] use manufacturing absorption.[8] Although there are significant differences in the point estimate of the entry elasticity with respect to the literature, all values are significantly below 1 implying that average export revenues increase with market size.[9]

Focusing on differences in the elasticities with respect to market size, the entry elasticity for firms is higher than for products suggesting that fixed costs at the product level are the dominant form of entry barriers, i.e. $(\gamma^F > \gamma^P)$. To test for significant differences, I estimate equation 2.5 with the dependent variable being number of firms and the number of products jointly within a Seemingly Unrelated Regressions (SUR) model. This estimation method allows for correlation in the entry determinants of firms and products. I reject the null hypothesis of no differences at the 1 percent level.

In column (3) and (4) in Table 2.4, I include additional control variables to see how entry behavior varies with respect to destination characteristics. We run the following regression

$$\begin{aligned} \log(N_{d,c,t}) &= \alpha + b_1 \log(\pi_{d,c,t}) + \gamma_1 \log(X_{d,t}) + \beta_1 \log(y_{d,t}) \\ &+ \beta_2 \log(dist_{d,c}) + \beta_3 \log(b_{d,c}) + d_{c,t} + \epsilon_{d,c,t} \end{aligned} \tag{2.6}$$

where entry of firms and products depends now on GDP per capita in desti-

7. I do not have data for the countries mentioned and can not compare the results by running the same regression for the respective countries.

8. If I use manufacturing absorption as a proxy for market size I obtain a firm entry elasticity of 0.45. Absorption is calculated from gross manufacturing output plus imports minus exports. Due to data limitations on gross manufacturing output the number of destinations shrinks to 150.

9. Below, I provide a sensitivity analysis where I investigate differences in the entry elasticities. The analysis shows that the entry elasticities increase with home market size implying that larger economies have higher entry elasticities. The reason why my estimate of firm entry is lower compared to the literature lies in the fact that my sample consists predominately of small economies compared to the literature and therefore biasing the estimate downwards. Taking the estimated relationship between home market size and entry elasticity from below, the results imply a firm entry elasticity of 0.74 for the United States and 0.62 for France.

nation d, $y_{d,t}$, the distance between trading partners, $dist_{d,c}$, and whether the countries share a common border, $b_{d,c}$. We expect that richer countries spend more per product and hence entry rates should increase in GDP per capita, $\beta_1 > 0$. Similarly, sharing a border and being close to each other increases the demand for products because of lower transportation costs. We expect entry decrease in distance, $\beta_2 < 0$, and increase when sharing a border, $\beta_3 > 0$.

Columns (3) and (4) in Table 2.4 confirm previous results. All coefficients are statistically significant. Based on equation 2.6, more firms and products enter in markets with a higher level of GDP per capita. Note that GDP is the product of population and GDP per capita. Thus, the effect of the log of population on entry is γ_1 and the effect of the log of income per capita is $\gamma_1 + \beta_1$. Since $\beta_1 > 0$, income per capita is more important than population for firm and product entry. Distance has a negative effect on entry implying that less firms enter in distant markets. Overall, column (3) and (4) show that the entry elasticity of firms is statistically significantly higher than for products even after controlling for income per capita and geography. Interestingly, the effects of expenditure shares, distance and sharing a border on firm entry are not significantly different from product entry.

2.4.1 Discussion of results

A firm can produce multiple products and a product can be produced by multiple firms. Depending on the ratio of firms to products in the small market, the reason for a higher firm than product elasticity can have two potential explanations in relation to product and firm fixed costs.

One explanations is that the number of firms per product increases with market size, which is consistent with fixed costs at the product level. Because of the low demand in small markets, firms export few products to these destinations. Also, export pioneers pay the product fixed cost only for product categories where they face little competition because subsequent entry of rival firms due to lower costs would vanish all its profits. Larger markets offer more demand for each product and the number of firms per product increases.

An alternative explanations is that the number of products per firm decreases with market size. Suppose few multi-product firms export many products to small markets. As market size increases, more firms are able to pay the firm fixed cost. Most of these firms are single product firms and they enter in product categories that multi-product firms export to the small market. This intensifies competition and forces multi-product firms to reduce their product range. Mayer et al. [2011] emphasize this mechanism. In this case the firm elasticity is higher than the product elasticity but the implication would be consistent with fixed costs at the firm level.

To distinguish between the two effects, I use equation 2.5 and regress the average number of firms per product and the average number of products per firm on market size and other destination characteristics. Column (3) and (4) in Table 2.8 of the appendix contain the results. I find that the number of products per firm is independent of market size, in accordance with the findings of Arkolakis and Muendler [2010] in the case of Brazil. On the other hand, the number of firms per product increases significantly in larger markets. Higher demand in larger economies reduces average costs of firms and leads to more entry. These findings support the claim that entry barriers operate on the product level.

The decomposition of exports into extensive and intensive margin, (equations 2.3 and 2.4), offers an alternative view on the mechanism behind fixed costs. Consider for a moment a model with free entry. Firms enter the market until the expected profit is zero, i.e. expected export revenues equal marginal costs plus fixed costs. This condition determines the number of exporters per market. The fact that entry elasticities are smaller than 1 implies that export revenues increase in market size, see equation 2.4. If average export revenues increase with market size, then the model implies, under the assumption of constant markups, that revenues are higher in larger markets because fixed costs are higher, for example setting up a distribution network is costlier. In this case we would have a positive correlation between entry and market size because market size proxies for fixed costs. To analyze whether the positive entry elasticities are triggered by a correlation between market size and fixed costs, I use additional control variables ($F_{d,t}$) that proxy for fixed costs. The resulting regression equation becomes:

$$\begin{aligned} \log(N_{d,c,t}) &= \alpha + b_1 \log(\pi_{d,c,t}) + \gamma_1 \log(X_{d,t}) + \beta_1 \log(y_{d,t}) + \beta_2 \log(dist_{d,c}) + \\ &+ \beta_3 \log(b_{d,c}) + \beta_4 \log(F_{d,t}) + d_{c,t} + \epsilon_{d,c,t} \end{aligned} \tag{2.7}$$

We expect that the coefficient β_4 is negative, i.e. higher fixed costs decrease the presence of firms. Important is the coefficient on γ_1. If $\hat{\gamma_1}$ differs from $\tilde{\gamma_1}$ previously estimated in Table 2.4 then fixed costs are correlated with market size. To assess the relationship between market size and the proxies of fixed costs, we use the fact that $\tilde{\gamma_1} = \hat{\gamma_1} + Corr(F_{d,t}, X_{d,t})$. If larger markets have higher fixed costs, then the estimated coefficient of market size should be lower given the presence of fixed costs.

To proxy fixed costs, I include Urban population (% of total), Land area (sq. km), Container port traffic (TEU: 20 foot equivalent units), Rail lines (total route-km), number of internet and cell phone subscribers (per 100 persons) and Electric power consumption (kWh per capita) from the World Development Indicators dataset provided by the World Bank. Urban population and land area proxy for retail distribution costs. A higher percentage of urban population facilitates distribution. On the other hand a larger land increases the costs to reach consumers. Rail and container port traffic proxy for transportation infrastructure. While transportation costs are also part of marginal costs, I use them as proxies for infrastructure fixed costs.[10] The number of internet subscribers controls for networking and communication costs. Finally, energy consumption proxies for higher retail costs. Due to missing observations, the sample reduces to 11096 observations.

Table 2.9 in the appendix reports the detailed results for each dependent variables. Note that better infrastructure proxied by container port traffic and km of rail lines increases entry both entry of firms and products. Land size and energy consumption reduce the entry of firms and products. A larger area requires more distribution costs and a higher energy consumption points to more fixed costs. Note that the elasticities of the number of firms and products with respect to destination market size decrease significantly. The reason is that market size is positively correlated with distribution costs, i.e. larger market have higher fixed costs and thus reduce the importance of market size on fixed costs. Overall, the firm elasticity is still significantly higher than the product elasticity suggesting that fixed costs operate preliminary on the product rather than firm level even when we control for "observable" fixed costs.

10. Removing rail lines and container port traffic from regression 2.7 does not change the results.

This paragraph contains alternative explanations for the positive relationship between firm entry and market size. Arkolakis [2010] provides a theory for firm entry based on marginal costs in form of a market access costs rather than fixed costs. Arkolakis and Muendler [2010] extend the analysis by relaxing the assumption of one firm produces one product. To reconcile the product and firm margin within the Arkolakis and Muendler [2010] framework, one needs to assume that a firm has to incur a marketing cost for each product it wants to export to each destination, otherwise firms would benefit from economics of scope. While in Arkolakis [2010] and Arkolakis and Muendler [2010] firms choose the marketing cost, Chaney [2011] develops a model based on information frictions. Firms have to search for foreign trading partners in order to trade. This characterizes a dynamic formation of an international network of importers and exporters. Because larger markets have more contacts, the number of exporting firms will increase in markets size. However, firms have to find buyers implying that the search costs occur at the firm level and not at the product level. One might expect that in the case of a match, the exporters will sell all its products. The finding that product fixed costs are the dominant entry barrier suggests that also demand considerations are an important factor in the entry decision of firms. Armenter and Koren [2009] develop a model where demand for products is uncertain. Eaton et al. [2012] develop a model with search and demand uncertainty. Exporters have to search for potential buyers in destination markets. The success in selling to a buyer reveals information about the appeal of the seller's product in the market, affecting the incentive to search for more buyers, so importers learn about the product. The combination of search and demand uncertainty is likely to replicate the above results.

The key aspect of my analysis is that the product fixed cost is not firm specific. Bernard et al. [2011] consider product fixed costs at the firm level. While their analysis focuses on multi-product firms and the determinants of firm's product scope after trade liberalization, their model allows for differences in the elasticities of firms and products with respect to market size. Parameterized accordingly, the quantitative model of Bernard et al. [2011] can account for differences in the firm and product elasticity. The central difference with respect to this paper is that their analysis does not allow for spillover effects across firms once new products enter export markets. In the following section, I present empirical evidence suggestive of spillovers and show that these effects are quantitatively important in explaining the entry behavior of firms.

2.5 Firm entry within products

In the previous part I presented evidence consistent with fixed costs operating at the product level. An important implication is that product fixed costs cause spillover effects that lower costs for subsequent exporters once the product entered a destination market. To shed light upon this mechanism, I analyze how firm entry evolves over time after a product enters a destination for the first time. Based on the results, I then present additional empirical evidence supportive of product fixed costs inducing spillovers that facilitate firm entry through lower fixed costs.

Given the definition of product fixed costs, we expect that once a firm successfully introduces a product in a destination market many firms will follow. To test the effect, I investigate how the entry rate of firms from an exporting country within a product category in a particular country varies over time. I define entry of a new product k from country c in a destination d at time t if the product is not exported in any period prior to the year of the first entry. The first year of product data I observe is 1995 and the first year of firm entry is 1998. Therefore, I will focus only on products that have not been exported to a destination prior to 1998. Another issue with the data is that the dataset does not contain information that is origin - destination - product - year specific, i.e. we do not know how many exporters from a particular country sell a particular product in a particular destination in a given year. To address this problem, I specify 2 regression models. In the first regression I analyze the firm entry rate aggregated over all products within a destination. In the second model I consider the firm entry rate per product aggregated over all destinations.

The first regression model analyzes the entry rate of firms from country c exporting to destination d at time t:

$$N_{d,c,t} = \sum_{s=1}^{6} \alpha_s l_{s,k,d,c} + d_{c,d} + d_{c,t} + d_{k,t} + \epsilon_{k,d,c,t}$$

The firm entry rate, $n_{d,c,t}$, is defined by the number of new exporters. I regress the firm entry rate on the a set of dummies, ($l_{s,k,d,c}$), that capture the firm entry rate over time after a product is exported for the first time to a destination. I set the dummy $l_{s,k,d,c}$ equal to 1 if product k from country c is exported to destination d s years after the product is introduced. The coefficient α_s captures the difference to the average firm entry rate in year s after the product is introduced. Given this specification, we expect that the entry rate increases significantly right after a product is introduced in an export market, i.e. $\alpha_1 > 0$. To test whether $\alpha_1 > 0$, I include a large set of control dummies: destination-origin ($d_{c,d}$), origin-time ($d_{c,t}$) and product-time ($d_{k,t}$) specific dummies. Origin-destination dummies control for geography. The origin-time dummies control for any origin country specific effects that generates easier firm entry into international markets, for example institutions, infrastructure, etc. Product-time dummies account for product demand effects common across countries .

In the second regression model, I analyze the firm entry rate within a product group across destinations. I estimate the following equation:

$$N_{k,c,t} = \sum_{s=1}^{6} \beta_s l_{s,k,d,c} + d_{c,k} + d_{d,t} + d_{c,t} + \epsilon_{k,d,c,t}$$

The firm entry rate, $n_{k,c,t}$, is defined by the number of new exporters of product k from country c in year t. I regress the firm entry rate on the same set of time dummies ($l_{s,k,d,c}$). The only difference is that I include origin-product ($d_{k,c}$) and destination-time ($d_{d,t}$) fixed effects instead of destination-origin ($d_{c,d}$) and product-time ($d_{k,t}$) dummies. The origin-product dummies account for supply side effects. For example, firm entry may be higher because a country is very productive in producing a particular product. The destination-time dummies control for macroeconomic conditions in the destination common to all products.

Table 2.5 plots the results of the 2 regression specifications. Average firm entry is given by the constant. The year dummies describe the estimated time effects on firm entry after a product is exported for the first time with respect to the mean. Looking at the coefficient of year_1 and year_2, firm entry increases significantly the first 2 years and then becomes either negative (column (1)) or insignificant (column (2)). Dividing the estimated coefficient by the average, we obtain that the entry rate in a destination increases by 2.5

percent and by 7 percent within the product group one year after a product is introduced. Given that the average number of firms per destination is 343, as shown in the summary statistics, the number of new firms in a destination increases on average by 8.6 firms. On the other hand, the average number of exporters per product is 27 implying that 2 additional new firms start to export after a product is exported for the first time.

Overall, the entry pattern is consistent with the interpretation that part of the fixed cost is sunk. Once a firm introduces a product into a destination, firms enter at a significant higher rate the following 2 years. These findings point to spillover effects from lower fixed costs for following exporters and are consistent with the definition of product fixed costs. To strengthen the evidence of spillovers across firms, the next paragraph discusses additional empirical implications.

Higher firm entry rates after product entry as shown in Table 2.5 may lead to business stealing effects. An export pioneer who opens up international markets reduces fixed costs for product market rivals and thus spurs entry. More entry increases competitive pressure and results in lower prices. Based on this argument, products with a higher number of exporters per destination should be negatively correlated with export prices. Also, the willingness of the pioneer to bear the product fixed cost increases in market size because of the higher demand in larger markets. If this effect dominates, the pioneer may not be willing to pay the fixed cost in the small market. We would expect a negative correlation between the number of exporters per product and the number of export markets penetrated. Because the number of destinations does not control for the market size of the export markets penetrated, I also include the rank of the export market with the largest and the lowest size as additional control variables. The business stealing effect predicts that products with lots of firms export only to lower ranked markets, i.e. the markets with the largest size.

To investigate business stealing effects at the product level, I use the following regression specification:

$$\begin{aligned} \log N_{k,c,t} &= \beta_1 \log \bar{p}_{k,c,t} + \beta_2 \log s_{k,c,t} + \beta_3 \log \bar{q}_{k,c,t} + \\ &+ \beta_4 \log M_{k,c,t} + d_k + d_{c,t} + \epsilon_{d,c} \end{aligned} \tag{2.8}$$

where $N_{k,c,t}$ is the average number of exporters per destination in product class k from country c in period t, $\bar{p}_{k,c,t}$ is the unit value our proxy for the export price of the product, $s_{k,c,t}$ is the survival probability of an exporter remaining an exporter the following year, $\bar{q}_{k,c,t}$ represents the per firm average quantity exported and $M_{k,c,t}$ stands for the number of destinations product k is exported to. d_k and $d_{c,t}$ are product and country-time fixed effects. Country-time fixed effects control for institutional differences and macroeconomic trends that are common across products. Product fixed effects control for any characteristics that are common across export destinations like demand, substitutability and potentially common fixed costs. [11] Note that the fixed effects will not capture the effect of pioneers on the product differentiability and revenues of rivals because these firms operate in different product categories in different countries.

Table 2.6 plots the results. The number of firms per destination for a given product is significantly negative correlated with the average export price. Since we control for demand by product fixed effects, the number of export destinations and the firm's average quantity exported, I consider this as supportive evidence for business stealing effects. Lower fixed costs spur firm entry and results in more product market competition.

Table 2.6 also shows that contrary to our expectations products with many firms per destination are exported to more destinations, column (1) and (2). A potential explanation is that the number of destinations proxies for comparative advantage. In the Eaton and Kortum [2002] model a country has a stronger comparative advantage in a product group if that product is exported to many destinations. Given this interpretation, countries export a product to many destinations because the average productivity of firms within this product category is high. Firm export participation is positively correlated with the number of destinations not because of lower fixed cost but due to lower marginal costs caused by comparative advantage.

The quantity coefficient in Table 2.6 is significantly and positive. Because of lower fixed costs, firms slide down the average cost curve, increase production

11. I assume that product demand is common across countries, i.e. that consumers in different destination markets have the same demand for a product. Under this assumption, product fixed effects will control for demand effect.

efficiency and export on average a larger quantity. Note that the quantity effect is significantly larger than the price effect implying that efficiency gains from lower fixed costs dominate business stealing effects from more entry. Artopoulos et al. [2013] provide anecdotal evidence that export pioneers acquire knowledge about foreign markets through their embeddedness in the business community of destination markets. The generated knowledge diffuses to rival firms within the same sector in the domestic market, lowers fixed costs to export and increases their efficiency. Firm participation and export sales per firm increase significantly. My finding is consistent with this argument. An addition implication of the knowledge spillover is that firms learn to conduct international business allowing them to remain an exporter in the next period. Including the survival probability of staying an exporter next period as an additional regressors, confirms this conjecture. Firms exporting products with many rival firms have on average a 22 percent higher probability of survival in export markets.

Artopoulos et al. [2013] argue that spillover effects are particular pronounced in sectors with a high degree of product differentiation. In product categories that allow for more product differentiation, firms can react to more product market competition by upgrading their own product through quality. The higher the degree of product differentiation, the lower is the competition pressure form product market rivals. We expect the negative relationship between export price and firm entry to be weakened, i.e. differentiated product groups should experience relative more product entry. In regression 2.8, I control for product differentiation by including product fixed effects. In a sensitivity analysis I re-estimate equation 2.8 for different types of products classified according to Rauch [1999]'s product differentiation index. The three groups are: homogeneous goods, reference priced goods and differentiated goods. Index 1 refers to homogeneous goods, 2 to reference prices goods and 3 to differentiated goods.

Table 2.7 contains the results. I test whether the sensitivity of price on the number of firms per destination is lower for differentiated products than for homogeneous products. In differentiated product the effect of price on the number of firms per destination is significantly lower than in the other 2 groups. Also, the probability of staying in export markets and the average export revenues are higher for firms exporting differentiated products. This is additional evidence for the argument of Artopoulos et al. [2013] that positive spillover effects are stronger in differentiated products.

In sum, the results suggest that consistent with product fixed costs, once a firm introduces a product into a market subsequent exporters face lower fixed costs. The time analysis shows that most firms enter the year right after a product was exported the first time. This finding is consistent with lower fixed costs due to the removal of part of the fixed cost to export. I also find that the associated spillover form the lower fixed costs leads not only to higher entry rates but is also accompanied by a business stealing effect, i.e. the higher firm entry reduces export prices. At the same time, lower fixed costs allow firm to produce at a more efficient scale and export a larger quantity. Overall, the larger quantity offsets the negative price effects and average export revenues per firm increase.

2.6 Conclusion

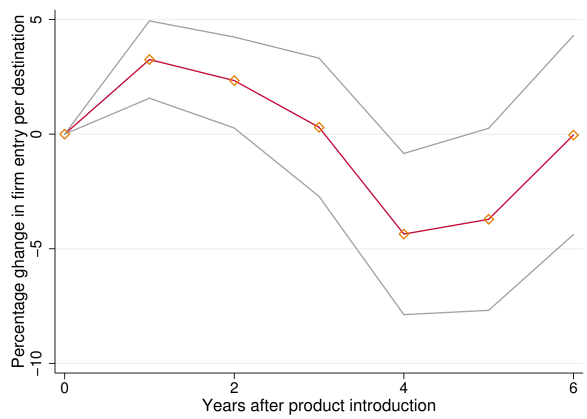
This paper develops an empirical framework to analyze fixed costs that exporting firms face in order to participate in international markets. The analysis distinguishes between fixed costs on the firm and product level. In the first part I study the presence of fixed costs by evaluating how the entry of firms and products varies with destination market size. The second part describes potential spillover effects at the firm and product level caused by fixed costs and presents empirical evidence consistent with these effects.

The results indicate that entry barriers operate at the product level. Taking cross country evidence into account, product fixed costs are even more important relative to firm fixed costs in countries with a large home market. Small countries have often only one firm within a product category, thus product fixed costs are identical to firm fixed costs. Overall, the results suggest that small economies are particularly affected from fixed costs. The low level of domestic demand implies that firms are not able to benefit from economies of scale resulting in relative high prices and a disadvantage in comparison to firms from larger economies. Moreover, because of low demand, few firms find it profitable to export to them. The limited entry results in even higher prices due to the lack of competitive pressure.

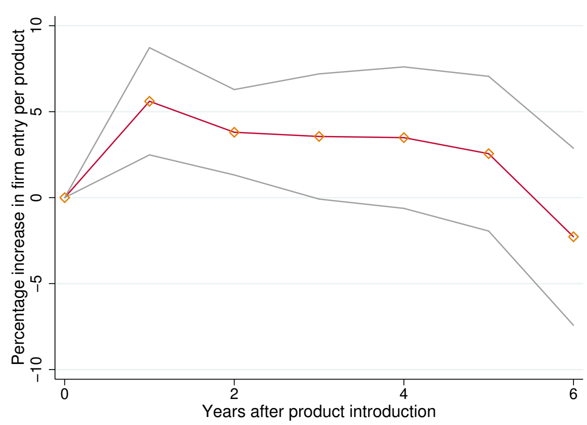
To investigate the effects of fixed costs on the entry decision of firms further, I analyze how firms change their product range with market size. The results show that the average number of products per firms does not change with the size of the destination market pointing to no cost advantages of expanding the product range in larger markets. I consider this as supportive evidence of product fixed costs. Including information on the timing of entry of products, I find consistent with product fixed costs that the entry of firm increases significantly the year after a product is introduced to a destination market. The higher entry of rival firms indicates lower entry barriers due to the removal of the product fixed costs. The additional entry introduces competitive pressure and lowers export prices, i.e. business stealing effects. The lower fixed cost allows rival firms to produce at a more efficient scale, increases their export revenues and results in a higher probability of staying in international markets the next period.

In conclusion, my findings have important policy implications. For the exporting country policies encouraging new product entry, for example advertising new products in destination markets through export promotion agencies, rather than firm entry would potentially lead to spillover effects that translate into higher level of firm participation and export growth. By paying part of the product fixed costs, the government increases incentives for firms to explore new export destinations and offsets part of the negative effects due to free riding of rival firms. The importing country can lower product fixed costs by reducing technical-barriers to trade. As a result, consumer surplus increases because of lower product prices due to competitive pressure. More generally, the existence of entry barriers to export implies that trade policy can effect market structure. When conducting policy experiments in the form of a reduction in trade costs, it is standard in the international trade literature to consider only a fall in marginal costs and evaluate the resulting impact on the patterns of trade and consumer welfare. However, an important component of the current EU-US free trade negotiations is the reduction of technical-barriers to trade by negotiating common product standards. Neglecting the existence of entry barriers and the resulting industry reallocations underestimates the impact of trade reforms.

2.7 Figures



(a) Firm entry within a destination



(b) Firm entry within a product group

Figure 2.2 – The firm entry rate over time after a product is exported to a market for the first time.

2.8 Tables

Table 2.1 – Exporting countries in the sample

Albania	Domenican Republic	Macedonia	Peru
Bangladesh	Ecuador	Malawi	Portugal
Belgium	El Salvador	Mali	Senegal
Bulgaria	Estonia	Mauritius	South Africa
Burkina Faso	Guatemala	Mexiko	Spain
Cambodia	Iran	Morocco	Sweden
Cameron	Jordan	Nicaragua	Turkey
Chile	Kenya	Niger	Uganda
Colombia	Laos	Norway	United Rep. Tanzania
Costa Rica	Lebanon	Pakistan	Yemen

Note: Data from the Exporter Dynamics Database provided by the World Bank

Table 2.2 – Importing countries in the sample

Afghanistan	Denmark	Kyrgyzstan	Samoa
Albania	Djibouti	Laos	Sao Tome & Principe
Algeria	Dominica	Latvia	Saudi Arabia
Angola	Dominican Republic	Lebanon	Senegal
Antigua & Barbuda	Ecuador	Liberia	Seychelles
Argentina	Egypt	Libya	Sierra Leone
Armenia	El Salvador	Lithuania	Singapore
Australia	Equatorial Guinea	Macao	Slovak Republic
Austria	Eritrea	Macedonia	Slovenia
Azerbaijan	Estonia	Madagascar	Solomon Islands
Bahamas	Ethiopia	Malawi	Somalia
Bahrain	Fiji	Malaysia	South Africa
Bangladesh	Finland	Maldives	Spain
Barbados	France	Mali	Sri Lanka
Belarus	Gabon	Malta	St. Kitts & Nevis
Belgium	Gambia, The	Marshall Islands	St. Lucia
Belize	Georgia	Mauritania	St.Vincent & Grenadines
Benin	Germany	Mauritius	Sudan
Bermuda	Ghana	Mexico	Suriname
Bhutan	Greece	Micronesia	Sweden
Bolivia	Grenada	Moldova	Switzerland
Bosnia & Herzegovina	Guatemala	Mongolia	Syria
Brazil	Guinea	Morocco	Taiwan
Brunei	Guinea-Bissau	Mozambique	Tajikistan
Bulgaria	Guyana	Nepal	Tanzania
Burkina Faso	Haiti	Netherlands	Thailand
Burundi	Honduras	New Zealand	Togo
Cambodia	Hong Kong	Nicaragua	Tonga
Cameroon	Hungary	Niger	Trinidad & Tobago
Canada	Iceland	Nigeria	Tunisia
Cape Verde	India	Norway	Turkey
Central African Republic	Indonesia	Oman	Turkmenistan
Chad	Iran	Pakistan	Uganda
Chile	Iraq	Palau	Ukraine
China	Ireland	Panama	United Arab Emirates
Colombia	Israel	Papua New Guinea	United Kingdom
Comoros	Italy	Paraguay	United States
Congo, Dem. Rep.	Jamaica	Peru	Uruguay
Congo, Republic of	Japan	Philippines	Uzbekistan
Costa Rica	Jordan	Poland	Vanuatu
Cote d'Ivoire	Kazakhstan	Portugal	Venezuela
Croatia	Kenya	Qatar	Vietnam
Cuba	Kiribati	Romania	Yemen
Cyprus	Korea, Republic of	Russia	Zambia
Czech Republic	Kuwait	Rwanda	Zimbabwe

Note: Data from Comtrade, Penn World Table and CEPII

Table 2.3 – Summary Statistics

	Obser.	Mean	Median	St. Dev.	Min	Max
Nr. of exporters	30164	343,8	39	1112,9	2	28981
Nr. of products	30164	297,1	55	564,8	1	4163
Nr. of exporters/product	30164	2,12	1,46	1,87	1	43,23
Nr. of products/exporter	30164	2,54	1,97	2,87	1	104,80
Av. revenues per exporter	30164	1,29	0,58	4,38	8,92E-06	755,4
Av. revenues per product	30164	1,27	0,49	4,57	2,85E-06	645,6
GDP in destination	1560	451909	65967	1358439	145	14400000
GDP per capita in dest.	1560	13303	92395	14071	192	91707
Expenditure share	30164	0,00125	0,00026	0,00783	1,97E-09	0,40083
Distance	30164	6873	6177	4343	86	19812
GDP in origin country	182	275916	165278	351089	8247	1516755
GDP per capita in origin	182	12105	7978	12090	559	54927

Note: Statistics are aggregated over all export destinations. Average expenditure per firm is total imports of destination per exporting country divided by number of exporting firms. Average expenditure per product is total imports of destination per exporting country divided by number of exported products. Average expenditure per firm and per product as well as GDP are measured in million International dollars. Expenditure shares are defined as a country's total value of imports per exporting country divided by the country's total expenditure, i.e. GDP. GDP per capita is measured in International dollars. Distances are in kilometers from capital city in country i to capital city in country j.

Table 2.4 – Entry of firms and products with respect to market size

Dependent variable	log(Number of firms) (1)	log(Number of products) (2)	log(Number of firms) (3)	log(Number of products) (4)
log(Market Size)	0.403***	0.317***	0.439***	0.357***
	[0.0116]	[0.0115]	[0.00595]	[0.00560]
log(Expenditure Share)	0.366***	0.360***	0.205***	0.199***
	[0.00631]	[0.00724]	[0.00475]	[0.00478]
log(Distance)			-0.828***	-0.847***
			[0.0232]	[0.0263]
log(GDP per capita)			0.139***	0.0953***
			[0.0113]	[0.0113]
Border			0.347***	0.311***
			[0.0225]	[0.0290]
Observations	30164	30164	30164	30164
R-squared	0.661	0.618	0.764	0.723

Note: The results from ordinary least squares regressions for the dependent variable normalized by the import expenditure share are noted at the top of each column projected on the covariates listed in the first column. All regressions include origin country, time and origin country-time fixed effects. Robust standard errors in parentheses: ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Table 2.5 – Fixed costs and the number of exporters per destination

Dependent variable	Firm entry per destination (1)	Firm entry per product (2)
year_1	0.943***	0.460***
	[0.249]	[0.130]
year_2	0.477**	0.312***
	[0.257]	[0.104]
year_3	0.086	0.292*
	[0.446]	[0.152]
year_4	-1.264**	0.286
	[0.520]	[0.172]
year_5	-1.077*	0.209
	[0.587]	[0.188]
year_6	-0.011	-0.186
	[0.642]	[0.0.215]
Constant	29.12***	8.210***
	[6.832]	[1.459]
Observations	3297489	2703038
R-squared	0,729	0,529

Note: The dependent variable is the number of entrants divided by the total number of exporters in a destination (column (1)) or within a product group (column (2)). The results are based on ordinary least squares regressions. All regressions include origin country - time fixed effects. The destination specific regression in column (1) includes product-time and origin-destination fixed effects whereas the product specific regression in column (2) includes product-country and destination-time fixed effects. Robust standard errors in parentheses (clustered by destination time in column (1) and by product time in column (2)): ***, **, * marks statistically significant difference from zero at the 1***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Table 2.6 – Fixed costs and the number of exporters per destination

Dependent variable	log(Av. Nr. Exporters per destination) (1)	(2)
log(Av. Unit value)	-0.01025***	-0.01085***
	[0.000703]	[0.000702]
Log(Av. Quantity)	0.0399***	0.0394***
	[0.000619]	[0.000625]
log(Nr. of destinations)	0.127***	0.117***
	[0.00154]	[0.00163]
Survival Probability	0.225***	0.224***
	[0.00443]	[0.00443]
Rank of largest market		-0.000786***
		[3.14e-05]
Rank of smallest market		0.000322***
		[4.06e-05]
Observations	201,788	201,788
R-squared	0.495	0.497

Note: The dependent variable is the average number of exporters per destination. The results are based on ordinary least squares regressions. All regressions include origin country, product and time fixed effects. Robust standard errors in parentheses (clustered by country time): ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Table 2.7 – Fixed costs and the number of exporters per destination for differentiated, less differentiated and homogeneous products

Dependent variable:	log(Av. Nr. Exporters per destination)		
Differentiation index	1	2	3
log(Av. Unit value)	-0.0389***	-0.0138***	-0.0046***
	[0.0033]	[0.0016]	[0.0005]
Log(Av. Quantity)	0.0145***	0.0173***	0.0237***
	[0.0017]	[0.0005]	[0.0009]
log(Nr. of destinations)	0.0333***	0.0342***	0.146***
	[0.0073]	[0.0037]	[0.0018]
Survival Probability	0.129***	0.152***	0.263***
	[0.0147]	[0.0081]	[0.0055]
Rank of largest market	0.000753***	0.000990***	0.000723***
	[0.0001]	[7.06e-05]	[3.61e-05]
Rank of smallest market	-0,000273	0.000360***	0.000295***
	[0.0001]	[8.49e-05]	[4.76e-05]
Observations	9682	40573	151369
R-squared	0,386	0,424	0,532

Note: The dependent variable is the average number of exporters per destination. The product differentiation index assigns a value of 1 to homogeneous goods, 2 to reference prices goods and 3 to differentiated goods. The results are based on ordinary least squares regressions. All regressions include origin country, product and time fixed effects. Robust standard errors in parentheses (clustered by country time): ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Table 2.8 – Relationship of market size and the number of firms and products including the decompostion of the extensive margin

Dependent variable	log(Number of firms) (1)	log(Number of products) (2)	log(Av. Nr. of Products per firm) (3)	log(Av. Nr. of Firms per per product) (4)
log(Market Size)	0.439***	0.357***	0.0181	0.100***
	[0.00595]	[0.00560]	[0.00979]	[0.00292]
log(Expenditure Share)	0.205***	0.199***	0.0210***	0.0269***
	[0.00475]	[0.00478]	[0.00206]	[0.00179]
log(Distance)	-0.828***	-0.847***	-0.185***	-0.166***
	[0.0232]	[0.0263]	[0.00937]	[0.00860]
log(GDP per capita)	0.139***	0.0953***	0.0112***	0.0546***
	[0.0113]	[0.0113]	[0.00401]	[0.00374]
border	0.347***	0.311***	0.0870***	0.123***
	[0.0225]	[0.0290]	[0.00890]	[0.00852]
Observations	30164	30164	30164	30164
R-squared	0,764	0,723	0,346	0,482

Note: Total firm-product combinations (T) are decomposed into $T_{d,c} = P_{d,c}\bar{p}_{d,c}$, where $P_{d,c}$ is the number of exported products from country c to destination d and $\bar{p}_{d,c}$ is the average number of firms per products exported. Equivalently, $T_{d,c}$ can also be decomposed into $T_{d,c} = F_{d,c}\bar{f}_{d,c}$ the number of exporting firms in c with shipments to destination d and the average number of products per exporter from c to d, $\bar{f}_{d,c}$. The results from ordinary least squares regressions for the dependent variable normalized by the import expenditure share are noted at the top of each column projected on the covariates listed in the first column. All regressions include origin country, time and origin country fixed effects. Robust standard errors in parentheses: ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Table 2.9 – Relationship between market size and the number firms and products including proxies for fixed costs

Dependent variable	log(Number of firms) (1)	log(Number of products) (2)
log(Market size)	0.195***	0.108***
	[0.0169]	[0.0178]
log(Expenditure Share)	0.240***	0.252***
	[0.00591]	[0.00622]
log(Distance)	-0.662***	-0.632***
	[0.0149]	[0.0156]
log(GDP per capita)	0.502***	0.574***
	[0.0285]	[0.0300]
Border	0.162***	0.0632***
	[0.0221]	[0.0233]
% of urban population	0.00263***	0.00391***
	[0.000294]	[0.000309]
log(Landsize km^2)	-0.0577***	-7,82E-05
	[0.00834]	[0.00878]
log(Container Traffic)	0.234***	0.225***
	[0.0102]	[0.0107]
log(Rail km)	0.156***	0.140***
	[0.0120]	[0.0126]
log(Nr. of internet subscribers)	0.03739	0.05841
	[0.04124]	[0.04183]
log(Electricity per capita)	-0.392***	-0.487***
	[0.0212]	[0.0223]
Observations	11.096	11.096
R-squared	0,867	0,843

Note: All regressions include time and origin country fixed effects. Robust standard errors in parentheses (clustered by country time): ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Chapter 3

The causal impact of migration on U.S. intermediate goods' trade: Evidence from a natural experiment

3.1 Introduction

Are international trade and migration substitutes or complements? An important strand of thought has maintained that trade and migration are, in fact, complementary, see Egger et al. [2012]. International migrants can increase international trade flows via two distinct mechanisms. First, migrants shift preferences towards the goods of their country of origin, thus generating demand for imports of those goods by their host country. Second, migrants reduce transaction costs between countries, either by holding information about relevant market characteristics or by attenuating frictions due to imperfect contract enforcement.

This paper studies the impact of immigrants on intermediate goods imports by studying geographical variations across US states for the period 1970 to 2005. To generate our results, we follow a gravity approach and regress the log of intermediate imports on the log of recent immigrants. We address reverse causality, and more broadly endogeneity, and find evidence of a significantly positive effect of immigrants on US intermediate imports. To estimate a causal relationship, we analyze the exogenous allocation of refugees within the US refugee resettlement program. Our results are robust to an alternative identification strategy, based on the large influx of Central American immigrants to the United States after hurricane Mitch. We find that a 10 percent increase in recent immigrants to a given US state raises imports from those immigrants' country of origin by 1.5 percent. Overall, our results suggest that immigrant networks play an important role in promoting trade across countries

There are two important differences between our analysis and previous studies. First, we focus on regional variations in immigration and intermediate imports *within* the United States. This approach allows us to estimate the trade-migration relationship leaving aside the potential correlation between trade and migration policy at the national level. Since there are no data on the distribution of intermediate imports across US states for the time period under review, we apply a procedure based on Anderson and Yotov [2010] to infer this distribution. Using data from the County Business Pattern and applying the techniques from Autor et al. [2013] for each US state, we calculate the intermediate imports as the sum of industry-specific imports. The latter are modeled as depending on the relative size of the industry and on bilateral trade costs with respect to each trading partner (geographical distance, border and the value-to-weight ratio). This method allows us to project the observed national import flows by industry onto the state level. To assess the appropriateness of our method, we use data available for recent years on the distribution of intermediate imports across US states and compare the constructed import flows with the observed ones. We find a high correlation between the two flows, i.e. a correlation coefficient of 0.85. In addition, using recent observed import data, we find estimates of the impact of migration on trade to be qualitatively similar to the ones based on constructed import flows for the 1970-2005 period. [1]

Second, and most importantly, we address endogeneity by focusing on exogenous shocks to immigration. Endogeneity arises because immigrants' decisions regarding their settlement within the United States are likely to be correlated with several variables, such as income, employment opportunities and/or preferences, which in turn are correlated with trade, see Borjas [1999]. An additional source of endogeneity arises in the form of reverse causality, i.e. immigrants from a given country are likely to settle in states that trade a lot with their country of origin. As a first step towards addressing these issues, we focus on recent immigrants and include the number of non-recent immigrants living in the state as a control variable. The non-recent immigrant population of a given origin will capture the unobserved preferences of immigrants of that origin for a specific location. It will also account for the impact of better income and employment opportunities in a state that are specific to immigrants of that origin, see Clark et al. [2007]. In addition, and most importantly, we address endogeneity and reverse causality by estimating an IV specification. First, we use the exogenous allocation of political refugees across US states within the refugee resettlement program. Second, we take advantage of an exogenous shock that forces people to migrate. In particular, we look at of the large influx of Central American immigrants to nearby US states following hurricane Mitch in October 1998. The IV approach removes the endogenous component of migration decisions whereby individuals might move to those regions with the best trading opportunities.

Starting with the pioneering work of Gould [1994], there exists ample empirical literature that argues that immigrants increase trade across international

1. Before 2008, only national US import data are available.

borders. Gould [1994] studies the effect of migration on aggregate US exports and imports for the years 1970-1986. He estimates a gravity model of trade on migration and finds evidence of a strong positive relationship. Many authors follow Gould and study immigration into a single country, for example Head and Ries [1998] examine Canada and Girma and Yu [2002] the United Kingdom. The drawback of focusing on the aggregate level is that these studies cannot account for correlation in migration and trade policies. More recent studies exploit the regional distribution of immigrants and look at the bilateral trade relationship between US regions and foreign countries, see Bardhan and Guhathakurta [2004], Dunlevy [2006] and Parsons and Vézina [2014]. While these studies focus exclusively on US exports, this paper is the first to present evidence on US imports at the state level.

In a recent literature review, Felbermayr et al. [2012] argue that the main concern for the identification of the causal effect of immigration is reverse causality. To deal with this issue, authors have adopted different approaches. Drawing upon the seminal work of Card [2001], several papers (see for example Peri and Requena-Silvente [2010]) instrument changes in immigrants at the sub-national level by applying the growth in immigrants from a given country of origin at the national level to the distribution of immigrants from a given country of origin across regions at a previous point in time. Still, historical migrant stocks can have direct effects on trade even many years after their arrival, therefore violating the exclusion restriction.

The only other paper that has recourse to a natural experiment for identification is Parsons and Vézina [2014]. Their strategy is based on the combination of an immigration shock caused by the exogenous allocation of Vietnamese Boat People across US states and a concurrent trade embargo. They use the cross section variation in Vietnamese immigrants and US exports across the 50 US states to estimate the impact of Vietnamese migration on US exports. Rather than focusing on a particular group of migrants at a given point in time, this paper uses data on political refugees to the U.S. from all countries between the years of 1970 and 2005. To establish the causal relationship, our analysis takes advantage of the fact that political refugees to the U.S. are exogenously allocated across locations once we control for the existence of families ties of migrants. As a robustness check, we use the migration flows from several Central American countries following hurricane Mitch. Overall, our findings show that the trade-enhancing effect of immigrants found by

Parsons and Vézina [2014] also holds in a general setting where we consider regional intermediate imports from many countries.

The rest of the paper is organized as follows. Section 2 describes the identification strategy as well as the details of the natural experiments, the political refugees and hurricane Mitch, which allow us to establish causality between trade and migration. Section 2 also contains the import demand model used to project the aggregate US data onto the state level. Section 3 covers the data and the summary statistics. Section 4 discusses the obtained OLS and IV results. Section 5 discusses the robustness of our findings and section 6 concludes.

3.2 Identification strategy

Our identification strategy looks at variations in intermediate imports and the number of recent immigrants across US states. Once migrants settle in a US state, international trade between the state of residence and their respective home countries is expected to increase. This effect is likely to take place through the following channels. First, migrants might provide information that reduces transaction costs. Second, trade might increase simply because migrants have a preference for the goods from their country of origin. However, in our analysis we do not explicitly distinguish between the two mechanisms.

The reason why we focus on intermediate imports is that, for our sample period, we have no information on regional trade data in the United States. Instead, we infer the distribution of trade flows across US states via an intermediate input demand model. This model allows us to project the observed aggregate trade data onto the state level. A key element in the calculation of regional intermediate imports is production data by region and industry. Since consumption data by region and industry is not available, we cannot assess regional imports for final goods. We then discuss the estimation equation and identification issues, followed by the procedure for inferring regional intermediate import data.

3.2.1 Regression specification

To analyze the impact of immigration on intermediate imports, we follow the literature, Felbermayr et al. [2012], and employ the gravity equation. The gravity equation relates trade flows between state i and country j to the relative sizes of the participating economies (which in our specification are captured by fixed effects). The log of the import flow of state i from country j for the period t, X_{ijt}, is given by

$$\begin{aligned} \log(X_{ijt}) &= \beta_1 \log(Imm_{ijt}) + \beta_2 \log d_{ij} + \beta_3 b_{ij} + \\ &+ \beta_4 \log PImm_{ijt} + f_{jt} + f_{it} + \varepsilon_{ijt} \end{aligned} \tag{3.1}$$

The regressor Imm_{ijt} indicates the number of foreign-born individuals above the age of 21 who immigrated from country j to state i in any of the 5 years prior to time t. The other regressors are the log of the distance, d_{ij}, between the capital of state i and the capital in country j, measured in kilometers. b_{ij} represents a dummy variable that indicates whether state i shares a border with country j. $PImm_{ijt}$ is the number of foreign-born individuals above the age of 21 who immigrated from country j to state i more than 10 years prior to time t. In addition, we also introduce time varying state and country fixed effects, f_{it} and f_{jt} respectively. The coefficient of interest is β_1. If $\beta_1 > 0$, then the presence of recent immigration increases intermediate imports.

Equation 3.1 represents the ideal regression in the case of import flows being observed. Given that we approximate import flows within an intermediate input demand model, we introduce a measurement error, u_{ijt}. The estimation equation becomes

$$\begin{aligned} \log(\hat{X}_{ijt}) &= \beta_1 \log(Imm_{ijt}) + \beta_2 \log d_{ij} + \beta_3 b_{ij} + \\ &+ \beta_4 \log PImm_{ijt} + f_{it} + f_{jt} + v_{ijt} \end{aligned} \tag{3.2}$$

and

$$\log(\hat{X}_{ijt}) = \log(X_{ijt}) + u_{ijt}$$

where $\hat{X}_{ijt}$ represents the estimated state import flows and v_{ijt} the error in the OLS regression. β_1 is consistently estimated if $Cov(Imm_{ijt}, u_{ijt}) = Cov(Imm_{ijt}, v_{ijt}) = 0$. Due to the presence of the measurement error, the standard errors are larger and the likelihood of not rejecting the null hypothesis of $\beta_1 = 0$ increases, see Wooldridge [2010]. Thus, a significant effect of migration on trade will reinforce our confidence in the results obtained. Equation 3.2 will be our main regression specification.

In terms of identification of β_1, it is important to control for geography, since both migration and trade are correlated with distance and border, see Head and Mayer [2013]. States import more from countries that are relatively close by and with whom they share a common border. At the same time, relatively more people migrate to neighboring countries that are close by. Neglecting these effects will introduce an omitted variable bias.

Note that, by analyzing the trade migration relationship across US states, we directly address the criticism of Hanson [2010] with respect to the earlier literature. He argues that "It is difficult to draw causal inference from results based on international trading and migration patterns, since immigration may be correlated with unobserved factors that also affect trade, such as the trading partners' cultural similarity or bilateral economic policies (e.g., preferential trade policies or investment investment treaties that raise the return to both migration and trade)". Trade policies and investment treaties are negotiated at the federal level and are thus controlled for by fixed effects specific to the country of origin, f_{jt}. These fixed effects also control for any effects that are common to all regions in the United States. For example, if a country experiences a positive productivity shock, trade might increase since all regions will face lower import prices from this country and emigration might decrease because of better employment opportunities.

A further concern for the identification of parameter β_1 is the presence of time-varying state-specific characteristics that may be correlated with trade flows as well as immigration. One such candidate is, for instance, economies of agglomeration, i.e more immigrants are likely to settle in larger states and those states have higher demand for intermediate goods. For this reason, we include time-varying state-fixed effects, f_{it}, that control for any state-specific effects, such as local demand and income shocks, which are common to all migrants.

By looking at regional variations and including state and country time-varying fixed effects, we follow the recent literature, see Dunlevy [2006] and

Bardhan and Guhathakurta [2004]. The key difference of this paper with respect to the literature is that we offer a new approach to resolving endogeneity. Endogeneity arises because immigrants' decisions regarding settlement within the United States is likely to be correlated with several variables, such as income, employment opportunities and/or preferences, which in turn are correlated with trade, see Borjas [1999]. An additional source of endogeneity arises in the form of reverse causality, i.e. immigrants from a given country of origin are likely to go to states that trade a lot with that country.

To tackle endogeneity, we proceed in two steps. First, we distinguish between different groups of immigrants according to their arrival date. Specifically, we focus on newly arrived immigrants and include the previous number of immigrants as an additional control variable, $PImm_{ijt}$. This allows us to account for the fact that immigrants are not randomly distributed across geographical locations. Consider for example Austrian immigrants, who have a specific knowledge of the production of skies and decide to emigrate to the United States in order to make use of this knowledge. They may prefer to go to Colorado rather than Texas because of the relatively higher demand for skies, and thus the greater business and employment opportunities for them in the Rocky Mountains. In a similar vein, Italians may prefer to move to California in order to start producing wine rather than to Montana where climatic conditions prevent wine from growing. In both scenarios, the settlement choice differs across migrants of different origins and is correlated with business and employment opportunities and/or preferences, which in turn are correlated with imports from their countries of origin. We mitigate the resulting bias by looking at the impact of recent immigrants on trade and, at the same time, controlling for the previous number of immigrants from the same country of origin already living in the region. Of course, this hinges on the assumption that recent immigrants' settlement preferences are not fundamentally different to those of previous immigrants.

As a second step, we focus on exogenous shocks to migration. These exogenous shocks address endogeneity arising from reverse causality, i.e. immigrants from a given country of origin are likely to go to states that trade a lot with that country. Gould [1994] argues that immigration occurs before the onset of trade and is therefore predetermined. This is true if the migration decision is based on current or past levels of trade. However, if the migration decision is forward-looking and dependent on expected future trade (for example people emigrate in order to take advantage of information arbitrage, which leads to trade), past immigration is endogenous. As a result, the number of immigrants and the level of trade are jointly determined by 2 equations. The first equation captures the preference/trade cost effect of immigrants, where trade is a function of the number of immigrants. The second equation describes the information arbitrage effect, where the number of immigrants is a function of expected levels of trade. The resulting simultaneity bias renders the OLS estimates inconsistent.

To solve for the resulting bias, the literature follows Card [2001] and instruments changes in immigrants at the sub-national level by applying the growth of immigrants from a given country of origin at the national level to the distribution of immigrants from that country of origin across regions at a previous point in time, see for example Peri and Requena-Silvente [2010]. Still, historical migrant stocks could have established long-standing trade relationships, with direct effects on the current level of trade. This will violate the exclusion restriction of the instrument. Instead, we take advantage of an exogenous shock that forces people to migrate independently of trading opportunities. This shock influences only the preference/trade cost equation and is independent of the information arbitrage equation, thus insulating our results from any endogenous migration decisions, whereby individuals might move to those regions with better trading opportunities.

The only other paper that provides evidence of a natural experiment on the trade migration relationship is Parsons and Vézina [2014]. Their identification strategy is based on the combination of an immigration shock driven by the exogenous allocation of Vietnamese Boat People by the US government across US states and a concurrent trade embargo. They use the cross section variations in the share of Vietnamese immigrants and US exports to Vietnam of the 50 US states. They show that after the end of the 1994 trade embargo US states with a higher share of Vietnamese immigrants exported significantly more to Vietnam. The key identification assumption is that the settlement choice of Vietnamese immigrants before the trade embargo is exogenous to US exports after the embargo. In contrast, our analysis considers 2 different natural experiments. First, we use the exogenous allocation of political refugees across US states within the refugee resettlement program. Second, we make use of a natural disaster, Hurricane Mitch. As we argue below in more detail, this hurricane led to the mass immigration of several hundred thousand people from Central America to the US.

The main differences with respect to Parsons and Vézina [2014] are the following. First, we focus on regional differences in US imports, whereas they focus on US exports. Second, our analysis resolves reverse causality by examining the panel structure of exogenous migration within two natural experiments. As a result, the number of observations increases significantly compared to Parsons and Vezina's cross section of 50 US states. A further benefit from working with the panel structure is that we can exploit the bilateral variation over time and control for any state and country time-varying fixed effects. The third distinction is that we include the pre-existing number of immigrants as an additional regressor. This allows us to control for unobserved settlement preferences. Finally, we use the exogenous allocation of the US refugee resettlement program, where migrants' only influence in the settlement decision is to decide whether to be close to family members already living in the United States or not. Given that we control for the number of previous immigrants, the exogenous variation in the migration decision is solely driven by the allocation of "free" application cases, i.e. those where the political refugee has no family ties and friends in the US.

Overall, we see our approach as complementary to that of Parsons and Vézina [2014] and as a test of external validity in a multi-country setting.Parsons and Vézina [2014] focus on a specific group of migrants at a given point in time. This paper uses data on refugees to the U.S. from all countries in the period of 1970 to 2005, leading to a more comprehensive sample, both in terms of years and countries of origin. Next, we describe our IV strategy in more detail. We start with the allocation of political refugees in the United States and then discuss hurricane Mitch.

3.2.2 Refugees

Refugees are people who have fled their home country and cannot return because they have a well-founded fear of persecution based on religion, race, nationality, political opinion or membership of a particular social group (Immigration and Nationality Act, Sect. 101[a][42]). Each fiscal year, the US government sets an overall refugee admissions limit based on regional allocations. The limit varies from year to year. For example, in 2005, 53,738 refugees were admitted to the United States, primarily from Laos (8,487), Somalia (10,106) and the former USSR (11,272). The total number of refugees per country of origin is given in Table 3.2. Figure 3.2 plots the regional distribution of refugees. Over the whole sample period, the state that took in the most refugees is California with 277,395 refugees. New York ranked second with 144,273 refugees, followed by Florida with 111,657 refugees.

The assignment of refugee settlement is not random. The US Refugee Admissions Program attempts to place refugees in the same community as their family members living in the United States, if the former stated a family relationship in the application process. If not, refugees are placed in a different state to avoid the concentration of ethnic groups, as was the case for Cuban refugees in Florida in the 1970s, see Kerwin [2012]. For this reason, we include the previous number of immigrants living in the state as a control variable.

The resettlement of refugees is decided by the US Department of State and it depends on the capacity of the local organizations to take care of the refugees, see David [2004]. To control for a potential correlation between a state's capacity to host refugees and the level of its income, and in further instance trade, we include time-varying state fixed effects. Controlling for time-varying state fixed effects and the number of previous immigrants living in the state will ensure the appropriateness of the instrument. This is true as long as the assignment process of refugees across states does not discriminate on the basis of the migrant's nationality (other than the family relationship).

In the first stage, we regress the log of the number of refugees who arrived in the 5 years prior to year t from country j and settle in state i (Ref_{ijt}) on the number of immigrants who arrived in the past 5 years (Imm_{ijt}). The first-stage equation is given by:

$$\begin{aligned} \log(Imm_{ijt}) &= \alpha_0 + \alpha_1 \log Ref_{ijt} + \alpha_2 \log d_{ij} + \\ &+ \alpha_3 \log PImm_{ijt} + f_{it} + f_{jt} + \epsilon_{ijt} \end{aligned} \tag{3.3}$$

where f_{it} and f_{jt} are country and state year pair fixed effects. As already mentioned, we include the number of previous immigrants, $PImm_{ijt}$, in state i from j to account for the fact that refugees may have a family member already living already in that state. The state-year effects, f_{it}, control for any state specific change in the allocation of refugees over time that is common to all countries. Country-year fixed effects, f_{jt}, control for country of origin specific effects that are common to all states in the United States, such as the nature of the conflict that forced people to emigrate or any other macroeconomic condition in the country of origin. Finally, d_{ij} measures the distance in kilometers between the capital of country j and the capital of state i. We exclude the border effect because there are no refugees from Mexico or Canada who migrated to the United States in our sample period

The main advantages of using refugees are that (1) refugees cannot freely choose where they settle in the United States, (2) the refugee sample includes a wide variety of countries and (3) these countries represent a large share of US imports, see Table 3.1 [2]. However, the correlation between refugees and migrants may potentially be weak because only a limited number of refugees are allowed to enter the United States every year. In addition, the underlying reasons why refugees flee their home country vary greatly across countries and individuals and we have no information on the underlying circumstances that trigger the migration choice.

3.2.3 Consequences of Hurricane Mitch for Migration

To address the concerns about the generality of migration and trade effects based on political refugees, we also focus on the trade-migration relationship with Central American countries and the United States after hurricane Mitch. We argue that the natural disaster forced people to migrate, while they would not have done so otherwise. We show that (1) migrants choose to settle in states that are close to their home country and (2) the increase in immigrants from hurricane-affected countries reduces the effect of distance on trade.

Hurricane Mitch hit Central America during the last week of October 1998. The countries impacted the strongest were Honduras and Nicaragua, but Guatemala and El Salvador (and to a much lesser extent Belize) were also affected. The hurricane destroyed a large part of these countries' transportation and social infrastructure, including roads, hospitals and schools, see Worldbank [2001]. According to the World Bank, people from Central America responded to the disaster by migrating to the United States. As a

2. We discuss the summary statistics in detail below after presenting the data sources.

formal response to the migration generated by hurricane Mitch, the Immigration and Naturalization Service (INS) announced in a news release the designation of Temporary Protected Status (TPS) to countries hit by the hurricane for a period of 18 months, which was later extended, see Kugler and Yuksel [2008]. As a consequence, Central Americans, who entered during this period, were not subject to deportation from the United States and, at the same time, were eligible to work in the United States. The INS estimates that by 2003, close to 50,000 Hondurans and Nicaraguans had been granted Temporary Protected Status (TPS) to allow them to stay and work in the United States. This massive inflow of Central Americans was accompanied by a large amount of illegal immigration, such that the number of Central Americans who came from these countries was probably higher than this official number. According to the immigration data from the US census, a total 637 thousand persons from El Salvador, Honduras, Nicaragua, Guatemala and Belize migrated to the United States, which was significantly higher than in previous years.[3]

Based on Figure 3.1, the migration decision is correlated with distance. Because these migrants have a preference for the goods from their home country and they carry information about the economic conditions in both countries, we consequently expect an increase in trade flows between the two regions. In particular, we expect a lower effect of distance on trade flows. The basic idea is that if immigrants move to states that are closer to their country of origin and immigrants have trade enhancing effects, then these regions will trade disproportionally more with the immigrants country of origin than those regions further away. As a result, we expect that the effect of distance on trade between the respective state and country of origin increases.

The first-stage equation for the IV estimates is:

$$\begin{aligned} \log(Imm_{ijt}) &= \alpha_0 + \alpha_1 \log d_{ij} dPost_{it} + \alpha_2 \log d_{ij} + \alpha_3 b_{ij} + \\ &+ \alpha_4 \log PImm_{ijt} + f_{it} + f_{jt} + \epsilon_{ijt} \end{aligned} \tag{3.4}$$

3. In the appendix we formally estimate the increase in the number of immigrants per US state following a difference in differences approach. The results show that the increase in the number of immigrants was 53 percent higher than the increase in immigrants from other Latin American countries.

where f_{it} and f_{jt} are country and state-year fixed effects. We include destination-specific time effects as well as destination origin specific effects to check whether the increase in the share of immigrants in close-by states simply is due to an ongoing trend or whether there is indeed a discernible break after Mitch. $dPost_{it}$ is a dummy that takes the value of 1 if the immigrant from country i arrived after hurricane Mitch and zero otherwise. d_{ij} is a variable which measures the Euclidean distance in kilometers from the capital in country j to the capital of state i.[4] b_{ij} indicates whether the state is bordering Canada or Mexico and $PImm_{ijt}$ is the previous number of immigrants from country j living in state i at time t.

Our instrument is represented by the interaction between a post-Mitch dummy and distance from Central America to various states in the United States. It captures the idea that after the hurricane close-by states received relatively more immigrants from Central America. This alters the effect of distance on trade and as a result close-by states trade more with Central American countries.

An important assumption for the validity of our instrument is that the effect of the hurricane was exogenous to export capabilities of the immigrants' country of origin and that the immigrants were the only factor that reduced the effect of distance on trade. Given our discussion about the hurricane destroying large parts of the local infrastructure and production facilities, this assumption is likely to be violated. Exports from hurricane-affected countries decreased significantly after Mitch. For this reason, we include a country of origin-year fixed effect (f_{jt}) in regression 3.4. As long as the reduction in exports does not differ significantly across US states, these dummies capture the reduction in trade after the hurricane.

In the previous paragraphs we described our estimation equation together with the IV strategy. We next describe the intermediate input demand model that generates the distribution of intermediate import flows across US states. The OLS and IV estimation results will follow in section 4.

4. 3The appendix contains detailed information on the method for calculating the distance between US states and the immigrants' country of origin

3.2.4 Inferring regional-industry level trade

To project trade flows onto the state level, we use data on the regional industry production structure and apply the intermediate input demand model put forward by Anderson and Yotov [2010]. The model relates industry specific trade flows between the exporting country and the geographical location in the United States. To begin, consider the unobserved true level of intermediate imports of state i from country j in year t, (X_{ijt}). These aggregate regional imports can be written as the sum over all industries within the state:

$$X_{ijt} = \sum_{k=1}^{K} w_{ijt}^{k} X_{jt}^{k}$$

where w_{ijt}^{k} is the regional import share of industry k in state i and X_{jt}^{k} represents the aggregate US import expenditure of industry k from country j. These aggregate US industry specific imports from country j are observed. Hence we estimate the unobserved regional import share. To do so, we follow Anderson and Yotov [2010]. Assume that final good producers in a particular state i have the following CES demand for the aggregate quantity of intermediate inputs in industry k (Q_{it}^{k}):

$$Q_{it}^{k} = \left(\sum_{j=1}^{J} \beta_{jt}^{k} \left(q_{ijt}^{k} \right)^{1-1/\sigma_k} \right)^{\sigma_k/(1-\sigma_k)}$$

where q_{ijt}^{k} is the quantity of intermediate inputs from country j demanded by industry k. β_{j}^{k} is a share parameter specific to the country of origin and the industry but common to all regions within the United States. σ_k is the industry specific elasticity of substitution. Total industry specific intermediate import expenditure, $X_{it}^{k} = \sum_{j=1}^{N} X_{ijt}^{k}$, is then given by the quantity demanded, Q_{it}^{k}, times the price index, P_{it}^{k}. Given the CES demand function, we can express the bilateral import expenditure, X_{ijt}^{k}, as the value of shipments at destination prices from the country of origin j to the geographical region i in industry k.

$$X_{ijt}^{k} = \left(\frac{\beta_{jt}^{k} p_{jt}^{k} t_{ijt}^{k}}{P_{it}^{k}} \right)^{1-\sigma_k} X_{it}^{k} \tag{3.5}$$

Destination prices are the product of the factory gate prices p_{jt}^{k} and trade costs $t_{ijt}^{k} \geq 1$ of the shipment of goods from i to j in industry k. The key assumption is that producers in country i charge the same factory gate price to all buyers within the United States. Prices within the United States vary only due to industry-region specific trade costs. The CES price index is given by $P_{it}^k = \left[\sum_{j=1}^J \left(\beta_{jt}^k p_{jt}^k t_{ijt}^k\right)^{1-\sigma_k}\right]^{1/(1-\sigma_k)}$.

Note that aggregate US country specific imports are simply the sum over all states, i.e. $X_{jt}^k = \sum_{i \neq j}^J X_{ijt}^k$. We write the normalized industry specific regional import share as:

$$w_{ijt}^k = \frac{X_{ijt}^k}{\sum_{i=1}^I X_{ijt}^k} = \frac{\left(t_{ijt}^k / P_{it}^k\right)^{1-\sigma_k} X_{it}^k}{\sum_{i=1}^I \left(t_{ijt}^k / P_{it}^k\right)^{1-\sigma_k} X_{it}^k} \tag{3.6}$$

which depends on the total expenditure of state i on the goods of industry k, the price level and trade costs. States with lower trade costs and higher expenditure import more. In order to calculate the regional import share, we follow Stumpner [2013] and assume the following trade cost function

$$\begin{aligned} \log(t_{ijt}^k) &= \delta_1 \log(\text{Distance}_{ij}) - \delta_2 \log(\text{Distance}_{ij}) \log(\text{Value-to-weight}_t^k) - \\ &- \delta_3 \text{Border}_{ij} \end{aligned} \tag{3.7}$$

Trade costs increase with distance and decrease if the state shares a border with Mexico or Canada and/or if the value-to-weight ratio is high. Modeling trade costs using an interaction between distance and the value-to-weight ratio will capture the heterogeneous effect of distance on trade flows across industries. Industries with a high value-to-weight ratio, for example computer chips, have relatively lower trade costs than industries with a low value-to-weight ratio, for example car engines. Stumpner [2013] presents evidence that this parameterization of trade costs reflects well US inter-state trade flows at the industry level. Unlike Stumpner, we include a dummy variable if the region borders Canada or Mexico to account for potential border effects.

The problem in estimating the trade elasticities $\delta_1 - \delta_3$ in equation 3.7 is that industry specific international trade flows at the state level are not available. To circumvent this problem, we offer two alternative approaches. First, we follow Stumpner [2013] and use his elasticities based on US inter-state trade: $\delta_1 = 0.667$, $\delta_2 = 0.081$ and $\delta_3 = 0$. However, these estimates are derived from trade between US states and do not take into account international flows. The resulting trade cost function may not be suitable if there are discontinuities in trade costs when crossing an international border. In an alternative approach, we estimate the trade elasticities using aggregate US industry specific international trade data and impose the estimated elasticities on the regional level. Following the estimation procedure described in detail in the appendix, we obtain elasticities similar to Stumpner, i.e. $\delta_1 = 0.691$, $\delta_2 = 0.154$ and $\delta_3 = 1.897$. Note that the results reported below are based on the elasticities using the international trade data and are robust to the values used by Stumpner [2013].

To calculate w_{ij}^k in equation 3.6, we require information on the intermediate input expenditure. We follow Autor et al. [2013] and assume that the industry's expenditure can be proxied by the total number of workers in industry k in region i. This assumption is reasonable if expenditure on intermediate inputs within an industry is proportional to total production, which again is proportional to employment.[5] Finally, we assume a common demand elasticity $\sigma_k = \sigma$ and that the law of one price within an industry k holds across the United States, i.e. $P_i^k = P^k$. The resulting weights are then given by

$$\hat{w}_{ijt}^k = \left(\frac{\left(t_{ijt}^k\right)^{1-\sigma} L_{jt}^k}{\sum_j \left(t_{ijt}^k\right)^{1-\sigma} L_{jt}^k} \right)$$

The intermediate import flow at the state level is simply the sum of the industry specific estimated weights times the industry specific national import flow.

$$\hat{X}_{ijt} = \sum_k \hat{w}_{ijt}^k X_{jt}^k \tag{3.8}$$

By summing over all industries, we minimize measurement errors due to the assumption that the law of one price holds at the industry level. This assumption can be problematic if relative prices across regions vary in a systematic way, i.e. a region has significantly higher industry prices than others. To capture aggregate price differences across states, we include time-varying state fixed effects in our main regressions. These fixed effects will control for state specific price differences common to all industries.

5. If we assume a Cobb-Douglas production function with labor (L), intermediate inputs and potentially other input factors, then demand for intermediate inputs is proportional to total labor employed.

The calculated import flows in equation 3.8 form the basis of our empirical analysis. The following section explains in detail the underlying data needed to apply equation 3.8. Afterwards, we discuss the estimation results.

3.3 Data and summary statistics

This section describes the data used and provides basic summary statistics. Particular emphasis is placed on trade flow and industry data. The latter allows us to map aggregate trade flows across states at the industry level. We then provide an overview of the other datasets and discuss the summary statistics

3.3.1 Import data

The import data comes from the US Census Bureau provided by Schott [2008] for the years 1972 to 2005.[6] The data contains import flows at the 5 digit SITC product level to 183 trading partners by mode of transportation, i.e. either air or sea. The import value of shipments is defined as the net selling value exclusive of freight charges and excise taxes. To calculate the intermediate import levels at the industry level, we use a correspondence table from the World Bank that converts 5 digit SITC codes into the Broad Economic Categories (BEC) developed by the United Nations. The BEC classification enables us to group each SITC code into intermediate and final goods so that we obtain product specific intermediate imports for each trading partner.[7] To make the trade flows directly comparable to the industry employment records at the state level, we convert the 5 digit SITC codes into SIC87 industry codes using the procedure described in Feenstra [1996]. As a result, we obtain the intermediate import value for each 4 digit SIC 87 industry code (409 industries). These values are used as X_{jt}^{k} in equation 3.8.

6. We consider the trade data of 1970 to be similar to the trade flows in the year 1972. This enables us to use an additional year of immigration data. In the appendix, we present the results when the year 1970 is excluded.

7. Intermediate goods are defined as the sum of the categories: Processed food and beverages (12), Industrial supplies (2), Capital goods (4) and Parts of transport equipment (53).

3.3.2 Trade cost data

The average value-to-weight ratio per SIC 87 industry is calculated as follows. For each 4 digit SIC 87 code, we divide the total import value by the total weight and convert it into tons. To compute the average value-to-weight ratio for each industry, take the average value-to-weight ratio across all trading partners for each mode of transportation, i.e. air or sea, and weight the importance of the transportation mode by its share in the total import value per industry.

To calculate the bilateral distances used in the trade cost function, we adopt the procedure used by Mayer and Zignago [2011]. d_{ij} is the distance between the capital of state i and the capital of country j measured in kilometers and calculated by the Great Circle Distance Formula. Like Mayer and Zignago [2011], we use 32.19 kilometers as inner-city distance. All data on latitudes and longitudes are from the Global Administrative Areas (GADM) database, see Hijmans et al. [2010].

3.3.3 Industry data

To measure the industry structure at the state level, we follow Autor et al. [2013]. We take the local employment structure from the County Business Patterns (CBP) for the years 1970, 1980, 1990, 2000 and 2005. The CBP is an annual data series that provides information on employment, firm size distribution and payroll by county and industry and covers all of the United States. Based on the data and the approach of Autor et al., we calculate total employment per state and SIC 1987 industry code and use it as a proxy for intermediate input demand in the regional import weight.[8]

3.3.4 Immigration data

Our measures of the immigrant population are based on data from the Census Integrated Public Use Micro Samples compiled by Ruggles et al. [2004] for the years 1970, 1980, 1990 and 2000, and the American Community Survey

8. Given a Cobb-Douglas production function, the demand for intermediate inputs will then be proportional to total employment per industry. As long as this proportionality factor does not vary across US states within an industry, total employment will be a proxy for intermediate input demand within an industry.

(ACS) for 2005. The 1980, 1990 and 2000 Census samples include 5 percent of the US population, while the pooled ACS and 1970 Census samples include 3 and 1 percent of the population respectively. The main explanatory variable in our regressions, i.e. recent immigrants, is defined as the number of immigrants who immigrated up to 5 years prior to the census year. In the robustness section, we control for pre-existing immigrants by all those immigrants who live in the respective state and immigrated ten years or more prior to the census year. We focus only on immigrants who are older than 21 at the census year. We then aggregate the number of immigrants at the state level using the census sampling weight.

As an alternative identification strategy, we use political refugees as an instrument for immigration in the import regression. Data on the number of refugees per US state come from the Office of Refugee Resettlement (ORR). The ORR provides the option to download refugee arrival data sorted by country of origin and state of initial resettlement in the United States for the years 1984 to 2005. Each fiscal year, the US government sets an overall refugee admissions limit based on regional allocations. The limit for refugee admissions varies from year to year depending on the Congress and the geopolitical situation. Figure 3.2 plots the number of refugees for each state over the period 1985 to 2005.[9] In order to make the refugee data compatible with the immigration data, we add up the refugee data per country of origin and state for all 5 years prior to the census years 1990, 2000 and 2005 census years. Table 3.2 contains the total number of refugees per country of origin who arrived in the United States during our sample period.

3.3.5 Country treatment

Finally, we should mention that several countries changed their names and borders during the 1970-2005 sample period. Since we cannot identify the origin of earlier immigrants, for example the number of Ukrainian immigrants prior to 1990 in the United States, we keep the country borders of the year 1970. We group all imports and migrants from former Soviet republics into

9. See US Department of Homeland Security. 2005 yearbook of immigration statistics. Washington, DC: US Department of Homeland Security. Available at http://www.dhs.gov/xlibrary/assets/statistics/yearbook/2005/OIS_2005_Yearbook.pdf.

one country, i.e. the USSR [10]. Immigrants from Bosnia and Herzegovina, Croatia, Macedonia, Serbia and Montenegro, Kosovo, and Slovenia are all included in Yugoslavia. Slovakia and the Czech Republic form Czechoslovakia and Eritrean immigrants count as Ethiopians even though Eritrea became a separate nation in 1993. The exception is Germany, which we assume to be unified as of 1970.

3.3.6 Summary statistics

Table 3.1 presents the summary statistics for each of the 5 census years, 1970, 1980, 1990, 2000 and 2005. Columns 1 to 5 show the total number of immigrants, the number of newly arrived immigrants and the trade statistics for each census year. The total number of immigrants to the United States grew from 9.7 million in 1970 at an annual growth rate of 3.8% to 36.1 million in 2005, which is reflected in the steady rise of recent immigrants, i.e. those immigrants who arrived up to 5 years prior to the census year. From 1965 to 1970, only 1.7 million people arrived, whereas from 2000 to 2005 that number grew to 8.1 million. Immigration became also more diversified. The number of different countries from which immigrants migrated rose from 67 in 1970 to 138 in 2005. During the same period, the total value of intermediate imports increased from USD 2.3 billion in 1970 to USD 82.7 billion in 2005. This increase corresponds to an annual growth rate of 10.7%, much higher than the growth rate of the immigrant population. Overall, the summary statistics suggest a positive relationship between the number of recent immigrants and intermediate imports.

To infer a causal link between migration and trade, we take advantage of the exogenous variation in migration in the form of a natural experiment. In particular, we use the emigration caused by hurricane Mitch and refugees from other geopolitical conflicts. For this reason, Table 3.1 splits the sample into three country groups: countries affected by hurricane Mitch, Latin American countries and refugee countries. Hurricane-affected countries are El Salvador, Guatemala, Honduras, Nicaragua and Belize. Latin American countries are all American countries beside the United States and Canada; they act as a control group for migration caused by hurricane Mitch. Refugee

10. The former USSR includes the following present-day nations: Azerbaijan, Armenia, Belarus, Estonia, Georgia, Latvia, Lithuania, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Uzbekistan, and Ukraine.

countries are all countries that sent refugees to the United States over the 1985-2005 period.

With respect to the different country groups, the total number of immigrants from hurricane Mitch affected countries only account for a small share of US immigration and trade. In 2005, they represented 6.7% of total immigrants, while their contribution to intermediate import flows was 0.18%. When compared to other Latin American countries, their share rises to 16% in terms of immigration and 1.4% in terms of total intermediate imports from Latin America to the United States. Given the low representation of hurricane-affected countries in overall trade, we compare the migration and import decisions of these countries relative to other Latin American countries. Below, we argue in more detail that these countries are more similar in their unobserved trade and migration characteristics, which helps to reduce potential omitted variable biases.

To further address concerns about the generality of the migration and trade effects in the case of Central America, we also focus on the trade-migration relationship for countries that sent political refugees to the United States. As in the case of migration caused by hurricane Mitch, refugees take their migration decisions independently of economic reasons and act as an instrument in our empirical analysis. The advantage of looking at the allocation of political refugees is that these refugees come from a wide range of countries, which trade a lot with the United States. Between 2001 and 2005, refugees from 83 countries settled across the United States and the combined import share of these countries was 72%.[11] However, the correlation between refugees and migrants may be weak because the share of refugees in the total number of immigrants is only 3.3%.

3.4 Results

This section shows empirically that the migration channel was important for increasing intermediate imports across US states. We start by a simple OLS regression of equation 3.2. The results are presented in Table 3.3. Columns (1) to (3) show the estimation results for all countries, where we observe

11. Table 3.2 contains the number of refugees per country of origin over the whole sample period, while figure 3.2 plots the distribution of refugees across states.

migration and trade data. Columns (4) to (6) contain the results for the restricted sample, where we only consider Latin American countries. Overall, we find a positive and significant migrant trade effect on intermediate imports across all specifications. The baseline OLS results in columns 1 and 4 suggest that a 10% increase in the number of recent migrants raises intermediate imports by 0.49% in the full and 0.78% in the restricted sample. When testing for significant differences between the restricted and the full sample, the results show that the effect is significantly higher for Latin American countries.

Both the magnitude of the migration elasticity for Latin American countries (0.0784) and for all countries (0.049) are significantly lower than the values found in the literature. Looking at the national level of US imports for the period 1870 to 1910, Dunlevy and Hutchinson [1999] found a value of 0.28. Other studies at the regional level also found significantly higher values. Combes et al. [2005] found a value of 0.12 in the case of France. Wagner et al. [2002] found a value of 0.28 using a variation across Canadian provinces and Bratti et al. [2011] found a value of 0.32 for Italy. There are at least two explanations for the low migration elasticities. First, the literature does not distinguish between intermediate and final goods. The above-mentioned papers focus exclusively on aggregate imports and consider both the preference effect and the trade cost reducing effect of immigrants. Second, this paper focuses on newly arrived immigrants. Thus, it sets their effect on trade apart from the already established trade network effects created by previous immigrants.

In columns (2) and (3) as well as (5) and (6), we add the number of previous immigrants as an additional control variable. It is likely that the migration decision of recent immigrants is not random and influenced by the presence of previous immigrants from the same country of origin. If the omitted variable drives both trade and migration, the previous results will be biased. For example, assume migrants move to states where there are already previous migrants for cultural or economic reasons. If these regions then experience positive demand shocks that lead to more production and employment, it is unclear whether the demand for intermediate imports increases because of the new immigrants or just because of the positive demand shock on the already established trade network. As a result, we expect the trade cost effect through the arrival of recent immigrants to be lower than in the absence of previous immigrants. The resulting bias from the omission of previous immigrants in the regression may overestimate the true effect.

To control for the number of previous immigrants, we follow two distinct approaches because immigrants are likely to change location between the year they arrived in the United States and the year the census took place. In columns (2) and (5) we add the number of immigrants who arrived 10 or more years ago and reported to be living in state i in the year the census was conducted. In columns (3) and (6) we use the number of immigrants before the first year of trade is observed. More precisely, we the number of immigrants who arrived up until in 1960 and reported to be living in state i in 1970. Previous immigration will be correlated with actual immigration if recent immigrants prefer to settle in states where there exists a large pre-existing community. However, the specification is less demanding because it is based on the distribution of immigrants in the year 1960.

Table 3.3 presents the results. When we control for pre-existing immigrants, the migration elasticity on intermediate imports increases slightly. For Latin American countries, the coefficient decreases from 0.078 to 0.062 when using the number of immigrants who arrived ten years ago and from 0.049 to 0.04 for the full sample. In the case of the immigrant distribution in the year 1960, the coefficient increases to 0.07 and 0.012 respectively. However, while these coefficients are not significantly different from the baseline results, they point to a potential upward bias in excluding the previous immigration. Note that the combined effect of previous and recent immigrants is larger and comparable to values found in the literature.

3.4.1 Instrumental variable results

Note that the OLS results might be biased, for example if people immigrate in order to take advantage of trading opportunities, i.e. trade causes immigration. To resolve the question of endogeneity, we consider the following two approaches. First, we use geopolitical conflicts that forced people to leave their home country. Second, we make use of a natural disaster, Hurricane Mitch. We start by discussing the results based on political refugees before showing the results in the case of hurricane Mitch.

3.4.1.1 Refugees

The first stage results are presented in columns (1) to (3) of Table 3.4. First, the number of refugees is positive and significantly correlated with the number of immigrants in all specifications. Note that the positive effect decreases in magnitude once we add the number of previous immigrants as an additional control variable, see columns (2) to (3). The reason is that part of the refugees are settled in states where they have family members living already in the US.

Before turning to the IV baseline results, we say a few words on the validity of our instrumental variable. To confirm the validity of the IV regressions, we include the Cragg-Donald F statistics and the Kleibergen-Paap test p-values at the bottom of our tables. The Kleibergen-Paap test examines whether the excluded instrument is correlated with the endogenous regressors conditional on the control variables. A p-value below 0.1 indicates that we can reject the null hypothesis of no correlation and that the instrument is statistically significant. The Cragg-Donald F statistic provides an indication of the significance of our instrument. If the instrument is only weakly correlated with the endogenous regressor, the IV estimator is not valid. To assess the weakness of our instrument we need to compare these F statistics with the Stock-Yogo critical values for the Cragg-Donald F-statistic with one endogenous regressor (Stock and Yogo [2002]). As a rule of thumb, an F-statistic above 10 indicates that the IV is acceptable.

Columns (4) to (6) in Table 3.4 contain the results of the second stage IV regression. The coefficient on the number of recent immigrants is significantly positive with an elasticity higher than in the OLS regression for all specifications. However, performing a Hausman test does not reveal any significant difference in the coefficients. Overall, our results suggest that a 10% increase in the number of immigrants in the past five years increases intermediate imports by 1.8%.

There are potential concerns that countries where political refugees are coming from do not form a representative sample for US trade. Table 3.2 shows that the majority of refugees are either from current or former communist countries (USSR, Vietnam, Cuba) or from countries with whom the US has political tensions (Cuba, Iran, Myanmar). Furthermore, correlation between refugees and migrants may potentially be weak because only a limited number of refugees are allowed to enter the United States every year, i.e. the share of refugees in terms of total number of immigrants is 3-5 percent. For this reason, we consider an alternative instrument, i.e. hurricane Mitch. The argument supporting the validity of the instrument is outlined in section 2 and parallels the argument in the case of political refugees. Hurricane Mitch forced people to migrate who would have not done so otherwise.

3.4.1.2 Hurricane Mitch

Columns (1) to (3) in Table 3.5 present the results of the first stage IV regression with different control variables. First, distance has a significant negative effect: immigrants move to states that are closer to their country of origin. The variable instrument, capturing the effect of distance after the hurricane, is significant and positive. While the overall effect of distance on the migration decision of Central Americans is still negative, migrants from Central America moved relatively further away from their country origin after the hurricane. Note that this effect is robust to the inclusion of the previous stock of previous immigrants in the state, see columns (2) and (3).

Focusing on the p-value of the Kleibergen-Paap test, we can reject the null hypothesis of no correlation between the endogenous variable and the instrument at the 1% level in all specifications. The Cragg-Donald F statistic is higher than 10 and therefore increases our confidence in the validity of the instrument.

Columns (4) to (6) in Table 3.5 contain the second stage IV regression, where the number of recent immigrants is instrumented by the log of distance interacted with a post-time dummy for hurricane Mitch. The coefficient on recent immigrants is significant and positive with an elasticity higher than in the OLS regression for all specifications. Thus, states that received more Central American immigrants after hurricane Mitch face a lower distance elasticity and import more. The estimated immigrant elasticity is 0.13 and comparable in magnitude to the effect when using political refugees as an exogenous migration variation.[12]

12. As noted in section 2, the identification assumption is that the only reason why the elasticity of trade with respect to distance between US states and Central American countries changed is because (1) these states received relatively more immigrants from

Overall, the similarity between the estimated coefficients leads us to conclude that immigrants have a significant and positive impact on intermediate imports. A 10 percent increase in the number of immigrants in the past five years increases intermediate imports by 1.5 to 1.8 percent.

3.5 Robustness

In this section, we check the robustness of our main findings using an IV strategy, conducting the following two falsification/robustness exercises. First, we address the potential measurement bias that arises from the fact that we construct intermediate import flows via a demand model. To do so, we use recent observed US state import data and compare them to the constructed import flows. Second, given our discussion about the fact that after hurricane Mitch Central American migrants moved predominantly to the United States, we should find that Central American exports to the United States increased more strongly than those to other countries. The following paragraphs address each of these points separately.

3.5.1 Observed import data

As mentioned above, we do not observe intermediate imports over the period from 1970 to 2005. However, the US Census provides state specific import flow data for the years 2008 to 2012. For these years, we can compare the estimated import flows with the observed import flows and check the robustness of our estimation method. We proceed in two stages. First, we use the import demand model from section 2 together with the national US intermediate import data and generate state specific import data for the years 2008 to 2012. We then compare this estimated import flow with the observed import flow for the same years. If the model works well, we should expect a high correlation between the two flows. Second, we compare the estimated migration elasticity from the observed import flows with the elasticity based on the estimated import flows in the previous section. If the measurement error of the demand model is uncorrelated with immigration, then the two elasticities should not be significantly different from each other.

Central America and (2) immigrants have trade enhancing effects by providing information or having preferences for goods of their country of origin.

To construct the import flow data, we apply equation 3.8 to the trade cost function specified in equation 3.6 and use the industry and aggregate US trade data for the years 2008 to 2012. We then calculate a correlation coefficient between the two series and obtain a coefficient of 0.85. This suggests that our procedure approximates the true import flows quite well.

We then use the estimation equation 2 and replace the calculated import flow with the observed import flow. We focus on the cross section of the year 2010. The reason for this is that we do not have trade data over a 5 year interval so as to use the number of immigrants who arrived in the past 5 years as an explanatory variable. [13] The precise regression model for the year 2010 is:

$$\begin{aligned} \log(X_{ij}) &= \beta_1 \log(Imm_{ij}) + \beta_2 \log d_{ij} + \beta_3 b_{ij} + \\ &+ +\beta_4 \log PImm_{ij} + f_i + f_j + v_{ij} \end{aligned} \tag{3.9}$$

Table 3.6 plots the OLS estimates for Latin American countries and for all countries. The results are similar to the OLS estimates in the constructed sample, see Table 3.3. The level of the coefficients is similar and the elasticity for Latin American countries is higher than for all countries. The only exception is that in the Latin American sub-sample the immigration elasticity is not significant when we include the number of immigrants who migrated 10 or more years before.

Since the hurricane occurred in 1998, our instrumental variable identification strategy for the year 2010 is solely based on the number of refugees who came to the United States between 2005 and 2010. Table 3.7 shows the results. The

13. For example, if we use the number of immigrants who arrived in the past 5 years as an explanatory variable in the year 2012 within the panel structure, then the previous import flow should correspond to the year 2007. This ensures independence between the previous trade flow and the arrival of immigrants in the past 5 years. If the two periods overlap, the number of immigrants and the trade flow are correlated. Note that we do not observe the import data for the year 2007. Once the trade data of the year 2013 become available, we can conduct a panel analysis and use the period 2008 to 2013. We experiment with the period 2008 to 2012 and use the number of immigrants who arrived in the past 4 years as an explanatory variable. The results are quantitatively similar to the ones reported in Table 3.6.

first stage results are reported in columns 1 to 3. We find that the number of refugees is positively correlated with the number of immigrants over the past five years in all specifications. The F statistics are larger than 10 and the Kleinbergen-Papp P-value are smaller than 0.05, which suggests that our instrument is appropriate. Columns 4 to 6 present the second stage IV. The number of immigrants has a significant impact on the trade level in all but column 5, where we include the number of immigrants who came at least ten years ago to the United States. Overall, the results and the magnitude of the coefficients are similar to our previous results. These findings suggest that our results are robust to the introduced measurement error of the calculated trade flows.

3.5.2 Effect on the exports of the country of origin

In the previous section, we use the variation in immigrants and Central American exports across US states. Next, we use the variation in the export revenue of Central American countries across all possible destinations in the world. According to our discussion in the previous section, Central Americans mainly responded to the disaster by migrating to the United States. Given this argument, Central American exports to the United States should be higher than those to other countries. We consider the following regression

$$\log(X_{ijt}) = \beta_1 D_{US,2000} + d_{ij} + d_{jt} + \epsilon_{ijt} \tag{3.10}$$

where the X_{ijt} is total intermediate exports from country i to country j in year t. Exporting countries are El Salvador, Guatemala, Honduras, Nicaragua and Belize. Importing countries are all countries in the world. To focus on long-term effects, we consider 5 year intervals. The years considered are 1975, 1980, 1985, 1990, 1995, 2000 and 2005. $D_{US,2000}$ is a dummy variable that equals 1 if the year is 2000 or 2005 and the importing country is the United States. Otherwise, the dummy takes the value zero. This dummy captures whether exports from hurricane-affected countries to the United States were significantly higher than to other countries after the hurricane in the year 2000. The model includes importing country year fixed effects, d_{it}, as well as exporting-importing country pair fixed effects d_{ik}. d_{it} controls for any exporting country-time specific effects, such as aggregate demand shocks. d_{ik} is a dummy variable that controls for any pair fixed effects, such as distance, preferences or other bilateral characteristics that are constant over time.

Table 3.8 shows the results. The dummy coefficient is significant, which shows that between 1995 and 2000 exports from hurricane-affected countries increased significantly in comparison to other countries. The average per year effect is: $(1.81/1)^{(1/5)} - 1 = 12.6\%$. While we cannot attribute the increase in exports entirely to the number of immigrants, it is consistent with the idea that migration fosters trade between countries.

3.6 Conclusion

Immigrants carry information about the goods and market conditions of the country of origin as well as the country of destination. By providing this information, immigrants reduce bilateral trade costs and increase trade between countries. This paper focuses on the trade cost reducing effect of immigrants by looking at the relationship between intermediate import demand and immigration across US States. Using the exogenous migration decision brought about by a natural experiment (hurricane Mitch), we establish the causal relationship between immigrants and imports. Our results indicate a strong positive impact of migration on trade. We find that a 10 percent increase in immigrants raises trade by 1.5 to 1.8 percent.

Policy makers may want to take into account the pro-trade effect of immigrants. By providing information on market conditions in both countries, the country of origin as well as the country of destination, they reduce transaction/trade costs for importers and exporters. As a result, industries can purchase cheaper intermediate inputs from abroad and produce more efficiently.

Taking a broader perspective, immigrants may also have knowledge of production techniques used in their country of origin, which can increase the comparative advantage of industries in their country of destination, see Bahar and Rapoport [2014] for empirical evidence on this issue. All in all, these results suggest that the mobility of people between countries can serve as a key element in enhancing industrial productivity growth. However, more research is needed to assess the long-term impact of immigration on the economy.

3.7 Figures

Figure 3.1 – The log of the number of immigrants who immigrated from hurricane Mitch affected countries (Nicaragua, El Salvador, Honduras, Guatemala and Belize) over the period from 1998 to 2005.

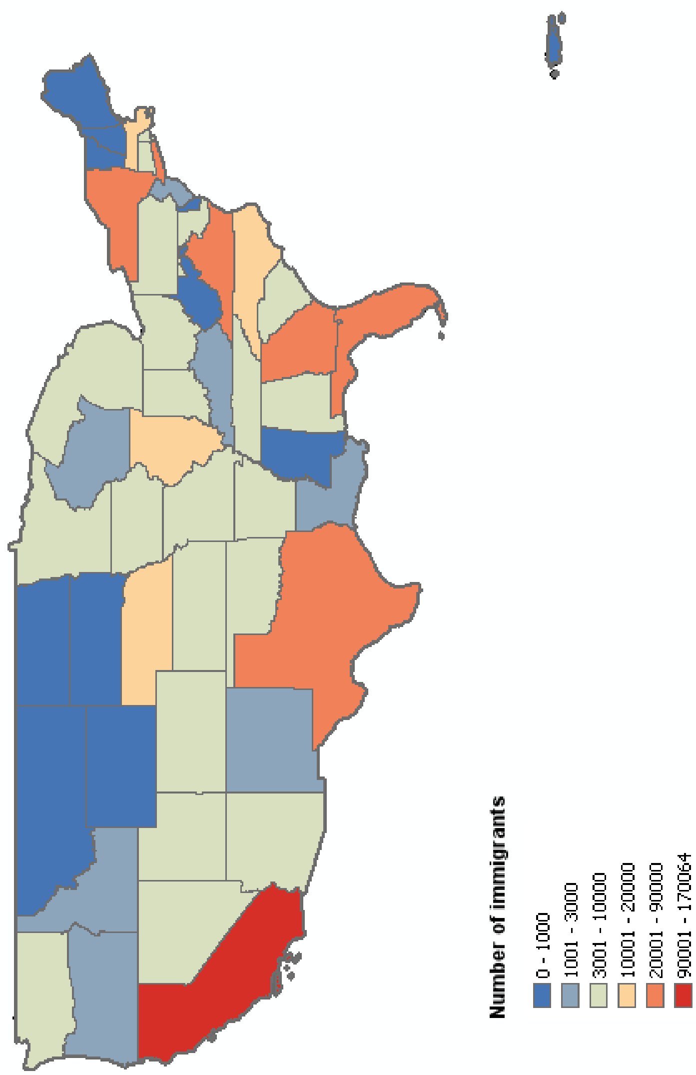
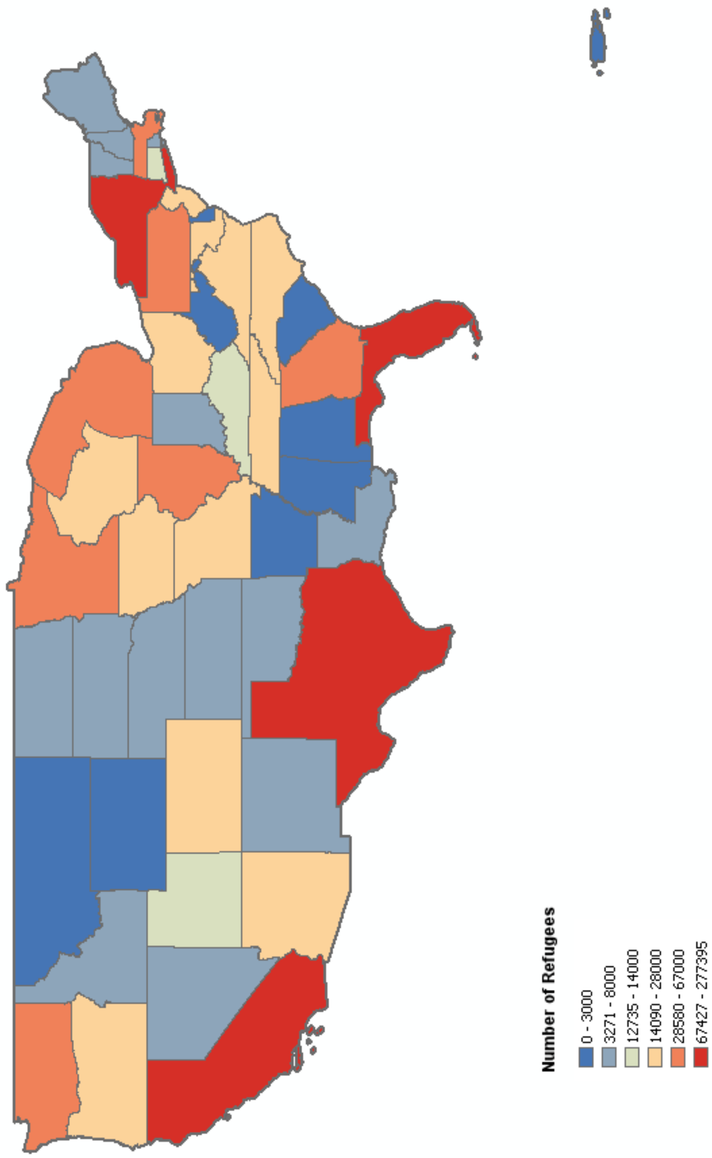


Figure 3.2 – The log of the number of refugees who immigrated to the United States over the period from 1985 to 2005.



3.8 Tables

Table 3.1 – Summary Statistics

		1970	1980	1990	2000	2005
Total	Nr. of immigrants	9,752	13,700	23,100	32,300	36,100
	Nr of immigrants last 5 years	1,708	3,054	5,508	7,783	8,056
	Value of imports (USD millions)	2330	8510	24500	66100	82700
	Nr. of nationalities among immig.	67	148	149	127	138
Latin American countries	Nr of immigrants	1,189	3,089	7,566	13,100	15,500
	Nr of immigrants last 5 years	360	997	2,292	3,682	4,125
	Share of immigrants (%)	12.20	22.55	32.74	40.47	42.88
	Share of immigrants last 5 years (%)	0.21	0.33	0.42	0.47	0.51
	Value of imports (USD millions)	133	621	2270	9550	12700
	Share of US imports	5.71	7.30	9.27	14.45	15.36
Hurricane affected countries	Nr of immigrants	80	253	1,051	1,871	2,260
	Nr of immigrants last 5 years	29	105	410	439	538
	Share of immigrants (%)	0.82	1.85	4.55	5.79	6.26
	Share of immigrants last 5 years (%)	1.72	3.43	7.44	5.65	6.68
	Value of imports (USD millions)	3.62	13.3	56.2	101	145
	Share of US Imports	0.16	0.16	0.23	0.15	0.18
Refugee countries	Nr of immigrants			8,664	12,567	25,993
	Nr of refugees last 5 years			288	425	266
	Share of immigrants (%)			37.51	38.91	72.00
	Share of refugees last 5 years (%)			5.23	5.46	3.30
	Value of imports (USD millions)			4340	15300	38600
	Share of US imports			17.71	23.15	46.67
	Nr. of nationalities among refugees			56	67	83

Note: Number of immigrants and refugees in thousands. Affected countries: Belize, El Salvadore, Guatemala, Honduras and Nicaragua. Value of exports and imports in current US dollars.

Table 3.2 – Refugee-Sending Countries over the 5 years prior to the census years: 1985 to 1990, 1995 to 2000 and 2001 to 2010

Refugees	Country	Refugees	Country	Refugees	Country
367995	USSR	267	Yemen	24	Senegal
284999	Vietnam	252	Central African Rep.	21	Lesotho
207944	Cuba	249	Cote d'Ivoire	20	Seychelles
168788	Yugoslavia	239	Syria	19	Bangladesh
113607	Somalia	216	Lebanon	15	Germany
108372	Laos	213	Jordan	14	Benin
85652	Iran	201	Sri Lanka	14	Morocco
77766	Iraq	187	Chad	11	Argentina
66419	Myanmar	179	Malaysia	11	Spain
55781	Cambodia	178	South Africa	10	UAE
46875	Liberia	162	Gambia, The	10	Saudi Arabia
40797	Sudan	161	Nepal	9	Madagascar
39682	Ethiopia	151	Egypt	9	Venezuela
32096	Afghanistan	148	Indonesia	8	Burkina Faso
30873	Bhutan	147	Cameroon	8	Greece
21088	Poland	123	Algeria	8	Italy
20596	Romania	109	Kuwait	7	United Kingdom
12856	Haiti	95	North Korea	6	Bahrain
12658	Sierra Leone	93	Ghana	6	Brunei
10909	Burundi	91	Guinea	5	Mali
9833	Thailand	88	Equatorial Guinea	4	Ant. and Bar.
9683	DR Congo	83	Djibouti	4	Maldives
5806	Czechoslovakia	74	Zimbabwe	4	Mexico
4207	Hungary	61	Philippines	4	Panama
2506	Colombia	61	Turkey	3	Canada
2476	Rwanda	51	Namibia	3	Iceland
1801	Kenya	50	Honduras	2	Belgium
1511	Nigeria	50	Mozambique	2	Bahamas
1333	Togo	46	India	2	Guinea-Bissau
1258	Rep. of Congo	40	Zambia	2	Peru
801	Bulgaria	37	El Salvador	2	Paraguay
666	Israel	36	Gabon	1	Switzerland
631	Nicaragua	34	Botswana	1	Guatemala
611	Albania	34	Costa Rica	1	Jamaica
427	Angola	34	Ecuador	1	Japan
404	Mauritius	31	Austria	1	Mongolia
387	Pakistan	31	Tunisia	1	Niger
373	China	31	Tanzania	1	Palau
344	Uganda	25	Libya	1	Portugal
269	Mauritania	24	Malawi	1	Taiwan

Table 3.3 – Gravity Equation with OLS estimates for the full sample including all countries and the restricted sample with Latin American countries only

Dependent variable	log(intermediate imports)					
	All countries			Latin American countries		
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of immigrants past 5 years)	0.0490***	0.0399***	0.0123*	0.0784***	0.0619*	0.0723**
	[0.00495]	[0.00560]	[0.00666]	[0.0189]	[0.0354]	[0.0313]
log(Distance)	-0.718***	-0.696***	-0.670***	-1.086***	-0.872***	-1.074***
	[0.0645]	[0.0653]	[0.0654]	[0.195]	[0.219]	[0.201]
Border	0.369*	0.349*	0.341	0.28	0.199	0.284
	[0.209]	[0.209]	[0.209]	[0.223]	[0.227]	[0.223]
log(No. of immigrants 10 years ago)		0.0208***			0.0461*	
		[0.00597]			[0.0279]	
log(No. of immigrants in 1960)			0.0577***			0.00969
			[0.00700]			[0.0395]
State x year fixed effects	yes	yes	yes	yes	yes	yes
Country of origin x year fixed effects	yes	yes	yes	yes	yes	yes
Observations	15633	15633	15633	2659	2659	2659
R-squared	0.923	0.942	0.932	0.933	0.956	0.945

Note: The dependent variables are total intermediate imports from country j to state i in period t. Distances are in kilometers from the capital city in country i to the capital city in state j. All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 3.4 – Gravity Equation with first and second stage IV estimates for all countries

Dependent variable	log(No. of immigrants past 5 years)			log(intermediate imports)		
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of refugees in past 5 years)	0.434***	0.270***	0.425***			
	[0.0220]	[0.0250]	[0.0220]			
log(No. of immigrants past 5 years)				0.160**	0.155***	0.133**
				[0.0634]	[0.0489]	[0.0622]
log(Distance)	-1.62***	-1.01***	-1.56***	-0.462	-0.24	-0.251
	[0.199]	[0.191]	[0.199]	[0.285]	[0.307]	[0.285]
log(No. of immigrants 10 years ago)		0.493***			0.0989	
		[0.0287]			[0.0924]	
log(No. of immigrants in 1960)			0.014***			0.0115**
			[0.003]			[0.00493]
State x year fixed effects	yes	yes	yes	yes	yes	yes
Country of origin x year fixed effects	yes	yes	yes	yes	yes	yes
Observations	1674	1674	1674	1674	1674	1674
R-squared	0.744	0.825	0.747	0.962	0.968	0.962
Cragg-Donald F	28.72	42.91	31.6	28.72	42.91	31.6
Kleibergen-Paap p-val	0.0195	0.0109	0.0128			

Note: The dependent variables are total intermediate imports from country j to state i in period t. The number of immigrants is instrumented by the number of refugees who arrived in the past 5 years in state i from country j. Distances are in kilometers from the capital city in country i to the capital city in state j. All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 3.5 – Gravity Equation with first and second stage IV estimates for Latin American countries

Dependent variable	log(No. of immigrants past 5 years)			log(intermediate imports)		
	(1)	(2)	(3)	(4)	(5)	(6)
log(Distance) x dPost	0.432***	0.163***	0.150***			
	[0.035]	[0.0335]	[0.0355]			
log(No. of immigrants past 5 years)				0.156**	0.131***	0.132***
				[0.079]	[0.060]	[0.064]
log(Distance)	-1.846***	-0.704***	-1.553***	-0.647*	-0.578*	-0.521
	[0.200]	[0.196]	[0.200]	[0.340]	[0.340]	[0.320]
Border	2.038***	0.758***	1.908***	0.821**	0.465	0.748**
	[0.241]	[0.212]	[0.238]	[0.377]	[0.353]	[0.365]
log(No. of immigrants 10 years ago)		0.559***			0.021*	
		[0.021]			[0.012]	
log(No. of immigrants in 1960)			0.022***			0.007*
			[0.002]			[0.004]
State x year fixed effects	yes	yes	yes	yes	yes	yes
Country of origin x year fixed effects	yes	yes	yes	yes	yes	yes
Observations	1946	1946	1946	1946	1946	1946
R-squared	0.809	0.88	0.814	0.948	0.953	0.948
Cragg-Donald F	15.24	17.9	15.1			
Kleibergen-Paap p-val	0.0352	0.0215	0.0318			

Note: The dependent variables are total intermediate imports from country j to state i in period t. The number of immigrants is instrumted by distance interacted with a time dummy for the period after Hurricane Mitch. Distances are in kilometers from the capital city in country i to the capital city in state j. All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 3.6 – Gravity Equation with OLS estimates for the 2010 sample including all countries and the restricted sample with Latin American countries only

Dependent variable	log(intermediate imports)					
	Latin American countries			All countries		
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of immigrants past 5 years)	0.0638*	0.0183	0.0623*	0.0352**	0.0203**	0.0366**
	[0.0349]	[0.0378]	[0.0349]	[0.0152]	[0.009]	[0.0152]
log(Distance)	-2.25***	-2.08***	-2.19***	-1.15***	-1.103***	-1.05***
	[0.227]	[0.233]	[0.230]	[0.0769]	[0.0776]	[0.0817]
Border	0.311	0.212	0.0464	1.85***	1.82***	1.81***
	[1.197]	[1.190]	[1.211]	[0.687]	[0.696]	[0.686]
log(No. of immigrants 10 years ago)		0.121***			0.0403**	
		[0.038]			[0.0165]	
log(No. of immigrants in 1960)			0.048			0.054
			[0.038]			[0.039]
State fixed effects	yes	yes	yes	yes	yes	yes
Country of origin fixed effects	yes	yes	yes	yes	yes	yes
Observations	710	710	710	3913	3913	3913
R-squared	0.798	0.799	0.798	0.787	0.789	0.787

Note: The dependent variables are total intermediate imports from country j to state i in period t. Distances are in kilometers from the capital city in country i to the capital city in state j. All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 3.7 – Gravity Equation with first and second stage IV estimates for all countries in the year 2010

Dependent variable	log(No. of immigrants past 5 years)			log(intermediate imports)		
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of refugees in past 5 years)	0.319***	0.229***	0.266***			
	[0.0372]	[0.0391]	[0.0500]			
log(No. of immigrants past 5 years)				0.274**	0.211**	0.225**
				[0.114]	[0.102]	[0.108]
log(Distance)	-1.988***	-1.191***	-1.873***	-1.017	-1.422**	-1.220*
	[0.398]	[0.372]	[0.227]	[0.794]	[0.690]	[0.676]
log(No. of immigrants 10 years ago)		0.388***			0.0354	
		[0.0393]			[0.140]	
log(No. of immigrants in 1960)			0.0745***			0.0789
			[0.0186]			[0.0817]
State fixed effects	yes	yes	yes	yes	yes	yes
Country of origin fixed effects	yes	yes	yes	yes	yes	yes
Observations	649	649	649	649	649	649
R-squared	0.816	0.868	0.832	0.769	0.777	0.766
Cragg-Donald F	22.14	23.1	24.55			
Kleibergen-Paap p-val	0.051	0.048	0.027			

Note: The dependent variables are total intermediate imports from country j to state i in period t. The number of immigrants is instrumented by the number of refugees who arrived in the past 5 years in state i from country j. Distances are in kilometers from the capital city in country i to the capital city in state j. All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 3.8 – Effect of Hurricane Mitch on total exports from hurricane-affected countries.

Dependent variable	log(Total exports)
Dummy_US_2000	0.811**
	[0.337]
Observations	948
R-squared	0.897

Note: Regression includes exporting country, importing country, time, importer-time and exporter-importer fixed effects. ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Bibliography

C. Aaron and M. Levy. The effect of mexican immigration on the wages and employment of u.s. natives: Evidence from the timing of mexican fertility shocks. *Paper presented at the 2013 Population Association of America*, 2013.

F. Alcalá and A. Ciccone. Trade and productivity. *The Quarterly Journal of Economics*, 119(2):613–646, 2004.

P. Allanson and C. Montagna. Multiproduct firms and market structure: An explorative application to the product life cycle. *International Journal of Industrial Organization*, 23(7):587–597, 2005.

T. Allen. Information frictions in trade. *Job-Market paper, Yale University*, 2011.

F. Alvarez and R. Jr. Lucas. General equilibrium analysis of the eaton kortum model of international trade. *Journal of Monetary Economics*, 54(6):1726–1768, 2007.

M. Amiti and J. Konings. Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *The American Economic Review*, pages 1611–1638, 2007.

J. E. Anderson and E. Van Wincoop. Gravity with gravitas: a solution to the border puzzle. *American Economic Review*, 93(1):170–192, 2003.

J. E. Anderson and E. Van Wincoop. Trade costs. *Journal of Economic Literature*, 42(3):691–751, September 2004.

J. E. Anderson and Y. V. Yotov. The changing incidence of geography. *American Economic Review*, 100:2157–2186, 2010.

A. Ardelean and V. Lugovskyy. Domestic productivity and variety gains from trade. *Journal of International Economics*, 80(2):280–291, 2010.

C. Arkolakis. Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy*, 118(6):1151–1199, 2010.

C. Arkolakis and M. A. Muendler. The extensive margin of exporting products: A firm-level analysis. Technical report, National Bureau of Economic Research, 2010.

C. Arkolakis and M. A. Muendler. Exporters and their products: A collection of empirical regularities. *CESifo Economic Studies*, 2012.

C. Arkolakis, P. Klenow, S. Demidova, and A. Rodriguez-Clare. Endogenous variety and the gains from trade. *American Economic Review Papers and Proceedings*, 98(4):444–450, 2008.

C. Arkolakis, A. Costinot, and A. Rodriguez-Clare. New trade models, same old gains? *American Economic Review*, 102(1):94–130, 2012.

R. Armenter and M. Koren. Economies of scale and the size of exporters. *FRB of Philadelphia Working Paper*, 2009.

A. Artopoulos, D. Friel, and J. C. Hallak. Export emergence of differentiated goods from developing countries: Export pioneers and business practices in argentina. *Journal of Development Economics*, 105:19–35, 2013.

D. H. Autor and D. Dorn. The growth of low skill service jobs and the polarization of the us labor market. *American Economic Review*, forthcoming, 2013.

D. H. Autor, D. Dorn, and G. H. Hanson. The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review*, Forthcoming, 2013.

D. Bahar and H. Rapoport. Migration, knowledge diffusion and the comparative advantage of nations. Technical report, Harvard University Working Paper, 2014.

B. Balassa. An empirical demonstration of classical comparative cost theory. *The Review of Economics and Statistics*, 45(3):231–238, 1963.

A. D. Bardhan and S. Guhathakurta. Global linkages of subnational regions: coastal exports and international networks. *Contemporary Economic Policy*, 22(2):225–236, 2004.

K. Behrens and Y. Murata. Trade, competition, and efficiency. *Journal of International Economics*, 87(1):1–17, 2012.

A. B. Bernard, J. Eaton, J. B. Jenson, and S. Kortum. Plants and productivity in international trade. *American Economic Review*, 93(4):1268–1290, September 2003.

A. B. Bernard, J. B. Jensen, S. J. Redding, and P. K. Schott. Firms in international trade. *Journal of Economic Perspectives*, 21(3):105–130, 2007a.

A. B. Bernard, S. J. Redding, and P. K. Schott. Comparative advantage and heterogeneous firms. *The Review of Economic Studies*, 74(1):31–66, 2007b.

A. B. Bernard, S. J. Redding, and P. K. Schott. Multiproduct firms and trade liberalization. *The Quarterly Journal of Economics*, 126(3):1271–1318, 2011.

S. Bertoli, J. Fernández-Huertas Moraga, and F. Ortega. Crossing the border: Self-selection, earnings and individual migration decisions. *Journal of Development Economics*, 101:75–91, 2013.

J. Bhagwati. The heckscher ohlin theorem in the multi commodity case. *Journal of Political Economy*, 1972.

C. Bo and D. S. Jacks. Trade, variety, and immigration. *Economics Letters*, 117(1):243–246, 2012.

G. J. Borjas. The economic analysis of immigration. *Handbook of labor economics*, 3:1697–1760, 1999.

H. P. Bowen, E. E. Leamer, and L. Sveikauskas. Multifactor, multicountry tests of the factor abundance theory. *American Economic Review*, 77(5): 791–809, 1987.

M. Bratti, L. De Benedictis, and G. Santoni. On the pro-trade effects of immigrants. *Review of World Economics*, pages 1–38, 2011.

P. Brenton and R. Newfarmer. Watching more than the discovery channel: export cycles and diversification in development. *Policy Research Working Paper*, 4302, 2007.

A. Briant, P. P. Combes, and M. Lafourcade. Product complexity, quality of institutions and the pro-trade effect of immigrants. 2009.

C. Broda and D.E. Weinstein. Globalization and the gains from variety. *Quaterly Journal of Economics*, 121(2):541–585, May 2006.

B. J. Bronnenberg, J. P. Dubé, and M. Gentzkow. The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, 102(6), 2012.

O. Cadot, C. Carrèère, and V. Strauss-Kahn. Export diversification: What's behind the hump? *The Review of Economics and Statistics*, 93(2):590–605, 2011.

O. Cadot, L. Iacovone, M. D. Pierola, and F. Rauch. Success and failure of african exporters. *Journal of Development Economics*, 101:284–296, 2013.

J. R. Campbell and H. A. Hopenhayn. Market size matters. *Journal of Industrial Economics*, 53(1):1–25, 2005.

D. Card. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19:22–64, 2001.

T. Cebeci, A. Fernandes, C. Freund, and M. Pierola. Exporter dynamics database. *World Bank Policy Research Working Paper*, (6229), 2012.

T. Chaney. Distorted gravity: the intensive and extensive margins of international trade. *The American Economic Review*, 98(4):1707–1721, 2008.

T. Chaney. The network structure of international trade. Technical report, National Bureau of Economic Research, 2011.

N. Chen, J. Imbs, and A. Scott. The dynamics of trade and competition. *Journal of International Economics*, 77:50–62, 2009.

D. Chor. Unpacking sources of comparative advantage: A quantitative approach. *Journal of International Economics*, 82(2):152–167, 2010.

X. Clark, T. J. Hatton, and J. G. Williamson. Explaining us immigration, 1971-1998. *The Review of Economics and Statistics*, 89(2):359–373, 2007.

P.-P. Combes, M. Lafourcade, and T. Mayer. The trade-creating effects of business and social networks: evidence from france. *Journal of International Economics*, 66(1):1–29, 2005.

A. Costinot and D. Donaldson. Ricardo's theory of comparative advantage: Old idea, new evidence. Technical report, National Bureau of Economic Research, 2012.

A. Costinot, D. Donaldson, and I. Komunjer. What goods do countries trade? a quantitative exploration of ricardo's ideas. *Review of Economic Studies*, 79:581–608, 2012.

M. David. *The United States Refugee Admissions Program*. Migration Policy Institute, 2004.

D. R. Davis and D. E. Weinstein. Economic geography and regional production structure: An empirical investigation. *European Economic Review*, 43(2):379–407, 1999.

D. R. Davis and D. E. Weinstein. An account of global factor trade. *American Economic Review*, 91(5):1423–1453, December 2001.

D. R. Davis and D. E. Weinstein. Market access, economic geography and comparative advantage: an empirical test. *Journal of International Economics*, 59(1):1–23, 2003.

J. De Loecker. Do exports generate higher productivity? evidence from slovenia. *Journal of International Economics*, 73(1):69–98, 2007.

J. de Sousa, T. Mayer, and S. Zignago. Market access in global and regional trade. *Regional Science and Urban Economics*, 42(6), 2012.

P. Debaere. Relative factor abundance and trade. *Journal of Political Economy*, 111(3):589–610, 2003.

A. Dixit and J. Stiglitz. Monopolistic competition and optimum product variety. *American Economic Review*, 67:293–308, 1977.

D. Dorn. *Essays on Inequality, Spatial Interaction, and the Demand for Skills*. PhD thesis, University of St. Gallen, 2009.

R. Dornbusch, S. Fischer, and P.A. Samuelson. Comparative advantage, trade, and payments in a ricardian model with a continuum of goods. *The American Economic Review*, 67(5):823–839, 1977.

J. A. Dunlevy. The influence of corruption and language on the protrade effect of immigrants: Evidence from the american states. *Review of Economics and Statistics*, 88(1):182–186, 2006.

J. A. Dunlevy and W. K. Hutchinson. The impact of immigration on american import trade in the late nineteenth and early twentieth centuries. *Journal of Economic History*, 59(4):1043–1062, 1999.

P. Dutt, I. Mihov, and T. Van Zandt. Does wto matter for the extensive and the intensive margins of trade? 2011.

J. Eaton and S. Kortum. Technology, geography, and trade. *Econometrica*, 70(5):1741–1779, 2002.

J. Eaton, S. Kortum, and F. Kramarz. An anatomy of international trade: Evidence from french firms. *Econometrica*, 79(5):1453–1498, September 2011.

J. Eaton, M. Eslava, D. Jinkins, C. J. Krizan, and J. Tybout. A search and learning model of export dynamics. *mimeo*, 2012.

C. Eckel and J. P. Neary. Multi-product firms and flexible manufacturing in the global economy. *The Review of Economic Studies*, 77(1):188–217, 2010.

P. H. Egger, M. Von Ehrlich, and D. R. Nelson. Migration and trade. *The World Economy*, 35(2):216–241, 2012.

W. J. Ethier. National and international returns to scale in the modern theory of international trade. *The American Economic Review*, pages 389–405, 1982.

S. J. Evenett and A. J. Venables. Export growth in developing countries: Market entry and bilateral trade flows. 2009.

FAO. Analysis of the medium-term effects of hurricane mitch on food security in central america. *Food and Agriculture Organization of the United Nations (FAO)*, 2001.

R. Feenstra. *US imports, 1972-1994: Data and concordances*, volume 5515. National Bureau of Economic Research, 1996.

R. Feenstra. *Advanced International Trade*. Princeton University Press, 2003.

R. Feenstra and H. Ma. Optimal choice of product scope for multiproduct firms under monopolistic competition. Technical report, National Bureau of Economic Research, 2007.

R. C. Feenstra, R. E. Lipsey, and H. P. Bowen. World trade flows, 1970–1992, with production and tariff data. *National Bureau of Economic Research Working Paper*, 5910, 1997.

G. J. Felbermayr and F. Toubal. Revisiting the trade-migration nexus: Evidence from new oecd data. *World Development*, 40(5):928–937, 2012.

G. J. Felbermayr, B. Jung, and F. Toubal. Ethnic networks, information, and international trade: revisiting the evidence. *Annals of Economics and Statistics / Annales d'Economie et de Statistique*, pages 41–70, 2011.

G. J. Felbermayr, V. Grossmann, and W. K. Kohler. Migration, international trade and capital formation: Cause or effect? Technical report, Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit, 2012.

A. C. Fieler. Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation. *Econometrica*, 79(4):1069–1101, 2011.

R. M. Friedberg and J. Hunt. The impact of immigrants on host country wages, employment and growth. *The Journal of Economic Perspectives*, pages 23–44, 1995.

M. Fujita and J.F. Thisse. *Economics of Agglomeration, Cities, Industrial Location and Regional Growth*. Cambridge University Press., 2002.

S. Girma and Z. Yu. The link between immigration and trade: Evidence from the united kingdom. *Review of World Economics/Weltwirtschaftliches Archiv*, 138(1):115–130, 2002.

S. S. Golub and C. T. Hsieh. Classical ricardian theory of comparative advantage revisited. *Review of International Economics*, 8(2):221–234, 2000.

H. Görg, M. Henry, and E. Strobl. Grant support and exporting activity. *The Review of Economics and Statistics*, 90(1):168–174, 2008.

D. M. Gould. Immigrant links to the home country: empirical implications for us bilateral trade flows. *The Review of Economics and Statistics*, pages 302–316, 1994.

J. Grogger and G. H. Hanson. Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, 95 (1):42–57, 2011.

H. G. Grubel and P. J. Lloyd. *Intra-industry trade: The theory and measurement of international trade in differentiated products*, volume 12. Macmillan London, 1975.

L. Halpern, M. Koren, and A. Szeidl. Imported inputs and productivity. *Hungarian Academy of Sciences, Central European University, and University of California, Berkeley*, 2009.

G. H. Hanson. International migration and the developing world. *Handbook of Development Economics*, 66:4363–4414, 2010.

J. Harrigan. Technology, factor supplies and international specialization: estimating the neoclassical model. *American Economic Review*, 87(4): 475–94, 1997.

J. Harrigan. *Handbook of international trade*, volume 1. Wiley-Blackwell, 2003.

R. Hausmann and D. Rodrik. Economic development as self-discovery. *Journal of Development Economics*, 72(2):603–633, 2003.

R. Hausmann, J. Hwang, and D. Rodrik. What you export matters. *Journal of Economic Growth*, 12(1):1–25, 2007.

J. Haveman and D. Hummels. Alternative hypotheses and the volume of trade: the gravity equation and the extent of specialization. *Canadian Journal of Economics/Revue canadienne d'économique*, 37(1):199–218, 2004.

K. Head and T. Mayer. The empirics of agglomeration and trade. *Handbook of Regional and Urban Economics edited by V. Henderson and J.F. Thisse.*, 4, 2004.

K. Head and T. Mayer. *Gravity equations: Workhorse, toolkit, and cookbook.* Centre for Economic Policy Research, 2013.

K. Head and J. Ries. Immigration and trade creation: econometric evidence from canada. *Canadian journal of economics*, pages 47–62, 1998.

E. Helpman. Trade, fdi, and the organization of firms. *Journal of Economic Literature*, 44:589–630, 2006.

E. Helpman, M. J. Melitz, and Y. Rubinstein. Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2):441–487, 2008.

A. Heston, R. Summers, and B. Aten. Penn world table version 6.3 center for international comparisons of production. *Income and Prices at the University of Pennsylvania*, 2009.

R. Hijmans, N. Garcia, and J. Wieczorek. Gadm: database of global administrative areas. 2010.

D. Hummels and P.J. Klenow. The variety and quality of a nations exports. *American Economic Review*, 95(3):704–23, 2005.

T. J. Kehoe and K. J. Ruhl. How important is the new goods margin in international trade? *Journal of Political Economy*, 121(2):358 – 392, 2013.

W. Kerr. Heterogeneous technology diffusion and ricardian trade patterns. *Working Paper*, 2008.

D. Kerwin. The faltering us refugee protection system: Legal and policy responses to refugees, asylum-seekers, and others in need of protection. *Refugee Survey Quarterly*, 31(1):1–33, 2012.

T. Khanna, K. G. Palepu, C. I. Knoop, and D. Lane. *METRO Cash & Carry*. Harvard Business School Case 707-505, 2009.

M. King, S. Ruggles, A. Trent, S. Flood, K. Genadek, M. Schroeder, B. Trampe, and R. Vick. Integrated public use microdata series. *Current Population Survey: Version*, 3(3), 2010.

P. Koenig. Agglomeration and the export decisions of french firms. *Journal of Urban Economics*, 66(3):186–195, 2009.

P. Koenig, F. Mayneris, and S. Poncet. Local export spillovers in france. *European Economic Review*, 54(4):622–641, 2010.

M. Koren and S. Tenreyro. Volatility and development. *The Quarterly Journal of Economics*, 122(1):243–287, 2007.

M. Koren and S. Tenreyro. Technological diversification. *American Economic Review*, forthcoming, 2012.

M. Koren and S. Tenreyro. Technological diversification. *The American Economic Review*, 103(1):378–414, 2013.

S. Krautheim. Heterogeneous firms, exporter networks and the effect of distance on international trade. *Journal of International Economics*, 87 (1):27–35, 2012.

P. Krugman. Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 16(4):670–675, 1980.

P. Krugman. Increasing returns and economic geography. *Journal of Political Economy*, 99:484–499, 1991.

A. Kugler and M. Yuksel. Effects of low-skilled immigration on us natives: evidence from hurricane mitch. Technical report, National Bureau of Economic Research, 2008.

M. Kugler and H. Rapoport. Migration, fdi, and the margins of trade. *Working Papers CID at Harvard University*, 222, 2011.

E. Leamer. *Sources of comparative advantage*. MIT Press,, Cambridge, Mass., 1984.

D. Lederman, M. Olarreaga, and L. Payton. Export promotion agencies: Do they work? *Journal of Development Economics*, 91(2):257–265, 2010.

A. A. Levchenko and J. Zhang. The evolution of comparative advantage: Measurement and welfare implications. Technical report, National Bureau of Economic Research, 2011.

A. Lileeva and D. Trefler. Improved access to foreign markets raises plant-level productivity for some plants. *The Quarterly Journal of Economics*, 125(3):1051–1099, 2010.

G. D. A MacDougall. British and american exports: A study suggested by the theory of comparative costs. part i. *The Economic Journal*, 61(244): 697–724, 1951.

A. M. Mayda. International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics*, 23(4): 1249–1274, 2010.

T. Mayer and S. Zignago. Notes on cepiis distances measures: the geodist database. *CEPII Working Paper 201125*, 2011.

T. Mayer, M. J. Melitz, and G. I. P. Ottaviano. Market size, competition, and the product mix of exporters. Technical report, National Bureau of Economic Research, 2011.

F. Mayneris and S. Poncet. Chinese firms entry to export markets: The role of foreign export spillovers. *World Bank Economic Review*, forthcoming, 2013.

M. J. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, 2003.

M. J. Melitz and G. I. P. Ottaviano. Market size, trade and productivity. *Review of Economic Studies*, 75:295–316, 2008.

E. Morales, G. Sheu, and A. Zahler. Gravity and extended gravity: Estimating a structural model of export entry. *unpublished working paper*, 2011.

G. I. P. Ottaviano, T. Tabuchi, and J. F. Thisse. Agglomeration and trade revisited. *International Economic Review*, 43(2):409–436, 2002.

Ç. Özden, C. Parsons, M. Schiff, and T. L. Walmsley. Where on earth is everybody? the evolution of global bilateral migration, 1960-2000. *World Bank Economic Review*, 25(1):12–56, 2011.

C. Parsons and P.-L. Vézina. Migrant networks and trade: The vietnamese boat people as a natural experiment. Technical report, Economic Research Institute for ASEAN and East Asia (ERIA), 2014.

N. Pavcnik. Trade liberalization, exit, and productivity improvements: Evidence from chilean plants. *Review of Economic Studies*, 69(1):245–76, January 2002.

G. Peri and F. Requena-Silvente. The trade creation effect of immigrants: evidence from the remarkable case of spain. *Canadian Journal of Economics/Revue canadienne d'économique*, 43(4):1433–1459, 2010.

F. Piguillem and L. Rubini. International trade and the firm size distribution. In *2012 Meeting Papers*, number 857. Society for Economic Dynamics, 2012.

J. E. Rauch. Networks versus markets in international trade. *Journal of international Economics*, 48(1):7–35, 1999.

J. E. Rauch and V. Trindade. Ethnic chinese networks in international trade. *Review of Economics and Statistics*, 84(1):116–130, 2002.

S. Redding. Specialization dynamics. *Journal of International Economics*, 58(2):299–334, 2002.

M. J. Roberts and J. R. Tybout. The decision to export in colombia: an empirical model of entry with sunk costs. *The American Economic Review*, pages 545–564, 1997.

J. Romalis. Factor proportions and the structure of commodity trade. *American Economic Review*, pages 67–97, 2004.

P. Romer. New goods, old theory, and the welfare costs of trade restrictions. *Journal of Development Economics*, 43(1):5–38, 1994.

S. Ruggles, M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. K. Hall, M. King, and C. Ronnander. Integrated public use microdata series (ipums): Version 3.0. *Minneapolis, MN: Minnesota Population Center. Available at: http://www. ipums. org. her, GA (1966)."*, 2004.

S. Sangita. The effect of diasporic business networks on international trade flows. *Review of International Economics*, 21(2):266–280, 2013.

P. K. Schott. One size fits all? heckscher-ohlin specialization in global production. *The American Economic Review*, 93(3), 686-708 2003.

P. K. Schott. Across-product versus within-product specialization in international trade. *The Quarterly Journal of Economics*, 119(2):647–678, 2004.

P. K. Schott. The relative sophistication of chinese exports. *Economic Policy*, 23(53):5–49, 2008.

S. Shikher. Capital, technology, and specialization in the neoclassical model. *Journal of International Economics*, 83(2):229–242, 2011.

JMC Silva and S. Tenreyro. The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658, 2006.

I. Simonovska and M.E. Waugh. The elasticity of trade: Estimates and evidence. Technical report, National Bureau of Economic Research, 2011.

C. Smith. The polarization of the us adult labor market and its effects on the demand for teenage labor. *Unpublished paper, Board of Governors, Federal Reserve System, May*, 2010.

M. I. Steinbaum. Migration and geography in a search and matching labor market. 2013.

J. H. Stock and M. Yogo. Testing for weak instruments in linear iv regression, 2002.

S. Stumpner. Trade and the geographic spread of the great recession. Technical report, Mimeo, University of Berkeley, 2013.

C. Syverson. Prices, spatial competition, and heterogeneous producers: An empirical test. *Journal of Industrial Economics*, 55(2):197–222, 2007.

C. M. Tolbert and M. Sizer. Us commuting zones and labor market areas: A 1990 update. *Economic Research Service Staff Paper*, (9614), 1996.

D. Trefler. The case of the missing trade and other mysteries. *The American Economic Review*, pages 1029–1046, 1995.

UN. *HS and SITC conversion and correspondence tables*. United Nations, 2012. URL http://unstats.un.org/unsd/trade/conversions/HS%20Correlation%20and%20Conversion%20tables.htm.

J. Van Biesebroeck. Exporting raises productivity in sub-saharan african manufacturing firms. *Journal of International economics*, 67(2):373–391, 2005.

P.-L. Vézina. Migrant networks as substitute for institutions: Evidence from swiss trade. Technical report, HEID Working Paper, 2010.

D. Wagner, K. Head, and J. Ries. Immigration and the trade of provinces. *Scottish Journal of Political Economy*, 49(5):507–525, 2002.

M. E. Waugh. International trade and income differences. *American Economic Review*, 100(5):2093–2124, 2010.

J. M. Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

Worldbank. *Hurricane Mitch: The Gender Effects of Coping with Crisis*. World Bank: Washington DC, 2001.

WTO. *Trade and public policies: a closer look at non-tariff measures in the 21st century*. World Trade Organization (WTO), Geneva, Switzerland, 2012.

C. Xiang. Diversification cones, trade costs and factor market linkages. *Journal of International Economics*, 71(2):448–466, 2007.

Appendix

Appendix Chapter 1

Lemma: The import expenditure distribution is Fréchet distributed. Import concentration on the intensive margin depends only on the degree of comparative advantage θ and the elasticity of substitution η.

Proof: Producers in country n import goods from country i if they obtain the lowest price for a particular good x in n

$$p(x)_{ni} = B(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})x_i^{\theta} \tag{4.11}$$

where $Bw_i^{\beta}p_{mi}^{1-\beta}$ are the unit costs of production in country i, $\kappa_{ni}\omega_{ni}$ are the trade costs from n to i, x_i is the realization of the random productivity draw and θ describes the variance of productivity draws. If we rewrite equation (3.8) in terms of x, the probability of country i to obtain a price smaller than p in country n is given by:

$$G_{ni}(p) = 1 - F_i(\lambda_i(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}} p^{\frac{1}{\theta}}) \tag{4.12}$$

where $F_i()$ is the cdf of the exponential distribution. The probability that p is the minimum price in country n is given by:

$$G_n(p) = 1 - \prod_{i=1}^{I}(1 - G_{ni}(p))$$

Substituting equation 4.12 into the previous equation, we can write the distribution function of minimum prices:

$$G_n(p) = 1 - \prod_{i=1}^{I}(e^{-\lambda_i(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}} p^{\frac{1}{\theta}}}) = 1 - (e^{-\sum_{i=1}^{I}\lambda_i(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}} p^{\frac{1}{\theta}}}) = 1 - (e^{-\varphi_n p^{\theta}})$$

where

$$\varphi_n = \sum_{i=1}^{I}\lambda_i(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}}$$

We proof two useful properties of the price distributions:

1.) The price distribution of country n is independent of source country i.

The probability that country i provides a good at the lowest price in country n is :

$$D_{ni} = \Pr[P_{ni} < min(P_{n1}...P_{ns}), s \neq i]$$

$$D_{ni} = \int_0^\infty \prod_{s\neq i}^{I}(1 - G_{ns}(p))dG_{ni}(p)$$

This means that the probability that country i obtains the minimum price in country n is the probability that no other country offers a lower price than p in country n, $\prod_{s\neq i}^{I}(1 - G_{ns}(p))$. To obtain the probability, we solve the following integral.

$$D_{ni} = \int_0^\infty (e^{-\sum_{s\neq i}^{I}\lambda_s(\frac{w_s^\beta p_{ms}^{1-\beta}}{\kappa_{ns}\omega_{ns}})^{-\frac{1}{\theta}}p^{\frac{1}{\theta}}})g_{ni}(p)dp$$

$$D_{ni} = \int_0^\infty \lambda_i(\frac{w_i^\beta p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}}\frac{1}{\theta}p^{\frac{1}{\theta}-1}e^{-\sum_{i=1}^{I}\lambda_i(\frac{w_i^\beta p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}}p^{\frac{1}{\theta}}}dp$$

Define $u = \sum_{s=1}^{N}\lambda_i(\frac{w_i^\beta p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}}p^{\frac{1}{\theta}}$ and $du = \varphi_n\frac{1}{\theta}p^{\frac{1}{\theta}-1}dp = \frac{1}{\theta}\frac{u}{p}dp$, we get:

$$D_{ni} = \frac{\lambda_i(\frac{w_i^\beta p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}}}{\varphi_n}\int_0^\infty \frac{1}{\theta}\frac{u}{p}(e^{-u})\frac{p}{u}\theta du$$

$$D_{ni} = \frac{\lambda_i(\frac{w_i^\beta p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}}}{\varphi_n}$$

Note that the distribution of prices that country i provides to country n is given by the conditional probability that i supplies the minimum price:

$$\Pr(p_n \leq p | p_{ni} \leq min_{j\neq i}(p_{nj})) = \frac{\Pr(p_n \leq p, p_{ni} \leq min_{j\neq i}(p_{nj}))}{\Pr(p_{ni} \leq min_{j\neq i}(p_{nj}))}$$

and note that $\Pr(p_{ni} \leq min_{j\neq i}(p_{nj})) = D_{ni}$ where $p_n = min_j(p_{nj})$. Now suppose that p_{ni} is the minimum price and calculate the joint probability by

$$\Pr(p_n \leq p, p_{ni} \leq min_{j\neq i}(p_{nj})) = \int_0^p \int_{p_{nj\neq i}}^\infty g(p_{ni})...g(p_{nj\neq i})dp_{nj\neq i}...dp_{ni}$$

By independence, we can solve for all integrals but $p_{n,i}$ and obtain

$$\Pr(p_n \leq p | p_{ni} \leq min_{j\neq i}(p_{nj})) = \frac{1}{D_{ni}} \int_0^p \prod_{s\neq i}^{N} (1 - G_{ns}(p_{ni})) dG_{ni}(p_{ni})$$

$$\Pr(p_n \leq p | p_{ni} \leq min_{j\neq i}(p_{nj})) = \frac{1}{D_{ni}} \frac{\lambda_i(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{n,i}\omega_{ni}})^{-\frac{1}{\theta}}}{\varphi_n} \int_0^p \varphi_n \frac{1}{\theta} q^{\frac{1}{\theta}-1} (e^{-\varphi_n q^{\frac{1}{\theta}}}) dq$$

Define $u = \sum_{s=1}^{N} \lambda_i(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}})^{-\frac{1}{\theta}} q^{\frac{1}{\theta}}$ and $du = \varphi_n \frac{1}{\theta} q^{\frac{1}{\theta}-1} dq = \frac{1}{\theta}\frac{u}{q} dq$, we get:

$$\Pr(p_n \leq p | p_{ni} \leq min_{j\neq i}(p_{nj})) = \frac{1}{D_{n,i}} D_{ni} \int_0^p (e^{-u}) du$$

$$\Pr(p_n \leq p | p_{ni} \leq min_{j\neq i}(p_{nj})) = 1 - e^{-\varphi_n p^{\frac{1}{\theta}}}$$

which is independent of the source country i. This implies that every country offers the same price distribution to country n.

2.) The import expenditure distribution in country n follows a Fréchet.

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

$$\Pr(imp) = \sum_{k\neq n}^{I} \Pr[P_{n,k} < min(P_{ns}), s \neq k]$$

$$\Pr(imp) = \sum_{k\neq n}^{I} \int_0^{\infty} \prod_{s\neq n}^{N} (1 - G_{ns}(p)) dG_{nk}(p)$$

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

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

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

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

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 countries:

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

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

$$q_n(x)p_n(x) = \left(\left(\min_{i\neq n} p_i(x)\right)^{1-\eta} p_{mn}^{\eta} q_n\right) = \left(\min_{i\neq n} B(\frac{w_i^{\beta} p_{mi}^{1-\beta}}{\kappa_{ni}\omega_{ni}}) x_i^{\theta})\right)^{1-\eta} p_{mn}^{\eta} q_n \tag{4.13}$$

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

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

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

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

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

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

Given that the price distribution is independent of the source country and follows a Fréchet distribution, we can calculate the corresponding concentration indexes analytically. The Theil index on the intensive margin for country n can be approximately written in terms of the continuos 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 in the density of the Fréchet distribution with location parameter $s_n = {k_n}^{-1}\varphi_n^{\theta(1-\eta)}$ and shape parameter $\alpha = \frac{-1}{\theta(1-\eta)}$, we get:

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

$\bar{R}_a$ is the mean import expenditure. Substituting for the mean of the Fréchet distribution $\bar{R}_a = s_n\Gamma(1-\frac{1}{\alpha})$, we can write the integral as:

$$T_n^W = \frac{1}{s_n\Gamma(1-\frac{1}{\alpha})}\left(\int_0^\infty \ln(R)\, Rf(R)dR - \int_0^\infty \ln\left(s_n\Gamma(1-\frac{1}{\alpha})\right) Rf(R)dR\right)$$

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

Performing the following change of variables $u = (R/s_n)^{-\alpha}$ and $dR = (-1/\alpha)u^{(-1/\alpha-1)}s_n du$ with $\lim_{R\to 0} u = \infty$ and $\lim_{R\to\infty} u = 0$, we get:

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

which simplifies to the following expression

$$T_n^W = \frac{1}{\Gamma(1-\frac{1}{\alpha})}\left(\int_0^\infty \ln\left(u^{(-1/\alpha)}\right) u^{(-1/\alpha)} e^{-u} du\right) - \ln\left(\Gamma(1-\frac{1}{\alpha})\right) \quad (4.14)$$

The Theil index on the intensive margin of imports does not depend on the scale parameter s_n and is hence 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 4.14 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.

The share of products exported

The concentration index of exports on the extensive margin 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 a country's probability to export a good is equal to the share of products exported. 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. $|\mathbf{B}_{jn} \cup \mathbf{B}_{kn}| = \mathbf{B}_{jn} + \mathbf{B}_{kn} - \mathbf{B}_{jn} \cap \mathbf{B}_{kn}$. 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

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

where A_i defines the event that a product is exported to destination i, i.e. A_i contains all products where country i obtains the minimum price in country n, $A_i = p_{n,i}(x) \leq \min_{j\neq n}[p_{n,j}(x)]$. 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

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

The last sum runs over all subsets L of the indices $1, ..., I-1$ 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 write the event to export to k destinations as the $L = k$ intersections of A,

$$a_k = (A_L)$$

and the set of products exported simplifies to

$$U_{ex} = \sum_{k=1}^{I-1} (-1)^{k-1} \left(\begin{array}{c} I \\ k \end{array} \right) a_k$$

where $I - 1$ is the total number of destinations. The resulting share of products exported can readily 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.

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

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 at home $p_{n,n} = Bx_n^\theta$. Notice that the only difference between prices offered is whether the country sells in the home market or in the 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 by the corresponding probability to obtain the minimum price in the respective destinations

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

By the properties of the exponential distribution, this probability is equal to

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

Proceeding in similar manner, one obtains 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

$$T_{i,X}^{ext} = -\ln\left(\sum_{i=1}^{I}(-1)^{k-1}\begin{pmatrix} I \\ k \end{pmatrix} a_k\right)$$

where

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

Trade Data

To build my empirical evidence, I use the Comtrade data set collected by the United Nations and choose the 6 digit HS 1992 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 to different varieties of the same product. I assume that the tradable goods sector corresponds to manufactures defined to be the aggregate across all 34 BEA manufacturingindustries. Using a correspondance table provided by the United Nations, I identify 4529 tradable manufacturing products. The baseline sample covers 160 countries representing all regions and all levels of development between 1992 and 2009 (16 years), including 129 developing countries, defined by the World Bank as countries with per capita GDP under $16,000 in constant 2005 PPP international dollars.

Following Waugh [2010], I assume that the tradable goods sector corresponds to manufactures. I construct trade shares D following Eaton and Kortum [2002] and Waugh [2010] in the following manner:

$$D_{i,j} = \frac{\text{Import}s_{i,j}}{\text{Gross Mfg. Production}_i \text{ - TotalExport}s_i + \text{Import}s_i}$$

In the numerator is the aggregate value of manufactured goods that country i imports from country j. These data are from the Unite Nations Comtrade database. Manufacturing goods are defined to be the corresponding HS codes of the the 34 BEA manufacturing industries used in Feenstra et al. [1997]. In the denominator is gross manufacturing production minus total manufactured exports (for the whole world) plus manufactured imports (for only the sample), see Eaton and Kortum [2002]. Basically, this is simply computing an expenditure share by dividing the value of inputs consumed by country i 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 from the World Bank. [15]

To get the traded goods of the manufacturing sector, I use the correspondance table provided by Feenstra et al. [1997] to identify manufacturing trade sectors at the 4 digit SITC rev.2 level. Using the correspondance table from the United Nations, I can then identify the corresponding HS 1992 6 digit manufacturing products. The resulting number of manufacturing products comprises 4529 product lines. [16]

SITC 4 digit industry classification

In the paper I analyzed total net trade flows at the 6 digit HS industry classification. This sections shows that the results found in the main part

15. The trade share matrix is the same as in Waugh but I since the product level data does not include South Africa, I renormalize the trade share such that the expenditure shares add up to 1.

16. The correspondance table between the SITC Rev.2 codes and the HS 1992 codes comes from the United Nations http://unstats.un.org/unsd/trade/conversions/HS Correlation and Conversion tables.htm

of the paper are not driven by the industry classification scheme and do apply in a more general sense. As a robustness check, I use the 4 digit SITC classification scheme which implies a total number of products of 642.

Table A.1 – Descriptive statistics for net trade flows based on the 4 digit SITC industry classification: 160 countries

	Gini		Theil Exports (X)			Theil Imports (M)		
	X	M	Ext. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	Total
Mean index (SITC 4 digit)	0.98	0.86	2.16	1.86	4.01	0.70	1.29	1.99
% share of overall concentration			54%	46%		35%	65%	
Mean index (HS 6 digit)	0.98	0.91	2.60	2.13	4.73	1.10	1.61	2.71
% share of overall concentration			55%	45%		40%	60%	
Correlation	0.95	0.88	0.94	0.91	0.95	0.50	0.82	0.64

Table A.1 present the descriptive statistics for the SITC sample and the correlation coefficient with the 6 digit HS net trade flow sample. The empirical estimates of the SITC industry classification are very similar. The correlation coefficient is 0.95 for exports and 0.64 for imports with respect to the 6 digit HS code sample. In addition, the cross country patterns are almost identical. Exports are more concentrated than imports on all margins. Strikingly, the L pattern of the extensive margin also appears when using the SITC classification. The differences across countries on the intensive margin are more pronounced as one would expect when choosing a higher level of aggregation. However, the overall pattern, with a cluster slightly above the 45 degree line, remains. Also, the shares of the decomposition are roughly the same. For exports, the intensive and extensive margin contribute almost to the same extent to the overall level of concentration. For imports, the intensive margin accounts for 66 percent of overall concentration.

Poisson parameter approach

The data contains 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 product units in the data. Suppose that the true level of disaggregation of Ricardian products, as defined in the Eaton and Kortum [2002] model, is unobserved and the classification scheme measures only an aggregate of those Ricardian products. For example, when products, in the sense of the Eaton and Kortum [2002] model, arrive at the boarder, custom agents aggregate those products into an industry according to the HS classification standard. The number of EK model products that custom 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, comprise one HS code (the observed product unit in the data).

To estimate the Poisson parameter, I proceed as follows. By 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. By independence, 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 just equals the share of products imported, N_M/N, we can use the definition of Theil index on the extensive margin,

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

to back out μ:

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

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 comprise an HS code.
code.

Empirical evidence and simulation results

In my simulation the total number of intermediate goods (N) is the product of the 4529 industries in the data times 0.94, the average number of products in an industry, $N = 4258$. 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.

Table A.2 – Mean concentration indexes for gross trade flows over 2676 country-year pairs

	Gini		Theil Exports (X)			Theil Imports (M)		
	X	M	Ext. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	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

Figure A.2 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 almost all countries and on all margins. The pattern on the extensive margin is displayed in Figure 4.2(d). Figure 4.2(c) shows the patterns on the intensive margin.

Turning our attention to 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 are different across the margins. In the case of the extensive margin, concentration decreases with respect to the net trade flow sample whereas on the 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 the respective product. 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. Table A.2 presents the summary statistics.

Table A.3 – Export and import concentration indexes for benchmark model and symmetric costs with Poisson parameter $\mu = 0.94$.

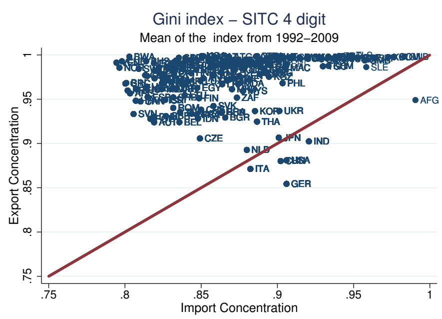
Parameters	Gini		Theil Exports (X)			Theil Imports (M)		
	X	M	Ext. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	Total
($\eta = 7.4$,$\kappa = 1$)	0.96	0.64	3.11	1.91	5.02	0.01	1.76	1.77
($\eta = 7.4$, $\kappa = 0.67$, NT=8)	0.99	0.96	1.83	2.55	4.38	0.94	1.76	2.70
Data (gross trade flows)	0.96	0.89	1.81	2.59	4.40	0.94	1.76	2.70
Data (net trade flows)	0.98	0.91	2.60	2.13	4.73	1.10	1.61	2.71

Table A.3 presents the simulation results. Given the benchmark parameters for θ, β and α, I calibrate the elasticity of substitution to the level of the intensive concentration of imports. I obtain a value of $\eta = 7.4$ slightly higher than in the net trade flow. The first row of table A.3 contains the resulting concentration indexes. The obtained concentration levels of exports are too high and too low for imports. In particular, the extensive concentration of imports is too low. For this reason, I introduce 50 percent trade costs ($\kappa = 0.67$). The introduction of symmetric trade costs does not reduce the gap between import and export concentration. Consequently, I reduce the number of trading partners to 8 partners in order to ease the competition in the export and domestic market. This increases the probability of obtaining the minimum price in foreign markets and reduces the extensive concentration of exports. The results are showns in row 2 in Table A.3 and match the data. Overall, the parameters chosen to replicate the mean concentration levels of gross trade flows are similar to the one found in the case of net trade flows. The exception are the number of trading partners and trade costs, which are with 0.67 significantly higher. On the other hand, the elasticity of substitution $\eta = 7.4$ is very close to the one found in the net trade version.

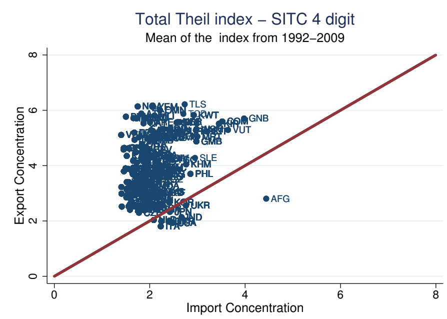
Table A.4 – Simulated concentration level with Poisson parameter $\mu = 0.94$ and exporter fixed effect

	Gini		Theil Exports (X)			Theil Imports (M)		
Model	X	M	Ext. Mar.	Int. Mar.	Total	Ext. Mar.	Int. Mar.	Total
Simulation	0.99	0.89	4.97	3.32	8.29	1.15	1.76	2.91
			60%	40%		39%	61%	
Data	0.96	0.89	1.81	2.59	4.40	0.94	1.76	2.70
			41%	59%		34%	66%	

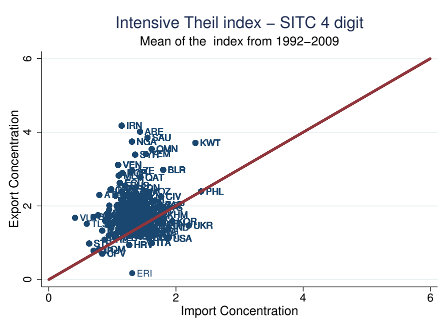
In this paragraph, I estimate trade cost and technology parameters based on bilateral trade shares using the models structure and check whether for given trade shares, the model is able to generate the observed specialization pattern in the gross trade data. Based on the estimated trade costs and technology levels, I simulate the model, calculate the resulting concentration indexes using the Poisson measurement device and compare the simulated results with the data. Table A.4 shows the results. Export concentration is higher than import concentration on all margins. With respect to the decomposition, similar to the net trade case, the extensive margin dominates overall concentration for exports and the intensive margin dominates for imports. The obtained concentation levels of 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, similar to the net trade results, the model cannot capture the negative relationship between import concentration and GDP.



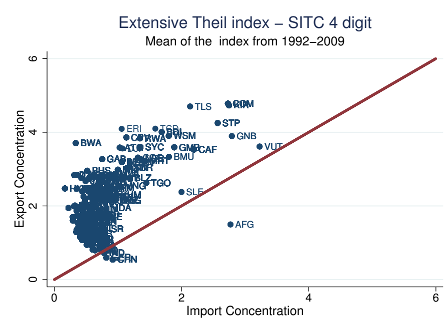
(a) Gini coefficient



(b) Theil index

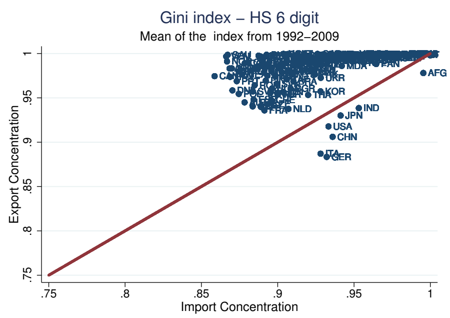


(c) Intensive margin

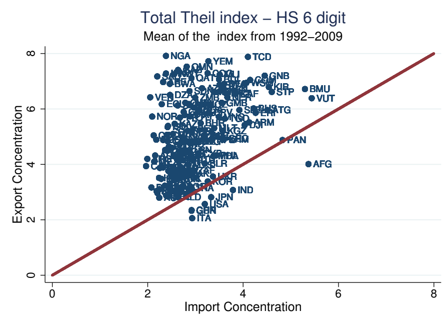


(d) Extensive margin

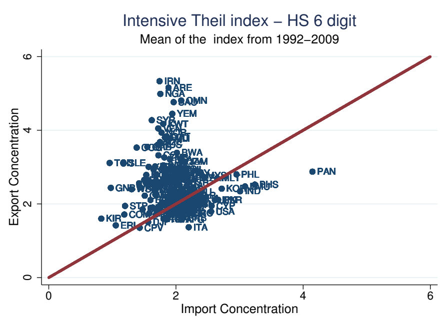
Figure A.1 – Export versus import concentration on the 4 digit SITC level



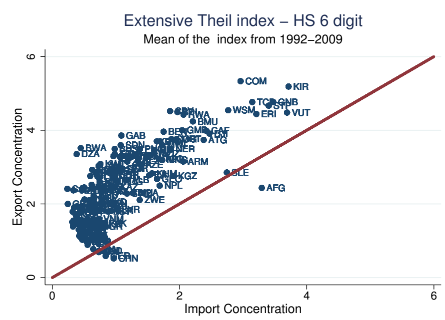
(a) Gini coefficient



(b) Theil index

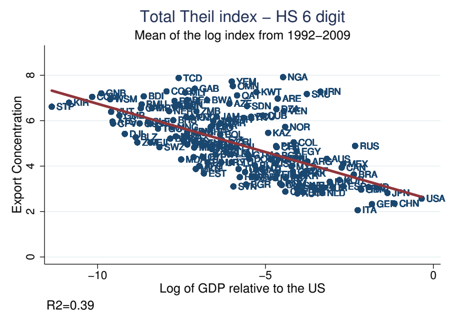


(c) Intensive margin

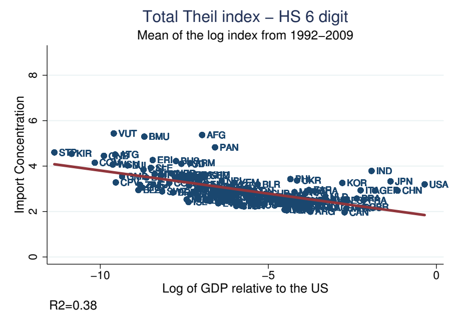


(d) Extensive margin

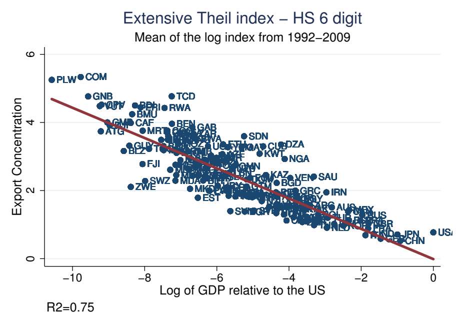
Figure A.2 – Export versus import concentration for gross trade flows



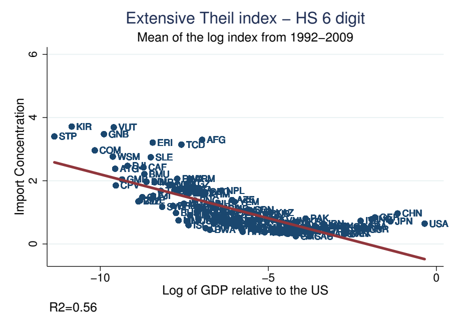
(a) Overall concentration of exports



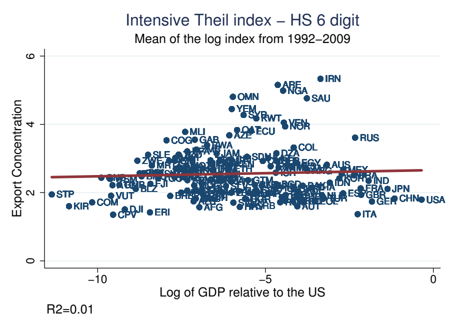
(b) Overall concentration of imports



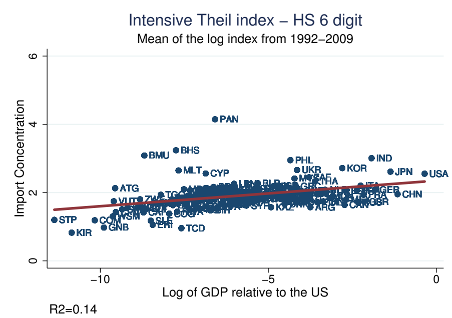
(c) Extensive margin of exports



(d) Extensive margin of imports



(e) Intensive margin of exports



(f) Intensive margin of imports

Figure A.3 – The relationship of export and import concentration verus GDP across 160 countries based on gross trade flows.

Appendix Chapter 2

Robustness

I conduct robustness tests to support the empirical result that product fixed costs are the dominant entry barrier. I also look at origin market characteristics. In particular, following the argument of Hummels and Klenow [2005] I suspect that the size of the home market is important to overcome fixed costs to export.

Cross-country differences

In this subsection, I analyze cross-country differences in the estimated entry elasticities. In regression 2.5, I pooled all observations across countries and reported a common entry elasticities for all countries. Now I impose less restrictions and allow for different entry elasticities depending on the origin country c. Instead of running country by country OLS regressions, I pool all observations to explicitly account for potential correlation across origin countries in the destination. I then test whether entry elasticities differ across exporting countries. I reject the hypothesis of a common slope coefficients on market size at the 1 percent level.[17] Given cross country differences, I test for differences between the product and the firm elasticity on a country per country basis. The results show that for 37 countries the entry elasticity of products with respect to market size is smaller than for firms, in 2 cases I do not find any significant differences and in the case of Niger the entry elasticity of firms is lower than for products. Overall, the results that fixed costs at the product level is the main entry barrier applies to the majority of the countries in the sample.

Digging deeper into cross-country differences, I investigate whether entry elasticities vary with the market size of the exporting country. Hummels and Klenow [2005] suggest that the size of the home market is important to overcome fixed costs to export. Economies of scale imply that firms make more profits in larger markets. Thus, if home sales are important to pay for fixed costs, firms from a larger home market have a competitive advantage over firms from a smaller market simply because they operate at a larger

17. This finding is contrary to Eaton et al. [2011], who do not find significant differences in the entry elasticity of firms for France, Denmark and Uruguay.

scale. As a result, firms from a larger home market will have higher entry rates than firms from smaller economies. To investigate whether home market size matters, I estimate regression 2.5 and include an interaction term of the log of destination market size with the log of home market size. We expect that the entry elasticity of firms with respect to destination market size is higher for countries with a larger home market. If there is no home market effect, only export sales are relevant and the change in the number of firms and products should be independent of home market size. [18]

I find that entry elasticities increase significantly with origin market size. Given an increase in demand (i.e. increasing the destination market size), relative more firms and products from larger home markets enter. The larger revenue in domestic markets facilitates firm and product entry in all destinations. This interpretation implies that products from larger markets have a production efficiency advantage over products from smaller markets because the exporting firms operate at a larger scale. Small economies are particularly affected from fixed costs. The low level of domestic demand implies that relative few firms are able to benefit from economies of scale resulting in less export activity and relative high prices for domestic consumers. Moreover, because of the low demand, few foreign firms find it profitable to export to them. The limited entry results in even higher prices due to the lack of competitive pressure.

Poisson Maximum Likelihood

Silva and Tenreyro [2006] argue that in the presence of heteroskedasticity the elasticity estimates in Table 2.4 are biased. Consider equation 2.3 with the respective elasticities. To allow for deviations from the theory, we write

$$N_{d,c,t} = \pi_{d,c,t}^{b} X_{d,t}^{\gamma} \eta_{d,c,t} \tag{4.15}$$

where $\eta_{d,c,t}$ is an error factor with $E(\eta_{d,c,t}|X_d, \pi_{d,c}) = 1$. As Silva and Tenreyro [2006] show the standard practice of log-linearizing equation 4.15 and estimating γ by OLS is inappropriate for mainly two reasons. First, $N_{d,c}$ can be 0, in which case log-linearization is infeasible. This is not an issue. If there are no exporters, then there is no trade. Second, even if all observations of $N_{d,c}$ are strictly positive, the expected value of the log-linearized error will

18. Table A.5 in the appendix shows the results.

in general depend on the covariates, and hence OLS will be inconsistent. To see the point more clearly, notice that equation 4.15 can be expressed as $y_{d,c,t} = \exp(\beta Z_{d,c,t})\eta_{d,c,t}$, with $E(\eta_{d,c,t}|Z_{d,c,t}) = 1$. Assuming that $y_{d,c,t}$ is positive, the model can be made linear in the parameters by taking logarithms of both sides of the equation, leading to $\log y_{d,c,t} = \beta Z_{d,c,t} + \log \eta_{d,c,t}$. To obtain consistent estimates of β, it is necessary that $E(\log \eta_{d,c,t}|Z_{d,c,t})$ does not depend on $Z_{d,c,t}$. Notice that the expected value of the logarithm of a random variable depends both on its mean and on the higher-order moments of the distribution. However under the presence of heteroskedasticity, the expected value of $\log \eta_{d,c,t}$ will also depend on the regressors, rendering the estimates of β inconsistent. For example, suppose that $\eta_{d,c,t}$ is log normally distributed with $E(\eta_{d,c,t}|Z_{d,c,t}) = 1$ and variance $\sigma^2_{d,c,t} = f(Z_{d,c,t})$. The error term of the log linearized representation will then follow a normal distribution, with $E(\log \eta_{d,c,t}|Z_{d,c,t}) = -1/2 \log(1 + \sigma^2_{d,c,t})$, thus implying inconsistency.

To estimate the elasticities, i.e. β, in equation 4.15 consistently, Silva and Tenreyro [2006] suggests a Poisson pseudo-maximum-likelihood estimator. Before applying the Poisson pseudo-maximum-likelihood estimator, I test for the presence of heteroskedasticity in equation 2.5 for each of the two different dependent variables. In all tests, I reject the null hypothesis of homoskedasticity at the 1 precent level.

Table A.6 plots the results for Poisson pseudo-maximum-likelihood approach. Qualitatively the results do not change. The signs of the coefficients do not change with respect to the log linear results. The elasticity of firms with respect to market size is significantly higher than for products implying that firm fixed costs are more important than product fixed costs. The key differences are quantitatively. All estimated elasticities slightly increase. The estimated entry elasticity of firms is 1.05 and 0.62 for products.

Table A.5 – Relationship between market size and the number firms and products taking into account cross-country differences

Dependent variable	log(Number of firms) (1)	log(Number of products) (2)
log(Market Size - Destination)	-0.215***	-0.0890***
	[0.0243]	[0.0262]
log(Market Size - Destination) * log(Market Size - Origin)	0.127***	0.0868***
	[0.00469]	[0.00505]
log(Distance)	-0.792***	-0.822***
	[0.00866]	[0.00933]
log(Expenditure Share)	0.200***	0.196***
	[0.00294]	[0.00317]
log(GDP per capita)	0.144***	0.0992***
	[0.00488]	[0.00526]
Border	0.332***	0.301***
	[0.0157]	[0.0169]
Constant	2.234***	2.675***
	[0.04]	[0.05]
Observations	28.978	28.978
R-squared	0,678	0,635

Note: All regressions include time and origin country fixed effects. Robust standard errors in parentheses (clustered by country time): ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Table A.6 – Pseudo Poisson Maximum Likelihood

Dependent variable	log(Number of firms) (1)	log(Number of products) (2)	log(Number of firms) (3)	log(Number of products) (4)
log(Market Size)	1.049***	0.619***	0.866***	0.557***
	[0.0177]	[0.00806]	[0.0155]	[0.00874]
log(Expenditure Share)	1.244***	0.966***	0.732***	0.594***
	[0.0244]	[0.0102]	[0.0189]	[0.0109]
log(Distance)			-0.836***	-1.036***
			[0.0357]	[0.0220]
log(GDP per capita)			0.371***	0.0785***
			[0.0252]	[0.0143]
Border			0.380***	-0,021
			[0.0566]	[0.0345]
Observations	30164	30164	30164	30164
R-squared	0,612	0,679	0,661	0,694

Note: The results from Poisson maximum likelihood regressions for the dependent variable noted at the top of each column projected on the covariates listed in the first column. All regressions include origin country, time and origin country - time fixed effects. Robust standard errors in parentheses: ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Appendix Chapter 3

Estimation of industry specific trade costs

To consistently estimate the effect of trade costs on trade flows, we follow the literature on estimation of gravity equations Head and Mayer [2013]. A wide variety of trade models yields an expression for trade flows that can be written in log form as follows:

$$\log(X^k_{ijt}) = \sigma \log(t^k_{ijt}) + f^k_{it} + f^k_{jt} + e^k_{ijt} \tag{4.16}$$

In this expression, f^k_i denotes all factors that promote exports of industry k in country i to all destinations, and f^k_j all factors that promote imports. σ represents the trade cost elasticity. We are only interested in estimation the bilateral trade costs independent of the origin and destination specific effects. To do so, we follow Stumpner [2013] and consider the following industry specific trade costs function

$$\log(t^k_{ijt}) = \delta_1 \log(d_{ij}) - \delta_2 \log(d_{ij}) \log(v^k_t) - \delta_3 b_{ij} \tag{4.17}$$

where d_{ij} is distance measured from capital in state i to the capital in country j. v^k is the value to weigh ratio calculated by the total value in dollars divided by the total weight in tons. b_{ij} indicates whether state i shares a border with a land with country j. We can now plug in the trade cost function, equation 4.17, into the trade flow equation, equation 4.16, and estimate the distance, value to weight and border elasticities.

$$\log(X^k_{ijt}) = \delta_1 \log(d_{ij}) - \delta_2 \log(d_{ij}) \log(v^k_t) - \delta_3 b_{ij} + f^k_{it} + f^k_{jt} + e^k_{ijt} \tag{4.18}$$

Since we do not observe regional trade flows and we cannot estimate equation 4.18 at the state level. Instead, we consider the following two alternatives to construct trade costs. First, we suppose that the inter state trade cost function from Stumpner [2013] holds on the international level as well. He finds $\delta_1 = 0.666$, $\delta_2 = -0.0803$ and $\delta_3 = 0$. Alternatively, we can estimate equation 4.18 at the international level using aggregate US international trade flows and and suppose that the parameters are valid at the regional trade level. To estimate the effect of trade costs on trade flows, we re-write equation 4.18 at the national level by aggregating trade flows over all US states. Aggregate US intermediate imports, X^k_j, from country j in year t are now given by

$$\log(X_{jt}^k) = \delta_1 \log(d_j) + \delta_2 \log(d_j) \log(v_t^k) + \delta_3 b_j + \delta_4 w_{jt} + d_t^k + e_j^k \quad (4.19)$$

where d_j is the distance from the capital of country j to the capital of the United States. b_j is a border dummy equal to 1 for Mexico and Canada and v_t^k is the value to weight in US dollars per ton shipped. In addition, we use industry-time fixed effects to control for industry specific variations. w_{jt} are control variables that are common to all regions within the United States but maybe correlated with distance and/or border. The bilateral control variables are a common currency indicator, a trade agreement indicator and an indicator whether the partner country shares a common language with the US. Furthermore, we include PPP converted GDP and population from the Penn World Table. Results are in Table A.7. As expected, they show a strong negative effect of distance on trade flows, which is muted in industries with high value-to-weight ratios. Sharing a border with the United States increases the exports to the United States.

Given the estimated parameters, we suppose that the estimated aggregate trade costs elasticities ($\hat{\delta}_1 - \hat{\delta}_3$) hold on the state level. As a result, we calculate industry specific trade costs using

$$\log t_{ijt}^k = 0.691 \log d_{ij} - 0.154 \log d_{ij} \log v_t^k - 1.897 b_{ij}$$

where d_{ij} represents the distance from the capital of state i to the capital of country j. b_{ij} equals one if the state i shares a land border with Canada or Mexico and v_t^k remains the average value-to-weight ratio of industry k at year t.

Calculation of the value-to-weight ratio

We calculate the average value to weight ratio per SIC 87 industry as follows. The data used are drawn from Feenstra [1996] for 1972 and 1980 and Schott [2008] for the years 1990, 2000 and 2005. They record the customs value of all US imports by exporting country and year from 1990 to 2005 according to SIC 87 codes and the corresponding value and weight of trade shipped by air and shipped by sea. For the years 1972 and 1980, we convert the seven-digit Tariff Schedule of the United States (TSUSA) to SIC87 codes following the correspondence table of Feenstra [1996]. Next, we compute the total import value and the average weight for each industry code and for each transportation mode, i.e. shipped by air or by sea. To compute the average value to weight ratio, we divide the total value shipped by its weight and convert it to tons for each mode of transportation and industry pair. We then weight the importance of the transportation mode by its share in the total per industry import value.

Calculation of intermediate imports

We classify goods according to the Broad Economic Categories (BEC) industry classification scheme of the United Nations. Intermediate goods are defined as the sum of the categories: Processed food and beverages (12), Industrial supplies (2), Capital goods (4) and Parts of transport equipment (53). To map the import data into the BEC categories, we use the concordance table from the World Bank that created a mapping between SITC Revision 2 codes and BEC categories. To convert the SITC Revision 2 codes into SIC 1987 codes, we follow the procedure outlined in Feenstra [1996]. Finally, intermediate imports are then simply the total value of all intermediate imports per SIC code.

Immigration response to hurricane Mitch

In this paragraph, we test whether hurricane Mitch increased immigration. We apply the difference in difference estimation procedure. Consider the following regression model

$$\log Imm_{ijt} = \beta_0 + \beta_1 X_{ij} + \beta_2 D_{ijt} + d_{ij} + b_{ij} + f_{it} + f_{jt} + \epsilon_{ijt}$$

where Imm_{ijt} is number of immigrants from country j living in state i at time t. X_{ij} is dummy variable equal to 1 if immigrants are from the hurricane affected countries. d_{ij} controls for distance and b_{ij} for border effects between state i and country j. f_{it} are state-year and d_{jt} origin country year fixed effects. D_{ijt} is a dummy variable equal to 1 if country j was affected by hurricane Mitch and the observation period is year 2000 or 2005. ϵ_{ijt} is an error term. The variable of interest is β_2 and measures the effect of hurricane Mitch on immigration from the affected countries, i.e. El Salvador, Guatemala, Honduras and Nicaragua, to the United States. The control group consists of Latin American immigrants, i.e. immigrants from the following countries: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, Mexico, Panama, Paraguay, Peru, Uruguay and Venezuela. We distinguish between 2 definitions of immigrants: classical and new. A classical immigrant is a foreign born individual that migrated to the US at any point in time. On the other hand, a new immigrant is a foreign born individual that migrated to the US within the last 5 years.

Table A.8 shows the results. The treatment dummy is significant and negative implying that the level of migration from the hurricane affected countries is significantly lower than from other Latin American countries. The coefficient on the difference in difference dummy is statistically significant. The number of immigrants and the number of newly arrived immigrants increased significantly after hurricane Mitch. Column (1) shows that while the number of immigrants from affected countries is 74.3 percent lower than from other Latin American countries, the number of immigrants from those countries significantly increased by 53 percent after the hurricane. Column (2) presents the results for the new immigrants who arrived during the last 5 years. The estimates reveal that new immigrants increased more strongly from affected countries in comparison to other Latin American countries. After the hurricane, the number of new immigrants per state grew even stronger by an additional 75.3 percent. Overall, the evidence in Table A.8 supports our prior that hurricane Mitch caused significantly more immigration from the affected countries to the United States in comparison to other Latin American countries.

Table A.7 – Effect of transport costs on industry trade flows in gravity model.

Dependent variable	log(intermediate imports)
log(Distance)	-0.691***
	[0.107]
log(Distance)*log(Value-to-weight)	0.154***
	[0.022]
Border	1.897***
	[0.430]
Trade agreement	0.420***
	[0.062]
Common language	0.235***
	[0.035]
Common currency	0.093*
	[0.057]
log(GDP of exporter)	0.990***
	[0.011]
log(Population of exporter)	0.024*
	[0.012]
Industry FE	Yes
Observations	61796
R-squared	0.449

Note: Regression includes exporting country, importing country, time, exporter-time and exporter-importer fixed effects. ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.

Table A.8 – Difference in Difference estimation: Number of migrants after Hurrican Mitch

Dependent variable	log(number of total immigrants)	log(number of immigrants last 5 years)
	(1)	(2)
Treatment Dummy	-0.743***	-0.760***
	[0.254]	[0.202]
Difference in Difference	0.530*	0.753*
	[0.283]	[0.385]
log(Distance)	-0.991***	-1.237***
	[0.0461]	[0.0376]
Border	1.111***	1.351***
	[0.127]	[0.117]
Constant	13.26***	15.64***
	[0.427]	[0.346]
Observations	15,633	15,633
R-squared	0.723	0.817

Note: All regressions include origin country, host country and time fixed effects. Robust standard errors in parentheses (clustered by country time): ***, **, * marks statistically significant difference from zero at the 1%, 5% and 10% level respectively.