Development Perspectives on Trade Growth at the Extensive Margin

Benjamin SHEPHERD

Thèse dirigée par M. leProfesseur Patrick Messerlin

Soutenue en mars 2008

Jury :
M. Chad P. Bown, Associate Professor of Economics, Brandeis University
M. Joseph F. Francois, Professor of Economics, Johannes Kepler Universität (Rapporteur)
M. Bernard M. Hoekman, Senior Advisor and Research Manager (Trade), The World Bank (Rapporteur)
M. Patrick A. Messerlin, Professeur des Universités à l’Institut d’Etudes Politiques de Paris (Directeur de recherche)
# Development Perspectives on Trade Growth at the Extensive Margin

## Contents

**Abstract**  
5  

**Résumé**  
6  

**Acknowledgements**  
7  

### General Introduction  
11  

Extensive Margin Growth and Firm Heterogeneity: A Development Perspective  
13  

Trade Policy, Regulatory Barriers, and Institutions  
17  

Structure of the Dissertation and Outline of Main Results  
20  

**Chapter One:** "Trade Costs, Barriers to Entry, and Export Diversification in Developing Countries"  
20  

**Chapter Two:** "Exporting to New Markets: An Empirical Investigation"  
22  

**Chapter Three:** "Export Variety and International Harmonization of Product Standards"  
23  

**References**  
24  

### I Trade Costs, Barriers to Entry, and Export Diversification in Developing Countries  
31  

**A Introduction**  
31  

**B Data and Stylized Facts**  
35  

1 Measuring Export Diversification  
35  

2 Measuring the Costs of Trade and Domestic Market Entry  
38  

3 Non-Parametric Evidence  
40  

**C Theoretical Model**  
42  

1 Consumption Block  
42
Development Perspectives on Trade Growth at the Extensive Margin

2 Production Block ......................................................... 43
3 Characterizing Export Diversification and Simulating the Model ...... 47
D Empirical Model and Results ........................................... 51
1 Instrumental Variables Strategy ........................................ 55
2 Alternative Measures of Diversification, Entry Costs, and Trade Costs . 58
3 Additional Independent Variables ...................................... 59
4 Narrower Country Samples ............................................. 60
5 Alternative Estimators ................................................... 61
E Conclusions and Policy Implications ................................... 61
F References ...................................................................... 63
G Tables ........................................................................... 67
H Figures ........................................................................... 73

II Exporting to New Markets: An Empirical Investigation ................. 77
A Introduction .................................................................... 77
B Theoretical Motivation .................................................... 81
1 Consumption Block ....................................................... 82
2 Production Block ......................................................... 84
3 Characterizing Geographical Diversification ........................... 86
C Preliminary Analysis: Firm Level Empirics ............................. 89
1 Productivity Estimates ................................................... 90
2 Productivity and the Number of Export Markets .................... 94
D Country Level Empirics .................................................. 96
1 Data ........................................................................... 97
2 Model, Estimation Results, and Robustness Checks ................. 100
### Development Perspectives on Trade Growth at the Extensive Margin

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Robustness Checks</td>
<td>102</td>
</tr>
<tr>
<td>4</td>
<td>Instrumental Variables Results</td>
<td>104</td>
</tr>
<tr>
<td>E</td>
<td>Conclusions and Further Research</td>
<td>107</td>
</tr>
<tr>
<td>F</td>
<td>References</td>
<td>109</td>
</tr>
<tr>
<td>G</td>
<td>Tables</td>
<td>113</td>
</tr>
</tbody>
</table>

### III Product Standards, Harmonization, and Trade: Evidence from the Extensive Margin 127

<table>
<thead>
<tr>
<th>Part</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Introduction</td>
<td>127</td>
</tr>
<tr>
<td>B</td>
<td>Theoretical Motivation</td>
<td>131</td>
</tr>
<tr>
<td>1</td>
<td>Consumption Block</td>
<td>133</td>
</tr>
<tr>
<td>2</td>
<td>Production Block</td>
<td>134</td>
</tr>
<tr>
<td>3</td>
<td>Modeling Product Standard Harmonization</td>
<td>136</td>
</tr>
<tr>
<td>4</td>
<td>Simulation Results</td>
<td>139</td>
</tr>
<tr>
<td>C</td>
<td>Data and Stylized Facts</td>
<td>141</td>
</tr>
<tr>
<td>1</td>
<td>The EU Standards Database</td>
<td>141</td>
</tr>
<tr>
<td>2</td>
<td>Measuring Export Variety</td>
<td>143</td>
</tr>
<tr>
<td>D</td>
<td>Empirical Model and Estimation Results</td>
<td>145</td>
</tr>
<tr>
<td>1</td>
<td>Instrumental Variables Results</td>
<td>147</td>
</tr>
<tr>
<td>2</td>
<td>Robustness Checks: Quotas, Income Levels, and Regional or Preferential Agreements</td>
<td>150</td>
</tr>
<tr>
<td>E</td>
<td>Conclusions, Policy Implications, and Future Research</td>
<td>152</td>
</tr>
<tr>
<td>F</td>
<td>References</td>
<td>154</td>
</tr>
<tr>
<td>G</td>
<td>Tables</td>
<td>158</td>
</tr>
<tr>
<td>H</td>
<td>Figures</td>
<td>162</td>
</tr>
</tbody>
</table>
General Conclusion

Policy Implications .......................................................... 165

Directions for Future Research .......................................... 169
  Endogeneity of Trade Costs ............................................. 169
  Data ................................................................................. 170
  Technological Change and Productivity Growth .................. 171

References .......................................................................... 172
Abstract

This dissertation examines the links between international trade and economic development through the lens of export growth at the extensive margin. By extensive margin growth is meant either trade in previously untraded products, or trade in existing products between new country pairs. The three chapters of this dissertation represent some of the first contributions to the literature in this area with an approach anchored in development policy concerns. They contribute to a clearer understanding of the links between firm heterogeneity, extensive margin growth, and development. In addition, they provide new insights into the importance of domestic regulations and institutions in terms of developing country trade performance. Chapter I shows that diversification in the product dimension is influenced by the costs of trade and market entry in the exporting country. Chapter II shows that similar factors also affect diversification in the geographical dimension. Chapter III takes the opposite approach, showing that product standards in importing countries can affect export diversification overseas. All three chapters use trade models incorporating heterogeneous firms to motivate the empirical work undertaken.
Résumé

Dans cette thèse, nous examinons les liens entre commerce international et développement économique à travers le concept de marge extensive du commerce. L’on considère comme participant à la croissance de cette marge, l’apparition de nouveaux flux de commerce, qu’ils concernent la commercialisation d’un bien jusqu’ici non échangé ou la commercialisation de produits déjà échangés mais entre des pays n’en faisant pas le commerce auparavant. Les trois chapitres de cette thèse constituent l’une des premières tentatives d’analyse de ces questions se concentrant sur la situation des pays en développement. Ils contribuent ainsi à une meilleure compréhension des liens entre hétérogénéité des entreprises, croissance de la marge extensive du commerce, et développement. En outre, ce travail permet de mettre plus clairement en évidence le rôle des réglementations et des institutions nationales dans la performance commerciale des pays en développement. Le premier chapitre montre que la diversification des exportations en termes du nombre de produits exportés est fonction des coûts de l’échange et des coûts d’établissement d’entreprise dans le pays exportateur. Le chapitre 2 élargit ces résultats en montrant que ces facteurs comptent, également, dans la diversification géographique des exportations. Enfin, le troisième chapitre aborde ces questions par une approche contraire, en étudiant l’impact des normes techniques d’un pays importateur sur la diversification des exportations de ses partenaires. Ces trois chapitres se basent sur des modèles théoriques prenant en compte l’hétérogénéité des entreprises qui justifient les travaux empiriques entrepris.
Development Perspectives on Trade Growth at the Extensive Margin

Acknowledgements

My sincere thanks to the members of my jury, Chad Bown, Joe Francois, and Bernard Hoekman, for having agreed to take part in my defense. I am particularly grateful to Joe and Bernard for accepting to act as rapporteurs.

Patrick Messerlin, my advisor, has been a constant source of support and encouragement, and I have always appreciated the time he has been willing to invest in providing advice and comments. The three years I spent at the Groupe d’Economie Mondiale in Paris were productive, rewarding, and very enjoyable. I am most grateful to Professor Messerlin for having made such a pleasant environment available to me, and for having integrated me into the research and teaching life of the GEM. A huge thank you is also due to Barbara Bender for her wonderful organizational skills, both during my time at the GEM and while finishing my dissertation and preparing my defense from the other side of the Atlantic.

It helped immeasurably to be able to share experiences (and pain) with other dissertation writers at the GEM, particularly Géraldine Kutas, and my former officemates Julien Hanoteau and Sébastien Miroudot. It was at the GEM with Daria Taglioni that I discovered the pleasures of firm heterogeneity, little knowing how much time I would end up spending with Pareto distributions, extensive margins, and all the rest of it.

Many people have been kind enough to read and provide comments on the various parts of my dissertation: Silja Baller, Allen Dennis, Simeon Djankov, Felix Eschenbach, Ana Margarida Fernandes, Matthias Helble, Daniel Lederman, Will Martin, Beata Smarzynska Javorcik, Christian Volpe Martincus, and John S. Wilson. I am very grateful for all of your efforts, and in particular for those of Bernard Hoekman who provided extremely helpful advice in relation to just about all of the work presented here (and a good deal more besides). Special thanks also to Allen Dennis, who allowed me to make use of our joint work in Chapter I.
There is no way that I could have finished this dissertation without the help, support, and (mostly good-natured) prodding of friends and family, even those—like my parents—far away. Robert and Marie-Françoise were incredibly tolerant of a new son-in-law who must have seemed unduly attached to his computer. I’m looking forward to being 100% present at family lunches and trips from now on!

Special thanks to Mitch O’Brien for his constant support, regular lunch company, and frequent raising of eyebrows. Tony and Sara McGarvey, and Anita and Seamus O’Hart rarely missed an opportunity to ask me when I would "really" be finished. If they hadn’t, I would probably still be touching up Chapter I.

Finally, I hardly know what to say to Mélanie.

For starters: IT’S FINISHED! REALLY!!

Thank you so much for all your love, understanding, and support. And for being patient at the right times, as well as impatient when I desperately needed it. I can’t wait to start having our evenings and weekends back!
General Introduction
General Introduction

This dissertation examines the links between international trade and economic development through the lens of export growth at the extensive margin. By extensive margin growth is meant either trade in previously untraded products, or trade in existing products between new country pairs.

Why focus on the extensive margin? One reason that it is relatively under-researched in general, having only emerged as a serious question in the literature with the observation of Haveman and Hummels (2004) that bilateral trade matrices tend to exhibit a very large number of zero entries: using data over the period 1970-1997, they find that in fully 80% of the observations, importers buy from less than 10% of the available suppliers. These zeros correspond to instances when a country either does not export a product at all, or exports it only to a subset of the available overseas markets. Thus, there would seem to be enormous scope for further extensive margin growth to take place in the future, provided that obstacles can be identified and, if necessary, dealt with via efficient policy interventions. To move in this direction, there is a clear need for research on the policy determinants of extensive margin growth.

More particularly, the development and development policy implications of extensive margin trade growth are still poorly understood. In a recent paper, Brenton and Newfarmer (2007) show that although it was intensive margin growth which dominated for developing countries over the period 1995-2004, extensive margin growth was far from negligible. They find the split to be approximately 80% intensive growth versus 20% extensive. Moreover, it is likely that their use of 5-digit SITC trade data tends to understate the level of extensive margin growth in the product dimension: the 3000 or so products identified by that nomenclature only represent around 60% of the products available at the 6-digit HS level, or around 30% of those available using the 8-digit Combined Nomenclature of the EU or its 10-digit US equivalent. Studies using different data—but
Development Perspectives on Trade Growth at the Extensive Margin

not focusing on developing countries—tend to report a stronger role for extensive margin growth: Evenett and Venables (2002) find that it accounts for 40% of recent trade growth, while Hummels and Klenow (2005) conclude that the extensive margin is responsible for 60% of the higher exports of big countries.

Set against the apparent importance of the extensive margin as a channel for trade growth, it is striking that there is relatively little work in this area from a development viewpoint. This dissertation is intended as a contribution towards bridging that gap. It does so by focusing on a range of possible policy determinants of extensive margin growth in developing countries. These policies are identified, and their effects predicted, using versions of the Melitz (2003) heterogeneous firms model. The first two chapters look at the role played by trade costs and domestic market entry costs in determining a country’s level of export diversification, which is conceptualized in terms of the extensive margin. Chapter I considers the new product dimension, while Chapter II focuses instead on geographical diversification. Finally, Chapter III examines the extensive margin impacts of the fixed costs of international trade, and in particular those associated with product standards in foreign markets.

In addition to the specific contribution made by each of the three substantive chapters, the dissertation as a whole can be seen as a contribution to the literature in two areas. First, as already noted, it represents an attempt to reinterpret and apply in a development context some recent advances in the theoretical and empirical literature on extensive margin growth. In addition, it can be seen as an extension of the existing literature on behind-the-border and regulatory barriers to trade: although tariffs figure in each chapter, the focus of the analysis is on other regulatory measures that usually lack an explicit trade motivation, such as domestic market entry barriers, and product standards.

The remainder of this introductory section briefly situates the dissertation with respect to these
two areas of the literature (sections 1 and 2). It then provides an overview of the dissertation’s structure, and the core results from each of the three substantive chapters (section 3).

**Extensive Margin Growth and Firm Heterogeneity: A Development Perspective**

Two central tendencies in the recent trade literature—both theoretical and empirical—represent important reference points for the work in this dissertation. The first is the increasing attention given to trade growth involving new products or new markets, as discussed above. The second is a renewed focus on the firm itself, and in particular the introduction of firm heterogeneity.\(^1\) Of course, it turns out that these two developments are in fact very closely linked: a trade model with firm heterogeneity and non-uniform market entry costs is broadly consistent with a trade matrix in which not all countries trade all products, and not all countries exporting a particular product do so to all available markets.

What are the potential implications of these changes in trade theory and empirics for development and development policy? The most obvious one has already been mentioned: the extensive margin represents a potentially important channel through which trade growth can take place, an important point for developing countries interested in promoting such growth as part of an overall development strategy. But there are at least two other closely related areas in which firm heterogeneity and the extensive margin could inform development policy: export diversification, and trade-induced productivity growth.

Export diversification and productivity growth are both core policy concerns for many development practitioners and policymakers. To see how they are related to firm heterogeneity and the extensive margin of trade, it is useful to briefly recall the intuition behind the Melitz (2003)\(^1\) See Bernard et al. (2007a) for a review.
heterogeneous firms model. The presentation here is deliberately schematic and informal, but is developed in full in the three substantive chapters of the dissertation.²

As in the previous generation of monopolistic competition trade models à la Krugman (1979), consumers in Melitz (2003) have love of variety preferences over a continuum of differentiated product varieties. Producers operate under increasing returns to scale, using a single, internationally immobile factor of production. They must pay one fixed cost to enter the domestic market, and an additional fixed cost for every overseas market entered. They also face per unit "iceberg" costs for international shipments. Zero profit conditions for the domestic market and each export market can be used to derive marginal cost cutoffs above which firms cannot profitably produce for a given market.

The major difference between this and the previous generation of models is that producers are not assumed to have a uniform level of productivity. Firms are heterogeneous, with their marginal cost drawn from a statistical distribution such as the Pareto. Melitz (2003) shows that under free entry, it is possible to derive a stable equilibrium for this type of model. Importantly, that equilibrium sees firms of different productivity levels coexisting in the same market. Only some of them—the most productive—export, due to the additional costs that must be incurred to do so. Thus, firms self-select into export markets based on their productivity draw and the costs they face.

These features of the Melitz (2003) framework fit well with some important stylized facts in the literature which cannot be easily obtained from a Krugman (1979) style model. First, there is strong evidence that only a small subset of active firms actually export. Bernard et al. (2007a) report that for the USA, only 4% of its 5.5 million active firms were exporters in the year 2000. Second, firm-level data generally suggest that exporters are larger and more productive than firms that produce for the domestic market only, and that these differences exist prior to export market entry.³ In the

² Alternative approaches to introducing heterogeneity are presented by Eaton and Kortum (2002) and Yeaple (2005). Bernard et al. (2007b) nest the Melitz approach in a two-factor model.
³ While there is also some empirical evidence of "learning-by-exporting" effects that would be consistent with a
US context, for instance, Bernard and Jensen (1999) find significant ex ante differences between exporters and non-exporters: firm size as measured by total shipments is 27%-54% higher in the former, while labor productivity is 7%-8% higher. Third, as already noted, not all countries export all products—or even all products that they produce—and when they do export a given product, it is usually only to a subset of the available overseas markets. Helpman et al. (Forthcoming) report that even aggregating across all products, only around half of the world bilateral trade matrix contains non-zero entries. Using firm-level data, Eaton et al. (2004) find that the modal French exporter deals with only one overseas market.

This third stylized fact is where the Melitz (2003) framework can be usefully linked to export diversification in both the product and geographical market dimensions. Such a link is important from a development policy point of view: export diversification has long been a concern for developing countries, but the literature has not provided a comprehensive framework within which to compare the impacts of different policies designed to promote it. Similarly, recent empirical work on the determinants of extensive margin trade growth has identified distance, tariffs, and market size among the relevant factors—but the link back to export diversification has not as yet been made.4

Moving beyond export diversification as such, there is also another sense in which heterogeneous firms models are of great potential importance to development policy. The productivity-based self-selection mechanism at the heart of Melitz (2003)-type models is suggestive of a direct link between trade costs and productivity growth.

Two things happen as trade costs fall. First, it becomes easier for home country firms to export—so more of them do, and this is expressed as export diversification (since each firm manu-

---

4This list is a consolidation of the results in Hummels and Klenow (2005), Feenstra and Kee (2006), and Debaere and Mostashari (2005).
factures a distinct variety). The second effect is the mirror image of the first: since it is now easier for foreign firms to export, import competition in the home market intensifies. This forces the least productive firms to exit, and so average productivity at home increases. The model therefore provides a precise, concrete mechanism by which trade is linked to growth—and it does so without relying on spillovers or learning-by-exporting effects that have found only patchy support empirically. It thereby paves the way for empirical work that can transcend the inconclusive "openness and growth" literature of the 1990s.\(^5\)

On a sectoral level, there is clear empirical support for the idea that lower trade costs lead to a resource reallocation away from less productive firms: see e.g., Pavcnik (2002) on the Chilean reforms of the 1970s and 1980s. Kee and Feenstra (2006) take a first step towards assessing the economic significance of the Melitz (2003) trade costs-productivity growth mechanism at a macroeconomic level. They analyze data on US imports from 44 countries over the period 1980-2000. Over these two decades, average export variety increased by around 4.6% per annum. They find that this variety growth was associated with an increase in total factor productivity of 4.5% in the exporting countries over the same period. Such a finding is of real economic significance from the point of view of development policy, given the vital role played by productivity growth in fueling sustained per capita income growth.

Against this background, the three substantive chapters of this dissertation can be seen as a contribution to the process of drawing out the implications of heterogeneous firms and extensive margin trade growth from a development policy perspective. Their focal point is the identification of policies that impact extensive margin growth, as well as a comparison of their relative impacts. While productivity growth as such is not the focus here, the results in all three chapters could be seen as laying the groundwork for the identification of trade-related policies that are likely to have significant growth promoting impacts. Importantly, the set of policies that is considered in the

---

\(^5\)For a critical review of this literature, see Rodrik and Rodriguez (1999).
remainder of this dissertation is not limited to border measures, such as tariffs, but is expanded to include a range of behind-the-border regulations, and even institutional factors, that can potentially influence the costs of domestic and foreign market entry.

**Trade Policy, Regulatory Barriers, and Institutions**

The three substantive chapters of this dissertation can also be seen as a contribution to the trade policy literature, and in particular to the part of it dealing with behind-the-border measures. The types of measures considered here include both trade-related regulations, such as customs procedures, and sectoral regulations, such as product standards. Consideration is also given to the regulation of market entry, which is an issue so central to the economy’s functioning that it could be argued to have almost an institutional character.

Regulatory measures and their potential trade effects have been moving steadily closer to the center of the international trade policy agenda over recent decades. In an environment of generally low and declining tariffs (at least for manufactured goods), it is natural that attention should shift towards other measures with potential trade consequences. This process was already well underway in the Uruguay Round, and is reflected in, for instance, the WTO Agreement on Technical Barriers to Trade and the Agreement on Sanitary and Phyto-Sanitary Measures. Regulatory barriers to trade are also at the core of the General Agreement on Trade in Services, since policies impacting services trade rarely take the form of tariffs.  

From a development point of view, it could be argued that regulatory barriers are particularly important because many developing countries benefit—at least in principle—from extensive preferential access to the main rich country markets. In theory, tariff barriers facing poor country exports to rich countries should be generally low, or even entirely absent for certain categories of products

---

6See Hoekman (2006) for a recent review of the literature on barriers to trade in services.
Development Perspectives on Trade Growth at the Extensive Margin

and countries. Thus, the EU’s "Everything But Arms" scheme provides for duty and quota free access for all exports originating in least developed countries. Together with the rules of origin that govern the ability to access these preferences, regulatory measures are therefore likely to be the main source of demand-side impediments to improved market access for developing countries.⁷

Two types of trade-related regulations are of particular interest in this context. On the one hand, the administrative procedures that must be followed when importing and exporting goods have a clear potential to influence trade flows. Simplification of such regulatory requirements is an important aspect of work programs in the area of trade facilitation.⁸ While Djankov et al. (2006) and Nordas et al. (2006) examine the trade impacts of the time delays at port and in clearing customs, there is as yet no empirical work on the direct impacts of these regulations themselves, or the costs they create. Chapters I and II of this dissertation contribute to the literature in this area by using new data from the "Trading Across Borders" section of the World Bank’s Doing Business database to examine this issue.

A second group of regulatory measures with potential trade impacts is product standards. Survey evidence from developing countries suggests that the compliance costs associated with rich country product standards can be substantial: Maskus et al. (2005) report that the fixed costs of compliance represent on average 4.7% of firm sales. However, the difficulty of assembling data on product standards means that only a relatively small number of empirical papers have examined their trade impacts, in particular from a developing country viewpoint.⁹ Chapter III of the present dissertation adds to this literature by using new data to examine the impact of CEN European standards on developing country exports. These standards have not previously been examined in the

---

⁷Two caveats are required here. First, preference coverage is far from universal across all developing countries and products. Significant barriers often remain in sectors such as agriculture and clothing, in particular for non-LDC developing countries. Second, the issue of South-South market access is also important from a development point of view. Tariffs remain much more central to this question than to the question of the South’s access to Northern markets.

⁸See Wilson et al. (2005) for an overview.

⁹Examples include Swann et al. (1996), Moenius (2004), Fontagné et al. (2005), and Disdier et al. (Forthcoming).
literature. Chapter III also evaluates the potential for international harmonization of product standards to ease the cost burden associated with product standards. While bilateral harmonization has been examined by Swann et al. (1996) and Moenius (2004), and regional harmonization has been considered by Chen and Mattoo (2004), there is no previous work on the impact of international harmonization as such.

Recent research has started going even further behind the border than the two groups of regulatory measures just discussed. It is increasingly argued that national institutions can also impact international trade: see, for example, De Groot et al. (2004), and Levchenko (2007). Indeed, Francois and Manchin (2007) find evidence that institutions and infrastructure—another regulatory issue—explain more of the observed variation in North-South trade flows than do tariffs.

The emergence of this trade and institutions literature can be seen as an extension of the recent interest in institutions from a growth perspective: see, for example, Acemoglu et al. (2001). As is the case for growth, it is important to "unbundle" the institutions-trade linkage (cf. Acemoglu and Johnson, 2005). This means identifying precise mechanisms by which institutional features can impact trade. Thus far, the focus has been on contract enforcement. Examples of work in that area include Anderson and Marcouiller (2002), Nunn (2007), and Ranjan and Lee (2007). All of these papers find that contract enforceability plays a significant role in shaping bilateral trade patterns.

Chapters I and II of this dissertation can be seen as building on the above work by focusing on the trade effects of a different but related institution: the regulation of market entry (see Djankov et al., 2002). While the distinction between regulations and institutions is by no means a clear one, it could be argued that the regulation of market entry plays such a fundamental role in structuring economic interactions that it should be placed in the latter category. Moreover, it is intimately related to the two core aspects of institutions, as that term is used by, for instance, North (1990): property rights, and contract enforcement. If the firm is seen as a nexus for contracts, then regulatory
measures that increase the cost of setting up a firm make those contracts more costly to conclude. Regulating entry also impacts on property rights more generally, by making it more costly to create a separate legal entity—the firm—which can hold property in its own name.

Market entry regulations have as yet received only very limited attention from a trade point of view. Helpman et al. (Forthcoming) include data on entry regulation in the first stage of their sample-selection gravity model, and find a significant impact on export propensity. The examination conducted in the first two chapters of this dissertation therefore represents one of the first attempts to analyze the trade impacts of market entry regulations, in particular in the development context.

**Structure of the Dissertation and Outline of Main Results**

The dissertation consists of three substantive chapters. All three are primarily empirical, but use adaptations of the Melitz (2003) theoretical framework to motivate the empirical investigation. The first two chapters are concerned with the determinants of export diversification, characterized as trade growth at the extensive margin. The first focuses on the "new products" margin, while the second deals with the "new markets" margin. Finally, the third chapter examines the impact of foreign product standards and international harmonization on the "new products" margin.

**Chapter I: "Trade Costs, Barriers to Entry, and Export Diversification in Developing Countries"**

The first chapter of this dissertation analyzes the policy determinants of export diversification in the product dimension. It develops a comprehensive theoretical framework that can be used to analyze and compare the impacts of various policies that have been put forward as possible instruments for promoting export diversification. Examples include infant industry protection, trade facilitation,
Development Perspectives on Trade Growth at the Extensive Margin

tariff preferences in importing countries, and promoting entrepreneurship and "discovery" (i.e., domestic market entry). The theoretical model is analyzed numerically to produce predictions as to the expected signs and relative magnitudes of export diversification elasticities with respect to these policies.

To examine the model’s predictions empirically, measures of export diversification are assembled for 118 developing countries in the year 2005. They are based on 8-digit EU import data which distinguish amongst over 10,000 individual products. Diversification is proxied using a count of the number of 8 digit product lines in each 2 digit sector for which non-zero exports are recorded. This measure is found to be closely correlated with common alternatives, such as the Herfindahl Index, and the measure of relative product variety due to Feenstra (1994).

Econometric results using a Poisson fixed effects model provide generally robust support for the model’s predictions. In the preferred specification, 1% reductions in the costs of exporting and international transport (proxied by distance) are associated with export diversification gains of 0.4% and 0.3% respectively.\footnote{Export cost data are sourced from the World Bank’s Doing Business dataset, and include all official costs of document preparation, inland transport, administrative fees, and port/customs charges.} Lower domestic market entry costs have a weaker effect, with an elasticity of -0.1.\footnote{These data are also sourced from Doing Business. They include all official costs of starting up and formally operating a local limited liability company with general industrial or commercial activities.} These results are robust to numerous specification checks, including use of narrower country samples, alternative diversification measures, and added control variables to account in greater detail for the exporting country’s economic structure. They continue to hold when the baseline model is re-estimated using colonization dummies (British, French, Spanish, Portuguese, and Russian) as instruments for export and entry costs, and the exporting country’s own tariffs.

These results show that policies to reduce export costs (i.e., trade facilitation), reduce international transport costs, and (to a lesser extent) promote domestic market entry are likely to have
a positive impact on export diversification. Equally striking, however, is what they do not show: tariffs in both the exporting and importing countries are found to have a statistically insignificant impact on export diversification. While there could be many reasons for this outcome—in particular the large number of tariff lines with zero rates under EU preferential schemes—such results should nonetheless give cause for reflection. Based on the data examined here, it is not at all clear that tariffs represent the binding constraint on export diversification in developing countries. Should this be confirmed by future work, it would be an important factor for policymakers to take account of both when designing policies to promote diversification, and when seeking reforms of trading partners’ policies with that same end in mind.

Chapter II: "Exporting to New Markets: An Empirical Investigation"

The second chapter complements the first by examining the determinants of an alternative dimension of export diversification: entry of new geographical markets. A multiple market version of the theoretical model from Chapter I is used to form predictions which are then tested using both firm- and country-level data.

Preliminary empirical analysis is undertaken using firm-level data from 98 countries, sourced from the World Bank’s Enterprise Surveys. Estimates of total factor productivity and labor productivity are related to information on the number of foreign markets to which each firm exports. Results provide some support for the productivity-based self-selection process at the heart of the model, although data limitations mean that it is necessary to be cautious in drawing strong conclusions from the firm-level analysis.

The second stage of the analysis proceeds with country-level data. Standard 6-digit international trade data are used to construct count measures of geographical export diversification for each of 177 developing countries, covering 97 2-digit Harmonized System sectors. A Poisson model with
fixed effects is used to relate these measures to market size, international transport costs (distance),
importer tariffs, export costs, and domestic market entry costs. In all cases except tariffs, robust
evidence is found of impacts that are both economically and statistically significant. These results
hold up to numerous specification checks, including first differencing, and use of data on colonial
history and official language as instruments for export and market entry costs.

In policy terms, the implications of the results in Chapter II are quite similar to those from
Chapter I. In particular, these results highlight the importance of policies designed to lower the
costs of trade and domestic market entry as potential means of promoting geographical export di-
versification. It is also interesting to compare the results from the two chapters in order to get a first
idea of the relative sensitivity of geographical and product market diversification with respect to
such policies. At this stage, the data suggest that the number of export markets is noticeably more
responsive to changes in these variables than is the number of products exported. While the reasons
behind this effect are not yet clear, the parallel with recent empirical work on the evolution of the
new product and new market margins is clear: Brenton and Newfarmer (2007) find that expansion
of the range of export markets has been a more important force behind developing country export
growth than has expansion in the number of products exported. The contrast between these conclu-
sions and some development policy discourse—in which it is product, not market, diversification
that is of most importance—is striking.

Chapter III: "Export Variety and International Harmonization of Product Standards"

Whereas the first two chapters examine the policy determinants of export growth at the extensive
margin, Chapter III uses observed differences in extensive margin growth to draw inferences about
the impact of foreign product standards and harmonization on the fixed costs of trade. Using a new
World Bank database of CEN European product standards in the textiles, clothing, and footwear

12Data sources and definitions are as in Chapter I.
sectors, it is shown empirically that international harmonization is associated with increased export
diversity in partner countries. International harmonization is proxied by the percentage of total EU
standards in a given sector which are equivalent to ISO standards. To deal with the potential for
endogeneity bias, sector size and geographical dispersion data are used as instruments for the total
number of standards and degree of international harmonization. The chapter’s core results prove to
be highly robust to this and other specification checks.

These empirical results are shown to be consistent with a theoretical framework in which prod-
uct standards impose fixed adaptation costs on exporting firms. Bringing such standards into line
with international norms means that a larger market can be accessed upon payment of a single fixed
cost. This effectively reduces the total fixed costs facing exporting firms, which in turn promotes
export growth at the extensive margin. Numerical simulations show that this reduction in cost
multiplicity dominates even when harmonization leads to a moderate increase in compliance costs.

The policy implications of these results are potentially far-reaching. They provide general
support for the WTO approach to technical regulations: international standards should be applied
whenever feasible and appropriate, in order to limit negative effects on third country exporters.
They also highlight the WTO’s limitations, however. The WTO legal regime is limited in direct
application to standards that are mandatory as a matter of law. It does not apply to standards like
CEN’s, which although not strictly mandatory, are nonetheless a commercial imperative in many
instances.

References

Economy, 113(5), 949-995.

Acemoglu, Daron, Simon Johnson, and James A. Robinson, 2001, "The Colonial Origins of
Development Perspectives on Trade Growth at the Extensive Margin


Disdier, Anne-Célia, Lionel Fontagné, and Mondher Mimouni, Forthcoming, "The Impact of Regulations on Agricultural Trade: Evidence from SPS and TBT Agreements", *American Journal*
Development Perspectives on Trade Growth at the Extensive Margin

_of Agricultural Economics._


SHEPHERD, Benjamin - Thèse IEP de Paris - 2008


Chapter I

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

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

**JEL codes:** F12; F13; O24.

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

A Introduction

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

---

13 This Chapter represents joint work with Allen Dennis (World Bank). I am grateful to my co-author both for a very fruitful collaboration, and for his agreement to include this paper in my dissertation. We thank the following for helpful discussions and suggestions: Simeon Djankov, Matthias Helble, Bernard Hoekman, Daniel Lederman, Will Martin, Beata Smarzynska Javorcik, and Christian Volpe Martincus. We also benefitted from comments by seminar audiences at the World Bank and the 2007 European Trade Study Group meetings.
Development Perspectives on Trade Growth at the Extensive Margin

Development objective. Indeed, Imbs and Wacziarg (2003) show that a higher level of per capita income tends to be associated with a more diverse production structure, at least until relatively late in the development process when specialization effects begin to dominate. Funke and Ruhwedel (2001) find that export diversification is positively related to per capita GDP and TFP growth in OECD countries. Similarly, Hausmann et al. (2007) report that exports of particular sets of products are associated with higher national income levels.

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

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

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

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

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

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

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

B Data and Stylized Facts

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

1 Measuring Export Diversification

At its simplest, export diversification means broadening the range of products that a country exports. We therefore equate it with export growth at the extensive margin, and draw on the recent literature in that area for guidance as to the ways in which diversification can be measured.\textsuperscript{14} As a starting point, we use a direct measure, namely a count of the number of products that a country exports. In practice, this approach is not as simple as it seems because individual "products" identified in the trade data usually map in reality to a number of distinct varieties. The most detailed trade

\textsuperscript{14}We limit consideration to growth in the number of products exported, and do not examine the number of markets to which they are exported. This is an important question in its own right, and we expect that the methods used in this paper could be adapted to build on the insights developed in this area by Evenett and Venables (2002).
data available on a worldwide basis are at the 6-digit level of the Harmonized System (HS), and distinguish amongst 5,000 or so different products. Casual empiricism suggests, however, that counts based on 6-digit data are likely to understate the true level of export diversity due to aggregation effects.

In this paper, we improve on the level of detail provided by the 6-digit HS classification by using an 8-digit classification that provides roughly twice as much product-level detail: 10,753 distinct product lines.\footnote{The data are classified using the EU’s Combined Nomenclature (CN), which is based on the Harmonized System (HS) but contains additional subdivisions at the 8-digit level.} We extract these data from a freely available Eurostat database covering exports from and imports into the European Union.\footnote{See \url{http://fd.comext eurostat.cec.eu.int/xtweb/}. To our knowledge, these data have not previously been used in product variety work. However, Fontagné et al. (1998), Henry de Frahan and Vancauteren (2006), and Manchin (2006) have exploited them in other contexts.} Our strategy is to use EU data on imports from developing countries to construct new "mirror" measures of export diversification in those same countries. Although it is true that we thereby measure export diversification vis-à-vis the EU and not the world as a whole, we believe that our measures remain highly relevant since the EU is one of the most important outlets for developing country exports. Moreover, this approach offers two concrete advantages over the use of cross-country export data at the 6-digit level. First, the aggregation problem is reduced (although not eliminated) due to the much greater level of detail in which products are defined and flows recorded. Second, import data from the EU are likely to be more reliable than the corresponding export data from developing countries due in part to stronger governance and customs agency capacity.

Taking 2005 as our base year, we start with a dataset of 470,035 observations across 246 countries and customs areas (including EU members), and 10,753 distinct products.\footnote{Although these data are available for a number of years, limited availability of trade and entry cost data means that we must work in a purely cross-sectional context.} In this paper, we focus only on the developing country component of that dataset, namely countries that are neither

---

SHEPHERD, Benjamin - Thèse IEP de Paris - 2008 36
members of the EU-25 nor the OECD. (We return to this definition in the context of robustness checks below.) Our first measure of export diversification, \( \text{lines} \), is a count of the number of 8-digit product lines in which a given country exported to the EU-15 in 2005. It has one observation per country. To provide greater detail, we also construct \( \text{lines}\_\text{cn2} \) following the same pattern as for \( \text{lines} \), but with counts by 2-digit sector rather than aggregated to the country level. \( \text{lines}\_\text{cn2} \) therefore has 97 observations per country (the number of 2-digit Chapters in the CN classification).

Given that the CN 8-digit classification scheme is inconsistent in the level of detail (i.e., individual "products") it accords each sector, we will need to take care to correct for this when using \( \text{lines}\_\text{cn2} \) as an indicator of export diversification.

In Table 1, we provide a list of the countries included in our sample divided up according to the quintiles of \( \text{lines} \). On average, they exported 1,138 8-digit product lines to the EU in 2005. However, the range is extremely wide: from 9 lines (Palau) to 8053 (China), out of a possible maximum of 10,753. In broad terms, the country rankings accord with the sensible prior that larger, more developed countries tend to have more diversified export bundles (see Hummels and Klenow, 2005). Thus, we find China, India, and Brazil at the top of the table, while Palau, Micronesia, and the Comoros are at the opposite end.

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

\[
\Lambda = \frac{\sum_{i \in J^H} p_i^{H} q_i^{H}}{\sum_{i \in J^W} p_i^{W} q_i^{W}}
\]  

(1.1)
The numerator in this measure is therefore the total value of world exports in product lines exported by country $H$, and the denominator is the total value of world exports across all products. When we calculate $\Lambda$, it turns out to be very strongly correlated ($\rho = 0.95$) with lines.

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

$$hh_{index} = \sum_{i=1}^{J} \left( \frac{p_i q_i}{\sum_{j=1}^{I} p_j q_j} \right)^2 \quad (1.2)$$

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

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

2 Measuring the Costs of Trade and Domestic Market Entry

Full details of the remaining elements in our dataset are presented in Table 2. Our bilateral applied tariff data from the ITC-CEPII Market Access Map (MAcMap) database take full account of preference schemes, and are available for 2004 only. We aggregate them from the HS 6-digit
Development Perspectives on Trade Growth at the Extensive Margin

level to the 2-digit level using the reference group methodology due to Laborde et al. (2007). This aggregation scheme minimizes endogeneity bias by calculating weights based on import shares in groups of similar countries, rather than in each country individually. We also draw historical and geographical data from CEPII (Mayer and Zignago, 2006), while our macroeconomic data are extracted from the World Development Indicators.

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

To measure domestic market entry costs, we use the "Starting a Business" component of Doing Business (see Djankov et al., 2002). This source includes indicators on the costs, time, and number of procedures required for an entrepreneur to start-up and formally operate a local limited liability company with general industrial or commercial activities. This includes legally required pre-registration, registration, and post-registration activities. Only official costs are considered, based on information gathered from the company law, commercial code, and specific regulation and fee schedules. Together, we refer to these as the costs of domestic market entry ("Entry Cost").

As far as we are aware, this is the most comprehensive source of cross-country information on business start-up costs, and has previously been used in the trade context by Helpman et al. (Forthcoming): they find that higher entry costs are negatively associated with the probability that two countries engage in trade.

3 Non-Parametric Evidence

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

Before entering into the detail of our approach, it is useful to have a first look at the data using
non-parametric methods. Like Imbs and Wacziarg (2003), we use the locally weighted scatterplot smoother (Lowess) for this purpose.\textsuperscript{18} We use \textit{lines} as the dependent variable in all cases, resulting in a single observation per country. It is important to be clear that this analysis is intended to be exploratory only, since it is conducted one variable at a time, and does not control for the impact of other factors.

Results are presented in Figure 1. The first two plots verify the analysis of Hummels and Klenow (2005) to the effect that larger (GDP), richer (GDP per capita) countries tend to have a more diverse export bundle.\textsuperscript{19} We do not observe the inverse-U relationship between \textit{lines} and GDP per capita that would be expected on the basis of the results in Imbs and Wacziarg (2003), undoubtedly because our sample includes only developing countries, which may not yet have reached the relatively high level of income (approximately that of Ireland) at which specialization effects come to dominate. The remainder of Figure 1 discloses generally negative relationships between export variety and entry costs (upper right corner), great circle distance from the EU (as a proxy for transport costs), and export costs (lower right). In the case of distance, however, the relationship is a complex one involving a mid-sample "hump".

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

\textsuperscript{18}The Lowess smoother runs separate regressions of \textit{lines} on trade or entry costs at each data point, using 80\% of the total sample. For each regression, observations are downweighted according to their distance from the central data point around which the regression takes place. All calculations are performed in Stata 9.2SE.

\textsuperscript{19}The outlier in the upper right corner of the first plot is China. Results do not change if it is dropped.
C Theoretical Model

We now show that a heterogeneous firms trade model similar to Melitz (2003) can explain the stylized facts presented in the previous section.20 The major substantive difference between our approach and Melitz’s is that we allow for potentially asymmetric countries, trade costs, and market entry costs. We also derive some explicit results in relation to the extensive margin of trade, which we interpret in terms of export diversification. While the basic mechanics of the model remain unchanged, the introduction of these asymmetries prevents us from obtaining an analytical solution (Baldwin and Harrigan, 2007). We therefore proceed by numerical simulation using essentially the same parameter set as in Bernard et al. (2007b). We consider unilateral changes in the fixed and variable costs of exporting and importing, as well as in the costs of domestic market entry. We then consider a reciprocal change in the variable costs of trade, which we assimilate with a variation in international transport costs. Simulation results not only accord with the preliminary empirical evidence presented in section 2, but also suggest an ordering of export diversification elasticities with respect to the different types of costs we have identified. We test both sets of predictions in section 4.

1 Consumption Block

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

---

20Our presentation of the model draws heavily on Helpman et al. (2004), and Baldwin and Harrigan (2007).
specialization is excluded from consideration.

Identical consumers in both markets maximize the two-tier utility function (I.3). The first tier is Cobb-Douglas, with expenditure share $\beta_s$ on the differentiated product sector $s$ and $(1 - \sum_{s=1}^{S} \beta_s)$ on the homogeneous sector. The second tier for each differentiated goods sector is Dixit-Stiglitz with elasticity of substitution $\sigma_s$ across the set $V_s$ of available varieties.

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

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

\[
x [p(v)] = d^H_s [p(v)]^{-\sigma_s}
\] (I.4)

2 Production Block

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

\[
p(v) = \frac{\sigma_s}{\sigma_s - 1} c
\] (I.5)

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

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

\[
\pi_{d,s}^H = d_s^H \left( \frac{\sigma_s}{\sigma_s - 1} - c \right)^{1 - \sigma_s} - d_s^H \left( \frac{\sigma_s}{\sigma_s - 1} c \right)^{-\sigma_s} - c - \frac{f_d^{H}}{\text{fixed costs}}
\]

\[
\equiv d_s^H \left( \frac{\sigma_s}{\sigma_s - 1} c \right)^{1 - \sigma_s} - f_d^{H} \quad (I.6)
\]

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

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

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

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

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

$$
\pi_{x,s}^{HR} = d_s^R \left( \frac{\sigma_s}{\sigma_s - 1} \frac{R_H}{R_s} c \right)^{1-\sigma} - d_s^R \left( \frac{\sigma_s}{\sigma_s - 1} \frac{R_H}{R_s} c \right)^{-\sigma} \frac{R_H}{R_s} c - f_x^H \text{ fixed costs}
$$

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

$$
c_{x,s}^{HR} = \left( \frac{\sigma_s f_x^H}{d_s^R} \right)^{\frac{1}{1-\sigma}} \left( \frac{\sigma_s - 1}{\left(\frac{R_H}{R_s}\right)^{\sigma}} \right)
$$

The corresponding condition for firms based in $R$ is:

$$
c_{x,s}^{RH} = \left( \frac{\sigma_s f_x^R}{d_s^H} \right)^{\frac{1}{1-\sigma}} \left( \frac{\sigma_s - 1}{\left(\frac{R_H}{R_s}\right)^{\sigma}} \right)
$$

Firm heterogeneity is introduced via marginal production costs $c$, which follow a Pareto distribution with support $[0, \bar{c}]$.\textsuperscript{21} For simplicity, this distribution is taken to be identical in both countries and across all sectors. The most productive firm in any sector therefore produces with zero marginal cost, while the least productive firm produces with the highest possible marginal cost $\bar{c}$.\textsuperscript{22} This setup implies a cumulative distribution function and probability density function of the form:

$$
G(c) = \left( \frac{c}{\bar{c}} \right)^k
$$

\textsuperscript{21} It is immaterial whether firm heterogeneity is introduced in terms of marginal costs, as here, or in terms of labor productivity as in Melitz (2003). See Baldwin and Harrigan (2007).

\textsuperscript{22} The possibility of zero marginal cost production can be excluded by assuming random draws from a Pareto distribution truncated at a point $\zeta > 0$ (Helpman et al., 2007). This would make no material difference to our results.
and

\[ g(c) \equiv \frac{dG(c)}{dc} = \frac{k c^{k-1}}{c^k} \]  

(I.13)

where \( k > \sigma - 1 \) is a "shape" parameter that in this context is effectively an index of firm heterogeneity: a higher value of \( k \) is associated with less pronounced heterogeneity.\(^{23}\)

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

\[
\int_0^{c_{d,s}^H} \left( \frac{d_s^H}{\sigma_s} \left( \frac{\sigma_s}{\sigma_s - 1} \right)^{1-\sigma_s} - f_d^H \right) dG(c) \\
\text{Expected profits from domestic sales}
\]

\[
+ \int_0^{c_{d,s}^H} \left( \frac{d_s^R}{\sigma_s} \left( \frac{\sigma_s}{\sigma_s - 1} \right)^{1-\sigma_s} - f_d^R \right) dG(c) = f_e \tag{I.14}
\]

The corresponding set of \( S \) conditions for \( R \) is simply:

\[
\int_0^{c_{d,s}^R} \left( \frac{d_s^R}{\sigma_s} \left( \frac{\sigma_s}{\sigma_s - 1} \right)^{1-\sigma_s} - f_d^R \right) dG(c) \\
\text{Expected profits from domestic sales}
\]

\[
+ \int_0^{c_{d,s}^R} \left( \frac{d_s^H}{\sigma_s} \left( \frac{\sigma_s}{\sigma_s - 1} \right)^{1-\sigma_s} - f_d^R \right) dG(c) = f_e \tag{I.15}
\]

\(^{23}\)As usual, we assume \( k > \sigma - 1 \) for two reasons. First, since \( \sigma > 1 \), this assumption is sufficient to assure the existence of second moments for the marginal cost distribution. Second, it ensures the convergence of a number of integrals used below, for instance in the free entry conditions.
3 Characterizing Export Diversification and Simulating the Model

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

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

$$d_s^H = \frac{\beta_s E^H}{\int_{V^H_s} p(v)^{1-\sigma_s} dv}$$

$$= \frac{\beta_s E^H}{n_s^H \int_0^{c_{d,s}^H} (\frac{\sigma_s}{\sigma_s - 1} - c)^{1-\sigma_s} dG(c) + n_s^R \int_0^{c_{x,s}^R} (\frac{\sigma_s}{\sigma_s - 1} - c_{x,s}^{HR})^{1-\sigma_s} dG(c)}$$

(1.16)
Development Perspectives on Trade Growth at the Extensive Margin

\[
d_s^R = \frac{\beta_s E^R}{\int_{v \in V_s^R} p(v)^{1-\sigma_s} dv}
\]

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

\[
n_{x,s}^H = \frac{G(c_{x,s}^H)}{G(c_{d,s}^H)} n_s^H
\]

\[
n_{x,s}^R = \frac{G(c_{x,s}^R)}{G(c_{d,s}^R)} n_s^H
\]

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

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

Figure 2 presents summary results for this set of simulations in terms of approximate elasticities at each point in the simulation space. Our choice of simulation scenarios is driven by the desire to provide a solid motivation for the empirical work that we undertake in the next section, so we concentrate on parameter changes that can easily be mapped to cross-country differences in observed policies. Thus, the first scenario we consider (solid line) is a unilateral increase in H’s import-specific trade costs ($\tau^H_{s,s}$), which could be brought about, for example, by an increase in H’s own tariff rate in sector $s$. By insulating H’s market to some degree from international competition, the overall mass of active firms increases and so, therefore, does the number of exporters—thus, diversification in H increases. The dashed line shows the impact of this policy change from R’s

---

24 We initially use: $\sigma = 3.8, k = 3.4, L^H = L^R = 1000, \beta = 0.5, \tau = 0.2, f^e = 2$, and $f^H_d = f^R_d = 0.1$. To conform with the stylized facts that trade is costly and that only some firms export, we set $f^H_x = f^R_x = 0.15$ and $\tau^H_R = \tau^R_H = 1.1$.

25 We simulate over the following ranges: $1.1 \leq \tau^H_H \leq 1.2, 0.15 \leq f^H_s \leq 0.25$, and $0.10 \leq f^H_d \leq 0.145$. Full results, including Maple worksheets, are available on request.

26 For example, for the simulation based on changes to $f^H_d$, we calculate $\frac{\Delta n^H_{x,s}}{n^H_{x,s}} \Delta f^H_d$ at each simulation point.
perspective (i.e., a unilateral change in a partner country’s tariffs): predictably, a smaller effective market for firms in that country leads to a decrease in diversification there.

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

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

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

D Empirical Model and Results

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

$$f (\text{lines}_2 \mid X_{es}) = \frac{\exp (-\mu_{es}) \mu_{es}^{\text{lines}_2} e^{-\mu_{es}}}{\text{lines}_2!}$$ (I.20)

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

$$\mu_{es} = \delta_s \exp \left[ \beta_1 \ln (\text{entry}_e) + \beta_2 \ln (\text{export}_e) + \beta_3 \ln (1 + t_{es}^{EU}) + \beta_4 \ln (\text{dist}_e) + \beta_5 \ln (1 + t_{es}^{own}) + \beta_6 \ln (\text{gdp}_hs2_e) + \beta_7 \ln (\text{gdppce}_e) \right]$$ (I.21)

We use $\text{entry}_e$ to refer to restrictions on entry in a given exporting country ($f_d^H$), which we

---

27We estimate the model at the sectoral (2-digit) level, since three of the variables of interest—EU and own tariffs, and sectoral expenditures—vary at that level. However, we are conscious that the other variables of interest vary only at the country level, and so we adjust all standard errors for clustering by exporting country.

28We expect the sector fixed effects to account for influences such as trade-related measures applied on an MFN basis within our country sample (e.g., product standards), as well as the different numbers of 8-digit product lines included in each 2-digit sector within the CN classification.
proxy using the *Doing Business* entry costs data referred to above. Export-specific trade costs \((\tau_s^{RH} \text{ and } f_x^H)\) are proxied using MAcMap bilaterally disaggregated applied tariff data for the EU \((t_{es}^{EU})\) and *Doing Business* data on the costs of exporting \((export_e)\). While ad valorem tariffs impose only variable trade costs, the types of costs captured by *Doing Business* include both fixed and variable components: for instance, export documentation needs to be agreed and drafted prior to any export activity taking place (fixed cost), but then needs to be copied and slightly adapted for each shipment (variable cost). We use the distance between the exporting country and Europe \((dist_e)\) as an indicator of the extent of symmetric trade costs \((\tau_s^{RH} = \tau_s^{HR})\) such as international transport charges, while sectoral expenditures \((\beta_s E^H)\) and technology \((c\text{ and } k)\) are proxied by (respectively) GDP multiplied by a sectoral (2-digit) import share, and GDP per capita.\(^{29}\) We can therefore summarize the core contentions of the paper as derived in the previous section in terms of restrictions on the parameter space of the conditional mean function: \(\beta_5 > 0; \beta_1, \beta_2, \beta_3, \beta_4 < 0\) and \(|\beta_4| > |\beta_2|, |\beta_3| > |\beta_5| >> |\beta_1|\).

The model can be estimated using standard maximum likelihood techniques.\(^{30}\) The Poisson estimator has at least three desirable properties that make it particularly useful in this context. First, it has identical first order conditions to those obtained by running weighted non-linear least squares on (I.21) with \(lines - cn2\) in place of \(\mu_{es}\) (Gourieroux et al., 1984). Thus, the various \(\beta\) coefficients can be interpreted as elasticities, but the model does not suffer from the usual limitation of log-log models in relation to zero observations: these can simply be included in the estimation sample as usual. Santos Silva and Tenreyro (2006) exploit this property as a means of dealing with zero trade flows in the gravity model context.

\(^{29}\)To construct our tariff measures and sectoral expenditure proxies, we use the reference group approach of Laborde et al. (Forthcoming), which relies on observed imports for a group of similar countries so as to avoid the endogeneity inherent in using a simple import weighted average. Little turns on this choice, however, since our results do not change substantially if aggregate GDP or simple average tariffs are used instead. (Results available on request.)

\(^{30}\)See generally: Cameron and Trivedi (2001), Wooldridge (1997), and Winkelmann (2000).
weak assumptions. It is not even necessary that the dependent variable in fact follow a Poisson distribution, so long as the conditional mean function is correctly specified (Gourieroux et al., 1984; Wooldridge, 1997). Other count data models, such as the negative binomial (Klinger and Lederman, 2004 and 2006), do not have this property and require the data to follow a particular distribution in order to ensure consistency. Third, Poisson is one of relatively few non-linear maximum likelihood models in which standard, unconditional fixed effects do not lead to bias and inconsistency due to the "incidental parameters problem" (Greene, 2004). This is an important practical advantage in relation to the alternative empirical strategy of analyzing the extensive margin via a standard gravity model using a Heckman sample selection estimator (Helpman et al., Forthcoming), a probit model (Debaere and Mostashari, 2005) or the Tobit estimator.31 While it may be possible to avoid this problem by using a random effects estimator (Baldwin and Harrigan, 2007), the price that must be paid is a strong assumption as to the distribution of country- or sector-specific unobserved heterogeneity.

Poisson estimates of our baseline model are presented in column 1 of Table 3.32 Overall, the model provides a close fit to the data, with $R^2$ equal to 0.92.33 All coefficients carry the expected signs: entry costs, export costs, distance, and EU tariffs are negative, while own tariffs, GDP, and GDP per capita are positive. The coefficients on entry costs, export costs, distance, and GDP are all statistically significant at the 1% level, but the remaining coefficients are not significant at the 10% level. These results are broadly consistent with previous work examining the diversification impacts of GDP, per capita income, and trade costs, such as Hummels and Klenow (2005) and Feenstra and

---

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

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

33 We follow Wooldridge (1997) in using $R^2 = 1 - ESS/TSS$ as a convenient summary measure of fit. We prefer it to the more common (for count data) pseudo-$R^2$ measures due to its ease of interpretation. We are conscious, however, that it should be used with caution since the Poisson model does not contain a residual as such.
Kee (2006). Similarly, our finding that lower domestic market entry costs are associated with greater export diversification is consistent with the results of Helpman et al. (Forthcoming), who find a negative and statistically significant coefficient on market entry costs in the first stage of a sample selection gravity model.34

In terms of the magnitudes of the estimated coefficients on the trade and entry cost variables, Table 3 indicates that EU tariffs have the largest elasticity in absolute value terms (-0.6), followed by distance and own tariffs (-0.4 and 0.4 respectively), then export costs (-0.3), and finally entry costs (-0.1). These magnitudes are clearly of economic significance, with (for example) a 10% decrease in export costs being associated with a 3% increase in export diversification. While this ordering of effects should clearly be interpreted with caution due to the relative imprecision with which the two tariff effects are estimated, it nonetheless moves in generally the same direction as the predictions from our theoretical model. The only major difference with the expected ordering based on simulation results is that EU tariffs have a larger elasticity (in absolute value terms) than does distance, but a Wald test of equality between the two coefficients does not reject the null (prob. = 0.89). However, when we test the hypothesis that the coefficients on entry costs, export costs, EU tariffs, and distance are equal, we find that the null is rejected at the 5% level (prob. = 0.03). This supports our earlier contention that the impact of entry costs on diversification is likely to be considerably weaker than for trade costs.

34Our results in respect of domestic market entry costs might appear at first glance to contrast with those of Klinger and Lederman (2004, 2006). However, those authors use a different dependent variable, namely a count of the number of 3-digit SITC categories in which a country exports for the first time over a given period. Thus, they view diversification through the lens of export "discoveries", rather than as an increase in the proportion of nationally produced varieties that make it to the international market. Their explanatory variable set is also different, including a wider set of entry cost data such as labor market rigidity and contract enforceability indicators, but they do not consider trade costs.
1 Instrumental Variables Strategy

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

To deal with these possibilities, we adopt two complementary approaches. First, we use five year lags of GDP and per capita GDP, since these measures should be genuinely exogenous with respect to the current level of export diversification. Next, we use an instrumental variables strategy to identify exogenous variation in our measures of entry and export costs (including own tariffs).³⁵ To estimate the Poisson model with instrumental variables, we follow the procedure set out in Wooldridge (2002, pp.663-665), which is analogous to two-stage least squares. The first stage consists of OLS regressions in which the potentially endogenous variables—entry costs, export costs, and own tariffs—are used sequentially as the dependent variable, while the exogenous variables from the baseline model along with the instruments are used as the independent variables. The residuals from these regressions are retained, and entered into the baseline Poisson formulation (I.21) as additional regressors (i.e., retaining the potentially endogenous measures of trade and entry costs). Conditional on the use of appropriate instruments, inclusion of the first stage residuals removes any endogeneity bias. A simple test of the joint significance of the residual terms can be used as an indication of the endogeneity of the variables being instrumented for: rejection of the null hypothesis indicates that there is a serious endogeneity problem.

As usual, the most difficult aspect of this approach lies in identifying variables that satisfy the

³⁵We treat distance and EU tariffs as exogenously determined.
twin conditions for instrument validity: strong correlation with the potentially endogenous trade and entry costs, but orthogonality to current export diversification. Our strategy exploits variation in geography and colonial history to identify exogenous movements in the variables of interest. The idea is that a country’s economic and social institutions should be correlated with distance from the equator (Hall and Jones, 1999), as well as with the marks left by colonial rule on legal regimes and institutions (Djankov et al., 2002). Thus we expect both instruments to be correlated with the currently observed levels of trade and entry costs, while being exogenous to current export diversification. In our first stage regressions (Table 4), we therefore include distance from the equator and binary dummies for colonization by Great Britain, France, Spain, Portugal, the Netherlands, and Russia in addition to the exogenous variables from the baseline model. The first stage $R^2$ is 0.4 for entry costs, and 0.3 for export costs and own tariffs, while F-tests of the null hypothesis that all instruments jointly equal zero are 5.47***, 1.88*, and 4.65*** respectively. While our instruments do a fair job of explaining the observed pattern of entry and trade costs, it would obviously be desirable in future work to try and identify stronger instruments. This is particularly true in the case of export costs, for which the F-test only rejects the null hypothesis at the 10% level. Nonetheless, we continue with geography and colonial history here in order to get a first impression of the extent to which instrumental variables estimates might have the potential to alter our baseline results.

Results from the second stage regression are in column 2 of Table 3, with the coefficients on the three sets of first stage residuals suppressed for brevity. The coefficients on entry costs, export costs, distance, and GDP all carry the expected sign, and (with the exception of entry costs) are statistically significant at the 5% or 1% levels. EU and own tariffs, as well as GDP per capita, carry

---

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

37 As a robustness check, we also estimate the IV model using two-stage least squares, since that methodology allows us to test the validity of our overidentifying restrictions whereas Poisson does not. Estimated coefficients are similar to Poisson, and the overidentification test does not reject the null ($\chi^2_4 = 5.14$, prob. = 0.27). These results suggests that our instruments are genuinely exogenous, and satisfy the second-stage exclusion restriction.
unexpected coefficient signs, but are not statistically significant at the 10% level. While we are cautious in interpreting changes in coefficient magnitude given the relative imprecision with which some of them are estimated, we observe that the elasticities of export diversification with respect to export costs and distance increase in absolute value terms, but the reverse is true for entry costs. Finally, we test the joint significance of the three first-stage residuals using a standard Wald test, and find that the null hypothesis that they equal zero cannot be rejected at the 10% level (prob. = 0.24). This suggests that the empirical impact of endogeneity would appear to be relatively minor in this particular case.

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

In addition to examining potential endogeneity problems, it is also important to ensure that our results are robust in other respects. We now perform checks along four additional dimensions: alternative measures of export diversification and entry/trade costs, additional dependent variables, changes in the country sample, and application of alternative estimators.


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

The results discussed above use \textit{lines\_cn2} as the dependent variable, i.e. a count of the number of 8-digit product lines exported to the EU in each 2-digit sector. Table 5 shows that our results are not sensitive to this choice. The first two columns use a more narrowly defined dependent variable, in which 8-digit export flows are only considered to be non-zero if they are greater than $100,000$ or $1,000,000$ respectively. The third column uses the original definition, but excludes agriculture (chapters 1-24) from consideration. Although there are some small changes in coefficient estimates as a result of these alterations to the dependent variable, Table 5 shows that the substance of our results is unchanged: entry and export costs, as well as distance, exert a negative and statistically significant impact on diversification, while the positive impact of own tariffs and the negative impact of EU tariffs are statistically insignificant.

The final two columns of Table 5 take completely different approaches to measuring diversification. Column 4 uses the relative variety measure of Feenstra and Kee (2006), which is closely related to the measures used by Hummels and Klenow (2005) and Broda and Weinstein (2006). Column 5, on the other hand, replaces \textit{lines\_cn2} with the Herfindahl index of concentration that is frequently used as an inverse indicator of export diversification in policy circles (e.g., UNCTAD, 2006; and Amin Gutiérrez de Piñeres and Ferrantino, 1997). Both measures are calculated at the 2-digit sector level using 8 digit data, applying the formulae set out in section 2. As can be seen from the table, both alternative measures of diversification produce qualitatively identical results to those in our baseline formulation. The estimated signs are the same in all cases except EU tariffs, while entry costs, export costs, distance, and GDP remain statistically significant at the 5% or 1% level.

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

3 Additional Independent Variables

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

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

As a final check, we examine the robustness of our results to the use of alternative estimators. As suggested above, Poisson is consistent under relatively weak assumptions, and is not adversely affected by the inclusion of unconditional fixed effects. However, the possibility remains of gains in efficiency through the use of alternative estimators that have been employed in the literature. Columns 4-6 of Table 6 present results using standard OLS, Tobit, and the negative binomial model.\footnote{Tobit estimates should be interpreted with caution since we include unconditional fixed effects. For the negative binomial estimates, we employ the conditional fixed effects approach of Hausman et al. (1984). Robust standard errors are obtained using interval regression for Tobit, while bootstrapping (200 replications) is used for the negative binomial.} Once again, there is no substantive change from our baseline results in column 1 of Table 3. Entry costs, export costs, and distance all impact export diversification negatively and significantly (at the 1% level), while own tariffs have a positive impact which is statistically significant at the 5% level under OLS and at the 1% level using the negative binomial. EU tariffs enter with an unexpected positive sign, but are not statistically significant in any of the three models.

E Conclusions and Policy Implications

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

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

Subject to that caveat, our results can be used to sketch out some preliminary implications for development policy. On the one hand, policies aimed at reducing trade costs related specifically to exporting, as well as international transport costs, would appear to be promising avenues for developing countries aiming to promote export diversification. Trade facilitation measures, in the broad sense in which Wilson et al. (2005) use that term, would therefore be attractive. Similarly, measures aimed at reducing the costs of domestic market entry are also promising, although our results suggest that their impact will be less strong than for export costs or international transport costs. The unifying factor behind both sets of measures is that they envisage diversification not as a result of governments "picking winners" through industrial policy, but as a natural outcome of winners picking themselves through an intensification of the Schumpeterian process at the heart of the Melitz (2003) model.

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

**F References**


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


### G Tables

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

<table>
<thead>
<tr>
<th>Range</th>
<th>Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-118</td>
<td>Belize, Bhutan, Burundi, Central African Republic, Chad, Comoros, Djibouti, Dominica, Eritrea, Grenada, Guinea-Bissau, Haiti, Kiribati, Lesotho, Marshall Islands, Micronesia, Palau, Rwanda, Samoa, Serbia, Solomon Islands, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, São Tomé and Príncipe, Tajikistan, Tonga, Vanuatu</td>
</tr>
<tr>
<td>131-281</td>
<td>Afghanistan, Antigua and Barbuda, Armenia, Benin, Botswana, Burkina Faso, Equatorial Guinea, Fiji, Gambia, Guinea, Guyana, Iraq, Kyrgyz Republic, Malawi, Maldives, Mali, Mauritania, Mozambique, Nicaragua, Niger, Papua New Guinea, Seychelles, Sierra Leone, Sudan, Swaziland, Togo, Zambia</td>
</tr>
<tr>
<td>296-685</td>
<td>Angola, Azerbaijan, Bolivia, Cambodia, Cameroon, Cape Verde, Congo, Rep., El Salvador, Ethiopia, Gabon, Georgia, Guatemala, Honduras, Jamaica, Kazakhstan, Lao PDR, Mongolia, Namibia, Panama, Paraguay, Suriname, Tanzania, Trinidad and Tobago, Uganda, Uzbekistan, Yemen, Zimbabwe</td>
</tr>
<tr>
<td>746-1815</td>
<td>Albania, Algeria, Bangladesh, Belarus, Bosnia and Herzegovina, Colombia, Costa Rica, Côte d’Ivoire, Dominican Republic, Ecuador, Ghana, Jordan, Kenya, Kuwait, Lebanon, Macedonia, FYR, Madagascar, Mauritius, Moldova, Nepal, Nigeria, Oman, Peru, Senegal, Syria, Uruguay, Venezuela</td>
</tr>
<tr>
<td>1876-8053</td>
<td>Argentina, Brazil, Bulgaria, Chile, China, Croatia, Egypt, Hong Kong, China, India, Indonesia, Iran, Israel, Malaysia, Morocco, Pakistan, Philippines, Romania, Russia, Saudi Arabia, Singapore, South Africa, Sri Lanka, Thailand, Tunisia, Ukraine, United Arab Emirates, Vietnam</td>
</tr>
</tbody>
</table>
### Table 2: Data and sources.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Units (Yr.)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag. % GDP</td>
<td>Agriculture value added as % GDP.</td>
<td>% (2005)</td>
<td>WDI</td>
</tr>
<tr>
<td>Col_*</td>
<td>Equal to one if a country was colonized by * (UK, France, Spain, Portugal, the Netherlands, or Russia).</td>
<td>NA</td>
<td>CEPII</td>
</tr>
<tr>
<td>Distance</td>
<td>Average of the great circle distances between the main cities of the exporting country and Germany, weighted by population shares.</td>
<td>Km</td>
<td>CEPII</td>
</tr>
<tr>
<td>Entry Cost</td>
<td>Official cost of starting up and formally operating an industrial or commercial business in the exporting country.</td>
<td>USD (2006)</td>
<td>Doing Business</td>
</tr>
<tr>
<td>Export Cost</td>
<td>Official fees levied on a 20 foot container leaving the exporting country. Includes document preparation costs, administrative fees for customs clearance and technical control, terminal handling charges, and inland transit.</td>
<td>USD (2006)</td>
<td>Doing Business</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross domestic product.</td>
<td>USD (2005)</td>
<td>WDI</td>
</tr>
<tr>
<td>GDP Defl.</td>
<td>GDP deflator.</td>
<td>% (2005)</td>
<td>WDI</td>
</tr>
<tr>
<td>GDPPC</td>
<td>Per capita GDP.</td>
<td>USD (2005)</td>
<td>WDI</td>
</tr>
<tr>
<td>Lat.</td>
<td>Latitude of the main city in the exporting country (absolute value).</td>
<td>Deg.</td>
<td>CEPII</td>
</tr>
<tr>
<td>Lines</td>
<td>Number of 8-digit product lines in which a country has strictly positive exports to the EU.</td>
<td>NA (2005)</td>
<td>Eurostat</td>
</tr>
<tr>
<td>Lines_CN2</td>
<td>Number of 8-digit product lines in a 2-digit sector for which a country has strictly positive exports to the EU.</td>
<td>NA (2005)</td>
<td>Eurostat</td>
</tr>
<tr>
<td>Manuf. % GDP</td>
<td>Manufacturing value added as % GDP.</td>
<td>% (2005)</td>
<td>WDI</td>
</tr>
<tr>
<td>Real Int. Rate</td>
<td>Real interest rate.</td>
<td>% (2005)</td>
<td>WDI</td>
</tr>
<tr>
<td>Tariffs</td>
<td>Average applied ad valorem tariff by HS2 sector. Aggregated from 6-digit data using the reference group methodology of Laborde et al. (Forthcoming).</td>
<td>% (2005)</td>
<td>MAcMap</td>
</tr>
</tbody>
</table>
Table 3: Baseline estimation results.

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

Estimation is by Poisson, with dependent variable lines_cn2. Independent variables are in logarithms. All models include fixed effects by 2-digit sector. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%). Specification IV uses colonization dummies and distance from the equator as instruments for entry and export costs, and own tariffs. First stage $R^2$ is 0.4, 0.3, and 0.3 respectively.
### Table 4: First stage instrumental variables regressions.

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

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

2. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%).
Table 5: Estimation results using alternative measures of export diversification.

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

1. Estimation in columns 4-5 is by OLS with dependent variables lambda_cn2 and hh_index_cn2. Estimation in columns 1-3 is by Poisson, with dependent variable lines_cn2. Columns 1 and 2 only count export flows greater than $100k and $1m respectively, while column 3 excludes agricultural products (HS chapters 1-24). All models include fixed effects by 2-digit sector.

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

3. Robust standard errors, adjusted for clustering by exporter, are in square brackets. Statistical significance is indicated by * (10%), ** (5%), and *** (1%).
**Table 6: Additional robustness checks.**

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

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

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

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

Figure 1: Non-parametric regression results.

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

2. All results were obtained by Lowess regression with bandwidth=0.8.
Figure 2: Simulation results, expressed as approximate elasticities.

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

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

Exporting to New Markets:
An Empirical Investigation
II Exporting to New Markets: An Empirical Investigation

**Abstract:** Export costs, tariffs, international transport costs, and the costs of domestic market entry are all shown to be important determinants of geographical export diversification in a sample of 177 developing countries. These effects are generally robust to alternative specifications, as well as to estimation using instrumental variables. In addition, firm level data from the World Bank’s *Enterprise Surveys* are used to provide suggestive evidence of a positive link between firm productivity and the number of export markets entered. The empirical findings at both the firm- and country-level are consistent with the predictions from a heterogeneous firms trade model.

**JEL codes:** F13; F15.

**Keywords:** International trade; Heterogeneous firms; Economic geography.

**A Introduction**

Both China and Djibouti exported some electrical goods (HS Chapter 85) in 2004.\(^\text{39}\) The total values involved were of course very different: mirror data from the UN Comtrade database record a total export value of $206bn for China, compared with the more modest sum of $353,000 for Djibouti. One factor behind this enormous difference might be that while China and Djibouti both export the same goods to the same markets, China simply exports a greater quantity. The gravity model literature (e.g., Anderson and Van Wincoop, 2003 and 2004) provides a wealth of detail on how to identify the factors behind this, such as domestic market size, distance to international markets, tariffs, and non-tariff barriers (NTBs). Recent work on quality margins in trade could help elucidate the causes behind another possible dynamic, namely that China might export goods with a higher unit value than those exported by Djibouti. See, for example, Baldwin and Harrigan

\(^{39}\)I am very grateful to Felix Eschenbach (IMF) and Ana Margarida Fernandes (World Bank) for their many helpful comments on this Chapter. Jens Arnold (OECD) kindly supplied Stata code that was most useful in conducting the firm-level estimations.
Recent work on the extensive margin of international trade suggests that two additional factors might also be playing an important role. On the one hand, China exports a much wider range of product varieties than does Djibouti. Limiting consideration to the EU as destination, the dataset assembled by Dennis and Shepherd (2007) indicates that China exported 633 8 digit products in 2004, whereas Djibouti only exported 14. A number of factors might be contributing to such an outcome: e.g., market size and level of development (Hummels and Klenow, 2005), or trade costs and barriers to domestic market entry (Dennis and Shepherd, 2007).

In addition to all of these possibilities, there is also the striking fact that in 2004, China exported its HS Chapter 85 products to a total of 141 countries, while Djibouti only dealt with 20. Surprisingly, the analytical literature provides relatively little guidance as to why this should be the case. What are the economic mechanisms that could explain why exporting firms in Djibouti were able to link up with 20 external markets, but no more? And why could Chinese firms easily exceed the 20 market level—to arrive at a total over seven times higher? From a policy point of view, what measures could the government in Djibouti contemplate in order to help its firms access more overseas markets for products they already manufacture and sell elsewhere?

In this paper, I show that lower trade costs and restrictions on domestic market entry, as well as higher firm productivity and larger market size, are all associated with exports to a larger number of foreign markets. To do this, I use a combination of aggregate and firm-level data. By way of preliminary examination, I estimate firm-level productivity using developing country data from the World Bank’s Enterprise Surveys. I then relate these productivity measures to the number of export markets entered by individual firms, and find some support for the hypothesis that more productive firms enter a larger number of markets. This approach is complemented by a more detailed analysis using standard 6 digit trade data, from which I construct sectoral counts of the number of foreign
markets to which each of 177 developing countries exported in 2004. Empirical results confirm that lower costs of trade and market entry, as well as a larger domestic market and a higher level of domestic productivity, are all associated with greater export market diversification.

What is the economic intuition underlying these results? This paper develops a simple heterogeneous firms model of trade, in which firms self-select into export markets based on productivity. Allowing for country-specific fixed and variable trade costs makes it possible to produce an ordering of "market accessibility": the most accessible market has the lowest productivity cutoff, while the least accessible market has the highest productivity cutoff. If firms enter markets in increasing order of accessibility, then more productive firms will tend to serve a larger set of export markets. The largest set of export markets served by any single firm—which is coterminous with the set of export markets served by the country as a whole—is determined by the productivity level of the most productive firm in the economy. This makes it possible to derive an explicit expression for the number of export markets entered. Comparative statics go in the same direction as the empirical results presented here, suggesting that heterogeneity of firms and trade costs could be a powerful explanation for observed differences in geographical export diversification across countries.

From a policy point of view, these are important findings in an area that is still poorly understood due to the relative lack of empirical work on the determinants of export market diversification. The recent re-emergence of interest in export diversification and the extensive margin of trade has tended to focus on the product dimension: e.g., Hummels and Klenow (2005), Broda and Weinstein (2006), Hausmann et al. (2007), and Dennis and Shepherd (2007). While the sample selection gravity model approach of Helpman et al. (Forthcoming) is more general, it does not allow for particular factors to impact product and market diversification differently. A priori, it seems plausible that such differences could be significant, due to the different cost levels involved in developing a new product compared with marketing an existing product in a new country.
Recent work suggests that geographical diversification of exports is an important mechanism by which developing countries can become more integrated in the world trading system.\(^{40}\) Evenett and Venables (2002) report that around 1/3 of developing country export growth over the period 1970-1997 was due to the export of "old" goods to new markets. Using a different dataset and methodology, Brenton and Newfarmer (2007) suggest that the proportion was more like 18% for the period 1995-2004. While there is thus a legitimate discussion to be had as to the exact breakdown of export growth into its various components—such as quality upgrading, new products, new markets, and intensification—there is little doubt that geographical diversification is sufficiently important to developing countries’ trade prospects to merit further empirical investigation.

It is striking that very little empirical work focuses on the possible determinants of geographical export diversification.\(^{41}\) Eaton et al. (2005) use a database of French firms to analyze the determinants of export behavior. They find evidence showing that bigger firms (i.e., those with higher levels of sales in France) tend to export to a larger number of foreign markets. However, they do not directly examine the impact of policy-related factors such as trade costs.

Evenett and Venables (2002) examine the export growth of 23 developing countries to 93 foreign markets over the period 1970-1997. They work at the 3 digit level of the SITC classification. Conducting logit regressions separately for each 3 digit product, they find that the probability of exporting to a given destination is generally decreasing in distance, but increasing in market size. Exporting to proximate markets is found to be a significant predictor of geographical diversification, which Evenett and Venables (2002) argue could be consistent with learning effects. They also find some evidence that a common border and common language increase the probability of observing trade for a given country dyad.

\(^{40}\) To the extent that policy interest in product diversification is motivated by the desire to diversify price or terms of trade risk, then market diversification can also be seen as a potential source of welfare gains under certain circumstances: e.g., if demand shocks are not correlated (or are negatively correlated) across export markets.

\(^{41}\) Brenton and Newfarmer (2007) is largely a descriptive paper. The authors do not conduct an empirical analysis of the determinants of geographical diversification.
The present paper builds on and extends these contributions in a number of ways. First, it develops a comprehensive theoretical approach to geographical diversification via a heterogeneous firms framework à la Melitz (2003). The model, which is similar to the monopolistic competition case of the model in Eaton et al. (2005), can explain the important stylized fact that many countries export goods to only a limited subset of all available trading partners. It suggests that the reason for this lies in a combination of firm productivity and trade costs.

Second, the model’s predictions are tested using a combination of firm- and country-level data. Like Evenett and Venables (2002), this paper uses standard international trade data, though at the 6 digit HS level rather than the 3 digit SITC level. In addition, it uses data from the World Bank’s Enterprise Survey database to provide preliminary evidence supporting the model’s main firm-level prediction.

Finally, the analysis of geographical diversification conducted here uses a direct measure—a count of the number of export markets entered—rather than an indirect approach focusing on the probability of observing exports for a given country pair. This makes both estimation and interpretation considerably easier.

The paper proceeds as follows. In the next section, I develop a theoretical framework that provides the hypotheses to be tested in the remainder of the paper. Section 3 then conducts a preliminary empirical investigation using firm-level data. In section 4, more detailed empirical tests are conducted using country-level trade data and a range of policy variables. Section 5 concludes, and discusses some directions for future research in this area.

B Theoretical Motivation

In this section, I develop the theoretical framework behind the hypotheses that will be tested empirically in the remainder of the paper. To do this, I use a heterogeneous firms model of inter-
national trade with many countries and heterogeneous fixed and variable trade costs. My approach
can be seen as an adaptation of Helpman et al. (2004, Forthcoming), and is also very close to the
model in Eaton et al. (2005) under the special case of monopolistic competition. The main con-
tribution here is to develop a framework that links firm-level productivity to export market entry
decisions, and then to show how this setup can be used to make predictions that can be tested using
both firm-level productivity data and country-level trade flow data.

The intuition behind this approach is straightforward in terms of the standard heterogeneous
firms framework. By allowing for multiple countries and arbitrary trade costs, the self-selection
mechanism at the core of the model in Melitz (2003) becomes more complex: firms sort into those
that produce for the domestic market only, and those that produce for each potential export market.
They do so according to the zero profit productivity cutoff associated with each bilateral route.
These cutoffs can therefore be seen as an index of market accessibility. Assuming that all firms
enter markets sequentially—i.e., in decreasing order of accessibility—the number of markets entered
by each firm should be increasing in productivity.

It follows from the above that the set of export markets observed on a country-level is identical
to the largest firm-level set. It is therefore the productivity level of the most productive firm in each
sector—along with trade costs and market size—that determines the expected number of export links
in bilateral trade data.

1 Consumption Block

The world consists of a home country $H$ along with $(M - 1)$ potential export markets, for a
total of $M$ countries. To simplify, $H$ only exports to $0 < m < (M - 1)$ of those markets. This
is consistent with the stylized fact that countries rarely export to all other countries in the world,
thereby leading to a large number of zeros in the bilateral trade matrix. For instance, Haveman and
Hummels (2004) report that in 80% of the observations in their dataset of world trade at the 4 and 5 digit SITC level, importers buy from less than 10% of the available suppliers. I treat \( m \) as an endogenous variable, depending on the market entry decisions of individual firms.

Labor is the only factor of production. Country \( i \) has an endowment of \( L_i \) units. The economy consists of two productive sectors. The first sector produces a freely traded homogeneous good under constant returns to scale with one unit of labor required for one unit of output. Wages are therefore equal to one in equilibrium. The second sector produces a continuum of differentiated goods under increasing returns to scale and costly trade. Corner solutions (i.e., absolute specialization in either sector) are excluded from consideration.

Identical consumers in all \( M \) markets maximize the two-tier utility function (II.1). The first tier is Cobb-Douglas, with expenditure share \( \beta \) on the differentiated sector and \( (1 - \beta) \) on the homogenous one. The second tier for the differentiated goods sector is Dixit-Stiglitz with elasticity of substitution \( \sigma \) across the set \( V \) of potentially available varieties.

\[
U = q^{1-\beta} \left[ \int_{v \in V} x(v)^{1-\frac{1}{\sigma}} dv \right]^{\frac{\beta}{\frac{1}{\sigma} - 1}} \tag{II.1}
\]

I use \( d_i = \frac{\beta E_i}{\int_{v \in V_i} p(v)^{1-\sigma} dv} \) to denote country \( i \)’s demand shift parameter, where \( V_i \) is the set of varieties available in that country, and \( E_i = L_i \) is total expenditure by its consumers. It is therefore possible to write the demand function \( x \) facing a given differentiated goods producer as:

\[
x \left[ p(v) \right] = d_i \left[ p(v) \right]^{-\sigma} \tag{II.2}
\]
2 Production Block

As usual under Dixit-Stiglitz, firms engage in constant markup pricing such that \( p(v) = \frac{\sigma}{\sigma - 1} c \). They face a fixed cost of domestic market entry \( f_d \), which is identical in all countries. A typical firm therefore has domestic market profits equal to:

\[
\pi^i_d = d^i \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} \frac{\sigma - 1 - c}{\sigma - 1} c - d^i \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} \frac{\sigma - 1}{\sigma - 1} c\right) - f_d \quad (\text{II.3})
\]

Setting profits equal to zero and solving for marginal cost gives a maximum \( c_d^i \) above which a firm cannot profitably produce for the domestic market:

\[
c_d^i = \left( \frac{\sigma f_d}{d^i} \right)^{\frac{1}{\sigma}} \left( \frac{\sigma - 1}{\sigma} \right) \quad (\text{II.4})
\]

Goods that are produced in country \( i \) and shipped to country \( j \) incur two additional costs. The first is "iceberg" type variable costs \( \tau^{ij} > 1 \), which make the effective marginal cost of exported products \( c\tau^{ij} \). These costs could be related to international transport charges, or trade policy measures such as tariffs. They vary by country pair and direction of trade (\( \tau^{ij} \neq \tau^{ji} \)). In addition, firms must also meet a fixed cost of export market entry \( f_x^{ij} \) in order to set up a distribution network or adapt goods to meet foreign preferences or product standards. These costs are the sum of a fixed (non-market specific) component, and a market specific component: \( f_x^{ij} = f_x + s^{ij} \). The additional firm profits due to exporting from country \( i \) to country \( j \) are therefore:

\[
\pi_x^{ij} = d^j \left( \frac{\sigma}{\sigma - 1} \tau^{ij} c \right)^{1-\sigma} \frac{\sigma - 1}{\sigma - 1} c\right) - d^j \left( \frac{\sigma}{\sigma - 1} \tau^{ij} c \right)^{1-\sigma} \frac{\sigma - 1}{\sigma - 1} c\right) - f_x^{ij} \quad (\text{II.5})
\]

In the same way as for the domestic market cost cutoff \( c_d^i \), it is possible to derive a maximum
cost level \( c_{x}^{ij} \) above which a firm cannot profitably export from \( i \) to \( j \):

\[
c_{x}^{ij} = \left( \frac{\sigma}{\sigma - 1} \right) \frac{1}{\sigma} \frac{\sigma - 1}{\sigma t^{ij}}
\]

(II.6)

As is common in this literature, I introduce firm heterogeneity through the marginal cost channel. I assume that marginal cost \( c \) follows a truncated Pareto distribution with support \([c_{low}, c_{high}]\), probability density function \( g (c) = \frac{k c^{\sigma - 1}}{c_{high}^{\sigma} - c_{low}^{\sigma}} \) and cumulative distribution function \( G (c) = \frac{c^{\sigma} - c_{low}^{\sigma}}{c_{high}^{\sigma} - c_{low}^{\sigma}}. \)

After paying a fixed cost \( f_{e} \) to enter the marginal cost "lottery", firms decide which markets (domestic and export) to enter based on the bilateral cutoff conditions above. Firms in \( i \) with \( c > c_{d}^{i} \) exit immediately, and do not produce at all. Those with \( c_{x}^{ij} < c \leq c_{d}^{i} \) for all \( j \) produce for the domestic market only, while firms with \( c \leq c_{x}^{ij} \) for at least one other country \( j \) start exporting. The set of markets to which they export is determined by the set of cutoffs for which \( c_{x}^{ij} \) holds.

To close the production block of the model, I assume that free entry equates expected operating profits from domestic and export market sales with the lottery cost \( f_{e} \).

\[
\int_{c_{low}}^{c_{d}^{i}} \left( \frac{d_{i}}{\sigma} \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} - f_{d} \right) dG (c) + \sum_{j \neq i} \int_{c_{low}}^{c_{d}^{ij}} \left( \frac{d_{j}}{\sigma} \left( \frac{\sigma}{\sigma - 1} t^{ij} c \right)^{1-\sigma} - f_{x}^{ij} \right) dG (c) = f_{e}
\]

(II.7)

The above system of \( M \) free entry conditions, \( M(M - 1) \) export cutoff conditions, and \( M \) domestic cutoff conditions can be solved for the total of \( M \) demand shifters, \( M \) domestic market cutoffs, and \( M(M - 1) \) export market cutoffs. Equilibrium sees firms self-selecting through the mechanism described above into those which serve a particular market, and those which do not.

For full details of the equilibrium properties of this class of models, see Melitz (2003).

---

42 This formulation is very close to the one used by Helpman et al. (2007). As Baldwin and Harrigan (2007) point out, it is immaterial whether a probability distribution is assumed for productivity (as in Melitz, 2003) or for marginal costs (i.e., inverse productivity) as here.
3 Characterizing Geographical Diversification

The above framework is useful for describing the pattern of bilateral trade relations amongst a set of countries. For instance, Helpman et al. (Forthcoming) use it as the basis for developing a gravity model that accounts for the presence of zeros in the bilateral trade matrix. It does not, however, provide direct information on the total number of markets entered by firms in the exporting country. To obtain such information, it is useful to frame the export market cutoff conditions in a slightly different way. Consider an ordering of the set of export cutoff conditions for country \( i \) in terms of "market accessibility", i.e. from easiest to enter (highest) to hardest to enter (lowest).\(^\text{43}\) Thus, if \( i \) has a trading relationship with \( m \) markets, then the ordering means \( c_x^1 \geq c_x^2 \geq \ldots \geq c_x^m \geq c_{low} \), where the final inequality follows from the fact that export market entry continues as long as it is possible for at least one firm to do so profitably. Since \( c_{low} \) represents the lower bound on marginal costs in \( i \), it is not possible for any export market cutoff associated with non-zero exports to have \( c_{ij} < c_{low} \). If firms enter markets in decreasing order of accessibility, then it follows that:

**Result 1:** Conditional on trade costs and market size, the number of foreign markets to which a firm sells is increasing in that firm’s productivity.

The market accessibility ordering makes it possible to re-formulate the set of export cutoff conditions in terms of the number of markets entered. Assuming that firms enter markets in increasing order of restrictiveness, the condition for entering one export market only (\( c_x^1 \)) is simply equal to the condition for the most accessible market, i.e. \( c_x^1 = c_x^1 \). The condition for entering two markets takes a more complex form. Recalling that \( f_{ij} = f_x + s_{ij} \), the additional profits from entering

---

\(^{43}\)While this approach seems reasonable a priori, it must be recognized that the work of Eaton et al. (2005) on French data provides only mixed support for this mechanism. Those authors find that only 27% of French firms satisfy such a strict ordering in their choice of export destinations.
Development Perspectives on Trade Growth at the Extensive Margin

exactly two markets can be written as:

\[
\begin{align*}
\pi_x^2 &= \frac{d^1}{\sigma} \left( \frac{\sigma}{\sigma - 1} \tau i^1 \right)^{1-\sigma} - \frac{d^2}{\sigma} \left( \frac{\sigma}{\sigma - 1} \tau i^2 \right)^{1-\sigma} - f_x^2 \\
&= \frac{1}{\sigma} \left( \frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \left[ d^1 \left( \tau i^1 \right)^{1-\sigma} + d^2 \left( \tau i^2 \right)^{1-\sigma} \right] - 2f_x - \sum_{j=1}^{2} s^{ijj} \quad \text{(II.8)}
\end{align*}
\]

Setting the above expression equal to zero and solving for the two market cutoff condition \(c_x^2\) gives:

\[
\begin{align*}
\frac{1}{\sigma} \left( \frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \left[ d^1 \left( \tau i^1 \right)^{1-\sigma} + d^2 \left( \tau i^2 \right)^{1-\sigma} \right] - 2f_x - \sum_{j=1}^{2} s^{ijj} &= 0 \\
c_x^2 &= \left( \frac{\sigma}{\sigma - 1} \right)^{\frac{1}{1-\sigma}} \left( 2f_x + \sum_{j=1}^{2} s^{ijj} \right) \quad \text{(II.9)}
\end{align*}
\]

Repeating this process for all remaining markets with non-zero exports produces \(m\) such conditions. The last one, for entering exactly \(m\) export markets takes the following form:

\[
\begin{align*}
c_x^m &= \left( \frac{\sigma}{\sigma - 1} \right)^{\frac{1}{1-\sigma}} \left( \frac{\sigma - 1}{\sigma} \right) \geq c_{low} \quad \text{(II.10)}
\end{align*}
\]

where the final inequality follows once again from the fact that if an export market is entered by at least one firm, the cutoff for that market can be no lower than the lower bound on marginal cost. To simplify the analysis, I assume that the export cutoff for \(m\) markets is precisely equal to the lower bound of the marginal cost distribution, i.e. \(c_x^m = c_{low}\). This is a reasonable approximation if the number of countries is large and the country specific market entry costs \(s^{ijj}\) are sufficiently close together at all relevant points.

Using the above condition makes it possible to derive an explicit expression for the maximum
number of export markets entered by any firm in the economy, in terms of model parameters and the set of demand shifters:

\[
m = \frac{1}{f_x} \left[ \left( \frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \frac{c_{low}^{1-\sigma}}{\sigma} \sum_{j=1}^{m} d^j (\tau^{ij})^{1-\sigma} - \sum_{j=1}^{m} s^{ij} \right] \tag{II.11}
\]

For a given country, an increase in \( m \) equates to greater export market diversification. To examine the determinants of geographical diversification, some simple comparative statics are useful.\(^{44}\) It simplifies things greatly to hold all demand shifters constant. One way of looking at this is as a sort of small country assumption: small changes in the parameters facing one country do not have any impact on the global equilibrium represented by the full set of demand shifters. This approach yields the following results:\(^{45}\)

\[
\frac{dm}{d f_x \, dd^i = 0 \, \forall j} = -\frac{m}{f_x} < 0 \tag{II.12}
\]

\[
\frac{dm}{d s^{ij} \, dd^i = 0 \, \forall j} = -\frac{1}{f_x} < 0 \tag{II.13}
\]

\[
\frac{dm}{d \tau^{ij} \, dd^i = 0 \, \forall j} = \frac{1}{f_x} \left( \frac{\sigma}{\sigma - 1} \right)^{1-\sigma} \frac{c_{low}^{1-\sigma}}{\sigma} (1 - \sigma) \sum_{j=1}^{m} d^j (\tau^{ij})^{-\sigma} < 0 \tag{II.14}
\]

\[
\frac{dm}{d c_{low} \, dd^i = 0 \, \forall j} = \frac{1}{f_x} \left( \frac{\sigma}{\sigma - 1} \right)^{1-\sigma} (1 - \sigma) \frac{c_{low}^{1-\sigma}}{\sigma} \sum_{j=1}^{m} d^j (\tau^{ij})^{1-\sigma} < 0 \tag{II.15}
\]

These comparative statics accord well with the sensible prior that countries with more productive firms \((d c_{low} < 0)\) which face lower fixed and variable trade costs \((d f_x, d s^{ij}, d \tau^{ij} < 0)\) should

---

\(^{44}\)The following analysis is a simplification, since it treats the number of markets \( m \) as a continuous, rather than discrete, variable. It would be possible to deal with this more rigorously by using a continuum of countries, but that is an approach that is left for future work at this stage.

\(^{45}\)For simplicity, the derivative with respect to iceberg trade costs effectively assumes \(d \tau^{ij} = d \tau^{ik}\) for all countries \( \{j, k\} \neq i \). Thus, it is akin to multilateral rather than unilateral liberalization.
export to a wider variety of destinations. They are restated as Result 2, which is tested together with Result 1 in the remainder of this paper.

**Result 2:** Lower trade costs and higher productivity are associated with exports to a larger number of foreign markets.

### C Preliminary Analysis: Firm Level Empirics

Over the last decade or so, the empirical literature on international trade has expanded in an important new direction. Whereas traditional trade empirics focused on countries as the level of analysis, more recent work has complemented this approach with increased attention to the firm-level empirics of trade. Well known examples include Bernard and Jensen (1995) for the USA, Roberts and Tybout (1997) for Colombia, Pavcnik (2002) for Chile, and Clerides et al. (1998) for Colombia, Mexico, and Morocco. These and numerous other examples are reviewed by Bernard et al. (2007).

Most often, firm-level empirical work covers a single country, or at most a handful as in Clerides et al. (1998). This is because of the need for industry census data, which are collected on a national basis. Due to differences in survey instruments and accounting conventions, it can often be difficult to merge census data from multiple countries into a single, comparable dataset that can be used for cross-country analysis.

One attempt to fill this gap is led by the Enterprise Analysis Unit of the World Bank. The Bank engages private contractors to conduct firm-level surveys in developing countries, using instruments that can be matched to a common set of questions. This makes it possible to use firm-level data in a cross-country framework. To date, survey results are available for 98 countries, covering over 66,000 firms. Typically, the survey is administered to 250-1,500 businesses in each country. Firms are selected using stratified random sampling, although it is likely that large firms are over-sampled.
Development Perspectives on Trade Growth at the Extensive Margin

The sample includes firms engaged in international trade and those serving the domestic market only. A variety of manufacturing and services industries are included in each survey. Data can be downloaded freely from www.enterprisesurveys.org upon acceptance of confidentiality restrictions. This website also contains full details of the questionnaire and survey methodology used.

I exploit these Enterprise Survey data to conduct a preliminary exploration of the links between firm productivity and the number of foreign markets entered. (Result 1 above suggests a positive link between firm productivity and the number of export markets served.) The analysis proceeds in two steps. First, I extract firm level total factor productivity (TFP) as the residual from a standard production function. I then relate TFP to a measure of the number of foreign markets entered, based on each firm’s identification of its main export markets. As will become clear below, the dataset used here suffers from the severe limitation that firms were only asked to identify their three main export markets. It is therefore possible to obtain meaningful information for firms that enter only one or two foreign markets, but those recording three foreign markets might in fact be involved in trade with a far larger number of destinations. Although this difficulty can be partly dealt with by econometric means, it nonetheless greatly constrains the level of insight that can obtained using this approach. That is why the next section turns to more comprehensive trade flow data in an effort to supplement the preliminary analysis presented here, through a more rigorous test of Result 2.

1 Productivity Estimates

Indexing across exporting countries (e), sectors (s), firms (f), and time periods (t), a standard Cobb-Douglas production function takes the following form:

\[
\ln y_{esft} = \alpha_0 + \alpha_1 \ln l_{esft} + \alpha_2 \ln k_{esft} + \alpha_3 \ln m_{esft} + \varepsilon_{esft}
\] (II.16)
where \( l \) is labor, \( k \) is capital, and \( m \) is raw materials. For a given firm at time \( t \), an estimate of (log) TFP can be backed out of the above expression as \( \widehat{\alpha_0} + \widehat{\varepsilon_{ext}} \).

I estimate (II.16) for manufacturing industries using extracts from the Enterprise Survey dataset as set out in detail in Table 1. To measure output \( y \), I calculate net sales (i.e., total sales less interest and other costs). Labor input \( l \) is measured as the total number of permanent and temporary employees, capital \( k \) is the net book value of machinery and equipment, and raw materials \( m \) is the value of total raw materials purchases. Value data are converted from local currency to US dollars, then deflated using the US GDP deflator. Both conversion factors are taken from the IMF’s International Financial Statistics database.

Given the sample size and cross-country nature of the Enterprise Survey data, as well as their focus on developing countries, it is likely that this dataset is less reliable than those constructed by national statistics agencies on the basis of comprehensive census data for a single country. It is therefore particularly important to do some cleaning prior to estimation. I drop all observations reporting zeros or negative numbers for the above series. In addition, I exclude observations with ratios of wages and inputs purchases to net sales of greater than unity.

The full Enterprise Survey dataset for which responses match the standardized survey instrument consists of 30,330 manufacturing firms in 64 developing countries (i.e., countries other than those in the World Bank’s high income group). For most productivity related questions, firms are asked to provide data for three time periods: the previous year, two years ago, and three years ago. Thus, there is in principle a panel of 90,990 observations available. However, implementation of the first two exclusions above, combined with the fact that many firms do not answer all questions, greatly reduces the effective size of the sample.

Table 2 shows results from estimating (II.16) by OLS separately for each sector (but pooling across countries within sectors). Thirteen sectors have sufficient effective observations to make
such an approach feasible. The model performs well in all cases, with $R^2$ around 0.9. All estimated parameters have the expected positive signs and reasonable magnitudes (i.e., they sum to approximately one, indicating that returns to scale are either constant or slightly increasing). Raw materials and capital both have coefficients that are statistically significant at the 5% level in all sectors, whereas labor is only significant at the 10% level in seven of the thirteen sectors. For each sector, $t f p^{ols}$ is calculated as the sum of the intercept and the residual.

As is well known, a potential problem with the above estimates is that they may not be consistent if a profit maximizing firm observes a productivity shock in time to alter its use of production factors. This is essentially an endogeneity issue, for which a number of remedies exist. As a robustness check, I use the method suggested by Levinsohn and Petrin (2003) (LP). Those authors show that intermediate inputs can be used to control for the correlation between productivity shocks and factor use. They express (II.16) in terms of a composite error term $\varepsilon_{esft} = \omega_{esft} + \eta_{esft}$, of which the first part is the transmitted component of the productivity shock, and the second is random noise:

$$
\ln y_{esft} = \alpha_0 + \alpha_1 \ln l_{esft} + \alpha_2 \ln k_{esft} + \alpha_3 \ln m_{esft} + \omega_{esft} + \eta_{esft} \quad \text{(II.17)}
$$

The demand function for intermediates is a function of capital and the productivity shock. Thus, it can be inverted to give $\omega_{esft} = \omega_t (m_{esft}, k_{esft})$. Substituting $\phi_t (m_{esft}, k_{esft}) = \alpha_0 + \alpha_2 \ln k_{esft} + \alpha_3 \ln m_{esft} + \omega_t (m_{esft}, k_{esft})$ gives:

$$
\ln y_{esft} = \alpha_1 \ln l_{esft} + \phi_t (m_{esft}, k_{esft}) + \eta_{esft} \quad \text{(II.18)}
$$

Following Olley and Pakes (1996), Levinsohn and Petrin (2003) estimate (II.18) by OLS using a third order polynomial to approximate $\phi_t (m_{esft}, k_{esft})$. This gives a consistent estimate of the labor coefficient $\hat{\alpha}_1$. To identify the remaining coefficients, they assume that the productivity shock
follows a first order Markov process such that $\omega_{esf,t} = E [\omega_{esf,t} | \omega_{esf,t-1}] + \tilde{\omega}_{esf,t}$. Netting out labor and rearranging the production function gives:

$$\ln y_{esf,t} - \tilde{\alpha}_1 \ln l_{esf,t} = a_0 + a_2 \ln k_{esf,t} + a_3 \ln m_{esf,t} + E [\omega_{esf,t} | \omega_{esf,t-1}] + \tilde{\omega}_{esf,t} + \eta_{esf,t} \quad (II.19)$$

which can in turn be combined with a consistent estimate of $E [\omega_{esf,t} | \omega_{esf,t-1}]$ to obtain consistent estimates $\hat{\alpha}_0$, $\hat{\alpha}_2$, and $\hat{\alpha}_3$. Levinsohn and Petrin (2003) use a third order polynomial in $\omega_{esf,t-1}$ to approximate $E [\omega_{esf,t} | \omega_{esf,t-1}]$ and then propose a GMM estimator for the above expression using the moment conditions $E [\tilde{\omega}_{esf,t} + \eta_{esf,t} | m_{esf,t-1}] = 0$ and $E [\tilde{\omega}_{esf,t} + \eta_{esf,t} | k_{esf,t}] = 0.46$

Standard errors are obtained by bootstrapping.

Table 3 presents estimation results. Due to the more data intensive nature of the LP methodology as compared with OLS, a number of sectors drop out of the analysis. LP productivity estimates are therefore only available for nine sectors, rather than the thirteen analyzed using OLS. In line with the evidence presented by Levinsohn and Petrin (2003), the estimated labor coefficient using LP is generally quite close to that obtained under OLS. However, the capital and raw materials coefficients are sometimes substantially different. LP estimates of the labor and raw materials coefficients are generally precise: the labor coefficient is statistically significant at the 5% level in all nine regressions, and the raw materials coefficient is 10% significant in six out of the nine sectors. Results are more disappointing for capital, for which the estimated coefficient is only statistically significant at the 10% level in three sectors. In part, decreased precision is the price to be paid for the LP two-step estimation procedure. It possibly also reflects the application of different accounting standards when determining net book value for survey purposes.

Although the LP coefficients in a number of sectors sum to values that appear to be substantially

\footnote{The mechanics of their approach as implemented in Stata are set out in detail in Petrin et al. (2003). I use the accompanying do file to obtain the results reported here.}

SHEPHERD, Benjamin - Thèse IEP de Paris - 2008 93
less than one—which would be consistent with decreasing returns to scale—formal hypothesis tests provide greater nuance. Returns appear to be decreasing in four of the nine sectors, and increasing in one (chemicals and pharmaceuticals). This last result seems sensible, but it would be important in future work with these data to look more closely at the possibility of decreasing returns in the other sectors. However, given the exploratory nature of the analysis undertaken here, this is not an issue that is considered further at this stage.

2 Productivity and the Number of Export Markets

Result 1 above suggests that amongst exporting firms, higher productivity should be associated with a greater number of foreign markets entered. This claim is closely related to the more common one to the effect that exporters tend to be more productive than non-exporters (e.g., Clerides et al., 1998; Bernard and Jensen, 1999). The difference is that while the latter concentrates on the decision whether or not to access at least one foreign market, the former is more interested in the subsequent decision of how many of those markets to enter. There is also some similarity between the contention in this paper and the analysis of Eaton et al. (2005), in which the authors relate the number of export markets entered to firm size (domestic sales).

The productivity estimates derived above using OLS and LP provide the basis for a preliminary exploration of the empirical validity of Result 1. In order to expand the available sample, I use alternately a simple measure of labor productivity (net sales per employee). Unfortunately, the analysis can only be preliminary at this stage due to severe data limitations affecting the ability to measure the number of export markets entered by each firm. The Enterprise Survey instrument does not require respondents to identify all markets served, but only the three main ones served in the survey year.\footnote{This makes it possible to distinguish amongst exporters that deal with one, two, or at least three markets. Since this paper postulates a causal link running from productivity to the entrance of multiple foreign markets, we cannot utilize these data to analyze export market entry and exit decisions over time.}
number of export markets, this means that the dependent variable is both left- and right-censored, leaving only a very small range for uncensored observations.

To compound this difficulty, exporting firms frequently do not respond at all to the question asking them to list their most important markets. I adopt two approaches to deal with this. In the baseline specification, I simply drop all such observations from the dataset. Under this definition, 674 firms export to a single market, 1383 firms export to two markets, and 541 export to at least three. In an alternative specification, I code all exporters as serving a single foreign market unless it is explicitly indicated that they serve two or three. This results in a count of 8396 exporters serving one market, with unchanged numbers serving 2 or 3+ foreign markets.

Table 4 contains results of OLS regressions using the number of export markets entered as the dependent variable and productivity as the explanatory variable (both in logarithms). Fixed effects by country-sector account for factors such as trade costs and market size. For both the baseline and alternative export market counts, as well as all three productivity measures, Table 4 is suggestive of a weak, positive relationship between productivity and geographical export diversification. However, the relationship is only statistically significant at the 10% level in three of the six specifications in the table. The estimated coefficients vary somewhat, but are consistent with an interpretation in which a 1% increase in firm productivity is associated with a 0.01% to 0.04% increase in the number of export destinations.

OLS estimates may not be consistent in light of the censoring problems noted above. Tables 5-6 present Tobit estimates of the same model, with left censoring at unity and right censoring at three. Table 5 includes standard fixed effects in the country and sector dimensions, while Table 6 uses random effects by country-sector. The reason for presenting both sets of results is that fixed effects Tobit estimates can suffer from bias and inconsistency due to the "incidental parameters problem" that arises frequently in nonlinear maximum likelihood settings. In essence, it is due to
the impossibility of conditioning the fixed effects out of the likelihood function, as is possible for a limited number of other estimators such as OLS and Poisson (see Greene, 2004, for a discussion). While the random effects model does not suffer from this difficulty, it imposes a considerably more restrictive structure on the data.

In all but one case, the Tobit estimates reported in Tables 5 and 6 go in the same direction as the OLS estimates in Table 4: they are consistent with a weak, positive effect of productivity on the number of export markets entered. Estimated coefficients range from approximately 0.01 to 0.1, and five of the twelve are statistically significant.48 Taking the results in Tables 4-5 together, the Enterprise Survey data appear to provide some preliminary empirical support for Result 1 above, namely that the number of export markets served is increasing in firm productivity. But given the data limitations already mentioned, it will clearly be necessary to refrain from drawing any strong conclusions for the time being.

D Country Level Empirics

While it is to be hoped that future work with more detailed data will produce more robust firm-level findings, the remainder of this paper shifts the analysis to the aggregate level. As the theory presented above makes clear, Results 1 and 2 are closely related: the latter suggests that the maximum level of firm productivity, as well as trade costs and market size, should be positively associated with the number of export markets entered on an aggregate (country-sector) level. In this section, I conduct a detailed empirical investigation of Result 2 using standard international trade data. This allows me to cover a far wider range of countries than in the previous section. It also makes it possible to construct export market counts that are not truncated at a relatively low

48However, the fixed effects Tobit results should be interpreted with caution for two reasons. First, they are subject to bias and inconsistency as previously noted. Second, the standard errors are robust, but have not been adjusted for clustering. They may therefore be understated.
number, as was the case previously due to the way in which the Enterprise Survey instrument is designed.

The empirical strategy in this section is straightforward. I first use HS 6 digit trade data to construct a count $m_{es}$ of the number of export markets entered by each exporting country $e$ in each 2 digit HS sector $s$. An export relationship is considered to exist whenever strictly positive exports are recorded for at least one 6 digit product line in a 2 digit sector. I then relate these export market counts to trade costs, market size, and technology (productivity) using a Poisson model. After a series of robustness checks, I adopt two methods to deal with possible endogeneity issues. First, I eliminate sector fixed effects by measuring all variables relative to a common comparator country (the USA). Second, I use (alternately) official language and colonial history as instruments for trade and market entry costs, based on the idea that these data should be reflective of a country’s institutions as adapted in many cases from the colonial period. At all points in the analysis, the data provide strong support for the core contentions of this paper.

1 Data

Data and sources are set out in full in Table 1. Since the focus in this paper is on developing countries, the dataset excludes countries in the World Bank’s high income group. Many of the data used, such as GDP and GDP per capita, official language dummies, and colonization dummies, come from standard sources and do not require any particular discussion. However, three aspects are more novel and are discussed in detail here: export market counts, trade costs, and domestic market entry costs. Due to limited availability of trade cost data, the analysis takes place using data for a single year only (2005).

To construct counts of the number of export markets served by exporter-sector ($m_{es}$), I use

---

49This is a broad definition of what constitutes a developing country. It is narrowed in the context of robustness checks, below.
the BACI trade data included in the Market Access Map (MACMap) database produced by CEPII. BACI is based on standard UN Comtrade data at the 6 digit HS level. The main difference for present purposes is that BACI uses a harmonization methodology to reconcile mirror flows, thereby providing more complete geographical coverage than if only a single direction of Comtrade statistics were to be used. BACI’s approach to harmonization consists of computing a weighted average of mirror flows based on an estimated quality indicator for export and import declarations in each country. The methodology is set out in detail in Gaulier et al. (2007).

Trade costs can cover numerous dimensions. Here, I focus on three of the most important. As is common in the gravity literature, I use international distance as a proxy for transport costs. Since data are by exporter (not bilateral), I take the average distance of each exporter from the rest of the world. The second dimension of trade costs captured here is effectively applied tariffs (i.e., including preferences). These are sourced from the MACMap database for the year 2004, and aggregated to the HS 2 digit level using the reference groups methodology of Laborde et al. (2007). The essence of that aggregation approach is to limit endogeneity concerns by weighting tariffs according to the imports of a group of similar countries (not the importer itself). MACMap tariffs include both ad valorem and specific measures, with the latter converted to ad valorem terms using the reference group unit value. Again, I take the simple average across the rest of the world to obtain a figure for each exporter.

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

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

To proceed with the empirical analysis, it is assumed that \( m_{es} \) can be adequately represented by a Poisson process. This is appropriate given that \( m_{es} \) represents strictly non-negative integer count data. The mean and variance of that process are equal to \( \mu_{es} \) (where \( e \) indexes exporters and \( s \) indexes sectors), and its density conditional on a set of independent variables \( X_{es} \) is given by:

\[
f (m_{es} \mid X_{es}) = \frac{\exp (-\mu_{es}) \mu_{es}^{m_{es}}}{m_{es}!} \quad (I.20)
\]

In line with (I.11), a baseline formulation for the conditional mean function is:

\[
\mu_{es} = \delta_s \exp \left[ \beta_1 \ln(entry_e) + \beta_2 \ln(export_e) + \beta_3 \ln(dist_e) + \ldots \right.
\]

\[
\ldots + \beta_4 \ln(1 + t_{es}) + \beta_5 \ln(gdp_e) + \beta_6 \ln(gdppc_e) \right] \quad (I.21)
\]

Export costs, distance, and tariffs capture the trade costs faced by exporters, while the exporting country’s own GDP proxies the size of the home market. Per capita GDP in the exporting country is used as a proxy for the technology parameter \( c_{low} \). Even though entry costs do not appear explicitly in (I.11), they are included in the estimating equation (I.21) because of their potential to exert a second order influence on \( m_{es} \) through the set of national demand shifters \( d^i \). Finally, the sector fixed effects \( \delta_s \) control for unobservables that impact all exporters in a given sector in the same way. An important example of such a factor is worldwide demand by sector.

The comparative statics presented above (see equations I.12, I.14, and I.15) suggest that \( \beta_2, \beta_3, \) and \( \beta_4 \) should all be negative. It is also clear that \( \beta_6 \) should be positive, to the extent that higher per capita GDP is associated with a higher maximum level of firm productivity \( c_{low} \). Although this paper does not derive explicit comparative statics results for market size (GDP) and entry costs, by analogy with previous work such as Helpman et al. (Forthcoming) and Hummels and Klenow...
it is anticipated that $\beta_1$ will be negative but that $\beta_5$ will be positive.

Estimation of the model contained in (II.20) and (II.21) can be undertaken using standard maximum likelihood techniques. For details, see Cameron and Trivedi (2001) and Winkelmann (2000). Indeed, the Poisson estimator can be shown to be consistent under relatively weak assumptions. All that is required is correct specification of the conditional mean function. The data need not in fact be distributed as Poisson (Wooldridge, 1997). Unlike most nonlinear panel data models, Poisson does not suffer from the "incidental parameters problem" when fixed effects are included. As shown by Hausman et al. (1984), the fixed effects can be conditioned out of the likelihood function, which can then be maximized as usual.

Poisson results for the baseline specification appear in column 1 of Table 7. Overall, the model does a good job of explaining the observed variation in export destinations: $R^2$ is 0.70. All parameters carry the expected signs and have sensible magnitudes: entry costs, export costs, distance, and tariffs are all negatively associated with the number of export markets entered, while the two GDP variables exhibit a positive association. In terms of absolute value elasticities, the strongest trade cost impacts come from (in descending order) distance, tariffs, export costs, and finally entry costs. This ordering is unsurprising in light of the direct impacts that the first three variables have on $m_{es}$ through (II.11), as compared with the indirect impact of entry costs. In terms of precision, all parameters except entry costs and tariffs are statistically significant at the 10% level. The coefficient on entry costs is significant at the 20% level.

Thus far, the data tend to support the core contentions of this paper. Concretely, 1% reductions in international transport costs, export costs, and importer tariffs are associated with increases of 0.3%, 0.2%, and 0.3% respectively in the number of export destinations serviced. The effect of

\footnote{The coefficient of determination is calculated as in the usual linear model, i.e. $R^2 = 1 - \frac{ESS}{TSS}$. Although strictly speaking this is not appropriate for a Poisson model (which has no residual), it nonetheless provides a useful guide as to overall model fit (see e.g., Wooldridge, 1997). The more common pseudo-$R^2$ measures are considerably more difficult to interpret.}
market entry costs is weaker, however: a 1% reduction is associated with a 0.1% increase in export markets. In the remainder of this section, a number of specification checks are applied to ensure that these results are robust.

3 Robustness Checks

First, I conduct simple robustness checks covering two important dimensions: definition of the dependent and independent variables, and constitution of the estimation sample.

Columns 2 and 3 of Table 7 present results for the same baseline model, but with per capita GDP replaced by more direct measures of technology. Since (II.11) suggests that the number of export markets entered nationally depends on the productivity of the most productive firm in the economy, I use Enterprise Survey data to build such measures on a country-sector basis.\(^{51}\) Column 2 uses firm-level TFP estimates obtained using the Levinsohn and Petrin (2003) methodology outlined above. Since data availability severely limits the number of countries and sectors for which such estimates are feasible, Column 3 also reports results using a simple measure of labor productivity (net sales per employee). Results in both columns are suggestive of a positive association between maximum productivity and the number of export markets entered, since the relevant coefficients have positive signs in both cases. However, only when labor productivity is used is the estimated coefficient statistically significant at the 10% level. Indeed the labor productivity formulation is also preferable based on the sign of the estimated coefficient on export costs, which unexpectedly turns positive when TFP is used. These problems are undoubtedly due to the very small sample for the TFP regression (only 682 observations, compared with 11,808 in the baseline formulation). The labor productivity regression still has a small sample (2,372 observations), but one which is nearly four times as large as the TFP regression.

Next, I consider alternative dependent variable definitions. As previously noted, the dataset

\(^{51}\)The Enterprise Survey dataset identifies sectors verbally. I map these manually to Harmonized System codes.
used here counts an additional export destination whenever there is a non-zero export flow recorded in trade data. However, it could be objected that this risks introducing noise into the data by including very small flows that might be linked to misclassification or other problems of data quality. Further, it could be argued from a policy perspective that it is more relevant to focus on larger flows. To take account of these issues, I re-estimate the model using thresholds of $0.1m and $1m, below which a 6 digit trade flow is considered to be zero.

Results appear in columns 4-5 of Table 7. In both cases, coefficient signs are identical to those reported for the baseline model. Export costs, GDP, and per capita GDP remain statistically significant at the 10% level. Although the distance coefficient loses its statistical significance, tariffs become statistically significant once the higher thresholds are imposed. Interestingly, the magnitude of the estimated coefficients increases markedly in absolute value for export costs and tariffs, whereas it declines for distance. It is clear, therefore, that imposing thresholds on the export market count procedure tends to reinforce the baseline results.

Next, I address the issue that the country definition used above (all countries except those in the high income group) is very wide, and potentially groups together extremely different countries. The last two columns of Table 7 exclude (in addition) upper- and lower-middle income countries from the estimation sample. As can be seen, the reduced sample leads to some changes in estimated coefficients. Export costs and GDP per capita lose statistical significance when the sample excludes upper-middle income countries, whereas entry costs now become statistically significant at the 10% level. The distance coefficient loses statistical significance, and has an unexpected positive sign. These results carry over to the model which excludes in addition low-income countries, with the exception that tariffs become statistically significant and GDP per capita has an unexpected negative sign. The coefficient on entry costs is now significant at the 15% level only. All in all, these results are generally supportive of the baseline model, although the reduction in sample size
has a noticeable impact on the precision with which some coefficients are estimated. Although there are some sign changes, none are statistically significant. Thus, focusing on a narrower sample of countries does not fundamentally alter the results presented earlier.

4 Instrumental Variables Results

Endogeneity is obviously a crucial issue for a model like (II.21). I deal with it in two ways. First, I re-estimate the model after transforming all variables so that they are expressed as ratios with respect to a constant comparator country (the USA). This first difference formulation eliminates the sector fixed effects, and therefore avoids any problems that might result from endogeneity of those fixed effects. Second, I use colonization and language as instrumental variables to identify exogenous variation in entry and export costs. This strategy is based on the idea that such costs arise within an institutional context that is heavily influenced by colonial history in many countries; however, past colonization status should be exogenous to the current number of export partners.

Results for the first difference formulation are reported in column 1 of Table 9. In general, they are stronger than those reported for the baseline formulation. The coefficients on all entry and trade cost variables are now significant at the 5% level, with noticeably larger magnitudes (in absolute value terms) in the case of entry costs, export costs, and distance. The only unexpected result is that tariffs now have a positive coefficient that is statistically significant at the 1% level. This outcome clearly does not fit within the analytical framework of the paper. The factors behind it will need to be more closely investigated in future work, although as will be shown below it is not robust to a number of other changes in the estimation methodology. This suggests that it would be appropriate not to interpret this tariff coefficient too strictly at the present time.

Table 8 presents results using (alternately) colonization and language as instruments for export and entry costs. Since results are generally very similar regardless of which set of instruments is
used, the discussion here focuses on the language instruments since they perform more strongly in first stage regressions.

For the purposes of instrumental variables estimation, all variables except export and entry costs are treated as exogenous. This is a consistent approach under the assumption that world average tariffs are determined exogenously to individual export flows, which is likely to be a reasonable approximation at least for "small" countries. In order to ensure exogeneity of GDP (aggregate and per capita), these variables are lagged by 5 periods. As column 1 of the table shows, using 2000 rather than 2005 income figures on its own does not make any major difference to the estimated coefficients in terms of magnitudes, although the coefficient on entry costs is now statistically significant at the 5% level.

The following three columns use as instruments dummies for countries with English, French, Spanish, Portuguese, and Russian as an official language. Data are sourced from Mayer and Zignago (2006), and first stage estimation is by OLS. F-tests are strong in both cases, with the null hypothesis that all instrument coefficients are equal to zero rejected at the 1% level. Coefficients of determination are 0.32 for entry costs and 0.21 for export costs, suggesting that the instruments (together with the exogenous explanatory variables) do a reasonable job in accounting for the observed variation in the potentially endogenous explanatory variables.

The second stage Poisson regression in column 2 includes the residuals from the two first stage OLS regressions as additional regressors. This is a simple way of accounting for the effects of endogeneity, analogous to two stage least squares (see Wooldridge, 2002). To gauge the extent of possible endogeneity bias, I first test the hypothesis that the two residual terms have coefficients jointly equal to zero. The null is rejected at the 1% level, suggesting that endogeneity is indeed a serious issue in these data. In terms of magnitude, endogeneity would appear to have biased the coefficients on export and entry costs upwards (i.e., towards zero in an absolute value sense).
Instrumental variables estimates that correct for this bias result in seemingly stronger effects of entry and trade costs on export market diversification.\footnote{The standard errors reported in Table 8 likely understate the true values, since they are not corrected for the use of estimated residuals from the first stage. It is therefore necessary to be cautious in drawing conclusions as to statistical significance either of the coefficients themselves, or of the apparent changes between baseline and IV estimates.}

Columns 2-4 of Table 9 use the same instrumental variables approach but with all variables transformed by taking ratios with respect to the USA, thereby eliminating sector fixed effects. Results are very similar to the IV results using untransformed variables, with the exception of tariffs which again have an unexpected positive coefficient.

An important additional issue for IV estimation is establishing that the instruments are genuinely exogenous. Since the literature does not disclose an appropriate testing approach in the case of the Poisson IV model, I re-estimate the model in first differences using standard two stage least squares (Table 9 columns 5-7). Results are quite similar to the Poisson case, suggesting that the change in estimation methodology does not itself pose any major problems. While TSLS coefficients are slightly smaller in absolute value than their Poisson counterparts, the differences are not large given the precision of the estimates involved. Interestingly, the tariff coefficient remains positive, but is not statistically significant in this case–a result that better accords with expectations than those obtained using Poisson.

Most importantly, the standard tests for instrument validity in the TSLS context suggest that language dummies are appropriate in this case. The F-tests of joint significance in the first stage regressions reject the null at the 1\% level. In addition, the fact that the model is overidentified–five instruments for two potentially endogenous explanatory variables–makes it possible to test for instrument exogeneity. The chi square test does not reject the null at the 10\% level (prob. = 0.11), indicating that the exogeneity requirement is satisfied.

One possible reason for the relatively high overidentification test statistic value is that language dummies might themselves represent a dimension of foreign market entry costs, perhaps related
to the difficulty of acquiring information on foreign market conditions. If true, this would sug-
gest that they should be entered in the main regression equation, and therefore would not satisfy
the exclusion restriction necessary for valid instruments. While this is a legitimate concern, the
outcome of the overidentification test in this case suggests that the problem is not unduly serious
here. Nonetheless, it will be important in future work to identify alternative instrumental variables
strategies as a robustness check on the results reported here.\textsuperscript{53}

\section*{Conclusions and Further Research}

This paper has provided some of the first evidence on the factors driving the geographical
spread of developing country trade. On the one hand, international trade flow data have been used
to show that the number of foreign markets entered depends on the costs of trade and domestic
market entry, as well as technology (firm productivity) in the exporting country. In the preferred
econometric specification using language to instrument for trade and entry costs, it is found that 1%
reductions in entry and export costs are associated with increases of 0.4\% and 0.6\% respectively in
the number of export markets served. A similar decrease in international transport costs (proxied
by distance) is associated with a 0.7\% increase in the number of export destinations. In addition,
World Bank \textit{Enterprise Survey} data have been used to provide direct, though preliminary, evidence
on the links between productivity and export market diversification on a firm level. Results from
both the firm- and country-level analysis are consistent with a framework in which more productive
firms self-select into servicing a larger number of export markets. However, more work is clearly
needed on the firm level empirics before any strong conclusions can be drawn.

\textsuperscript{53}In additional tests (not reported), it was found that other likely instruments were generally too weakly correlated
with export costs to be useful. This is true for geographical variables such as latitude and country surface area, as
well as for the settler mortality data that Acemoglu et al. (2001) use to instrument for institutional quality. The
present sample does not have sufficient variance in legal origin to allow effective use of the instrumentation approach
of Djankov et al. (2002).
The policy implications of these findings are significant from a development point of view. As noted at the outset, export diversification is back on the policy agenda of many developing countries. This research, as well as the previous findings of Evenett and Venables (2002) and Brenton and Newfarmer (2007), confirm that it will be important for policymakers to be clear on the benefits and costs of measures designed to promote diversification in both the product and market dimensions. The results in this paper and in Dennis and Shepherd (2007) suggest that policies aimed at reducing trade costs and domestic market entry costs may well be effective tools in promoting diversification. Improved trade facilitation leading to lower export costs could also play an important role. Interestingly, both this paper and Dennis and Shepherd (2007) find only weak evidence supporting the expected link between improved foreign market access and export diversification. While there could be many reasons for this, such as inaccuracies in the tariff data or problems linked to aggregation, it is nonetheless striking that market access does not appear at this point to be the binding constraint on diversification in developing countries.

There are a number of ways in which future research could extend the results presented here. As previously noted, more extensive firm-level datasets should make it possible to measure the number of export markets entered in considerably more detail than has been possible here. For instance, the French data presented by Eaton et al. (2005) contain detailed information on this point. The price of using such data may well be that a single country framework is required. But combining single country results with the preliminary cross-country firm-level evidence presented here would be a considerable step forward.

There is also room for additional work at the aggregate level. Here, a pure cross-section has been used due to the focus on trade cost factors for which data are not yet available over a long time period. It could be thought, however, that geographical diversification is not an instantaneous process. There would therefore be some value in attempting to replicate these results over a longer
time period, albeit at the price of excluding *Doing Business* trade cost data and using less detailed (i.e., TRAINS) tariff data.

On a technical level, it would be desirable to experiment further with alternative instruments for trade and entry costs. Although language appears to have done a reasonable job here, there is a genuine question as to whether it adequately satisfies the second stage exclusion restriction. Indeed, the question of endogenous trade costs is one that needs to be addressed more generally in the empirical trade literature: gravity models that instrument for tariffs or other potentially endogenous variables are surprisingly rare.

Brenton and Newfarmer (2007) bring out the importance of market diversification as a channel for developing country export growth. A number of other papers (e.g., Hummels and Klenow, 2005) emphasize product diversification. Just as it is important to be clear about the relative importance of these channels, so too is there an important research question as to the relative impacts of policy changes on each type of export diversification. A simple comparison of Table 8 column 2 in this paper with Table 3 column 2 of Dennis and Shepherd (2007) suggests that there may well be interesting differences to be examined. For instance, domestic market entry costs, tariffs, and per capita GDP as a proxy for technology or firm productivity appear to exert a noticeably stronger impact on geographical diversification than on product diversification. Future work in which these two channels of diversification are combined both theoretically and empirically would therefore seem to be of particular interest from a policy point of view.

**References**


Dennis, Allen, and Ben Shepherd, 2007, "Trade Costs, Barriers to Entry, and Export Diversifi-


### G Tables

*Table 1: Data and sources.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Units</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital</td>
<td>Net book value of machinery and equipment.</td>
<td>LCU</td>
<td>Enterprise Survey</td>
</tr>
<tr>
<td>Col</td>
<td>Dummy variable for countries colonized by the British, French, Spanish, Portuguese, or Russians.</td>
<td>NA</td>
<td>Mayer &amp; Zignago (2006)</td>
</tr>
<tr>
<td>Distance</td>
<td>Average of the great circle distances between the main cities of the exporting country and all other countries.</td>
<td>km</td>
<td>Mayer &amp; Zignago (2006)</td>
</tr>
<tr>
<td>Entry Cost</td>
<td>Official cost of starting up and formally operating an industrial or commercial business in the exporting country (2006).</td>
<td>USD</td>
<td>Doing Business</td>
</tr>
<tr>
<td>Export Cost</td>
<td>Official fees levied on a 20 foot container leaving the exporting country, including document preparation, customs clearance, technical control, terminal handling charges, and inland transit (2006).</td>
<td>USD</td>
<td>Doing Business</td>
</tr>
<tr>
<td>GDP, GDP00</td>
<td>Gross domestic product in 2005 or 2000.</td>
<td>USD</td>
<td>WDI</td>
</tr>
<tr>
<td>GDPPC, GDPPC00</td>
<td>Per capita gross domestic product in 2005 or 2000.</td>
<td>USD</td>
<td>WDI</td>
</tr>
<tr>
<td>Labor</td>
<td>Total number of workers.</td>
<td>Count</td>
<td>Enterprise Survey</td>
</tr>
<tr>
<td>Lang</td>
<td>Dummy variable for countries with the following official languages: English, French, Spanish, Portuguese, and Russian.</td>
<td>NA</td>
<td>Mayer &amp; Zignago (2006)</td>
</tr>
<tr>
<td>Materials</td>
<td>Total purchases of raw materials (excluding fuel).</td>
<td>LCU</td>
<td>Enterprise Survey</td>
</tr>
<tr>
<td>Net Sales</td>
<td>Total sales less interest/financial charges and other costs.</td>
<td>LCU</td>
<td>Enterprise Survey</td>
</tr>
<tr>
<td>Sales/Empl.</td>
<td>Net Sales / Labor</td>
<td>NA</td>
<td>Enterprise Survey</td>
</tr>
<tr>
<td>Tariffs</td>
<td>Average applied ad valorem tariff in the rest of the world, by HS-2 sector. Aggregated by reference groups (Laborde et al., 2007).</td>
<td>%</td>
<td>MAcMap</td>
</tr>
</tbody>
</table>
Table 2a: Estimated production functions by sector (OLS).

<table>
<thead>
<tr>
<th></th>
<th>Textiles</th>
<th>Leather</th>
<th>Garments</th>
<th>Food</th>
<th>Beverages</th>
<th>Metals &amp; Machinery</th>
<th>Electronics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>0.03</td>
<td>0.175***</td>
<td>0.154**</td>
<td>0.05</td>
<td>0.17</td>
<td>0.07</td>
<td>0.152**</td>
</tr>
<tr>
<td></td>
<td>[0.056]</td>
<td>[0.054]</td>
<td>[0.062]</td>
<td>[0.032]</td>
<td>[0.098]</td>
<td>[0.039]</td>
<td>[0.068]</td>
</tr>
<tr>
<td>Capital</td>
<td>0.154***</td>
<td>0.102**</td>
<td>0.200***</td>
<td>0.235***</td>
<td>0.161**</td>
<td>0.124***</td>
<td>0.096**</td>
</tr>
<tr>
<td></td>
<td>[0.037]</td>
<td>[0.042]</td>
<td>[0.049]</td>
<td>[0.026]</td>
<td>[0.071]</td>
<td>[0.025]</td>
<td>[0.038]</td>
</tr>
<tr>
<td>Materials</td>
<td>0.766***</td>
<td>0.729***</td>
<td>0.666***</td>
<td>0.694***</td>
<td>0.799***</td>
<td>0.798***</td>
<td>0.744***</td>
</tr>
<tr>
<td></td>
<td>[0.043]</td>
<td>[0.055]</td>
<td>[0.057]</td>
<td>[0.026]</td>
<td>[0.069]</td>
<td>[0.025]</td>
<td>[0.023]</td>
</tr>
<tr>
<td>Constant</td>
<td>1.135***</td>
<td>0.929***</td>
<td>1.123***</td>
<td>1.206***</td>
<td>0.32</td>
<td>0.981***</td>
<td>1.104***</td>
</tr>
<tr>
<td></td>
<td>[0.096]</td>
<td>[0.161]</td>
<td>[0.148]</td>
<td>[0.109]</td>
<td>[0.239]</td>
<td>[0.123]</td>
<td>[0.105]</td>
</tr>
<tr>
<td>Observations</td>
<td>3143</td>
<td>935</td>
<td>4309</td>
<td>4398</td>
<td>231</td>
<td>3165</td>
<td>2527</td>
</tr>
<tr>
<td>Countries</td>
<td>27</td>
<td>19</td>
<td>25</td>
<td>25</td>
<td>10</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>R²</td>
<td>0.93</td>
<td>0.91</td>
<td>0.89</td>
<td>0.94</td>
<td>0.95</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>Coef. Sum</td>
<td>0.95</td>
<td>1.01</td>
<td>1.02</td>
<td>0.98</td>
<td>1.13</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>CRS (F-Test)</td>
<td>5.08**</td>
<td>0.03</td>
<td>0.40</td>
<td>0.67</td>
<td>3.33</td>
<td>0.19</td>
<td>0.07</td>
</tr>
</tbody>
</table>

1. Dependent variable is net sales. All variables are in logarithms.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%). Coef. Sum indicates the sum of the coefficients on labor, capital, and raw materials. CRS reports the F-statistic for the null hypothesis that the three coefficients sum to unity.
Table 2b: Estimated production functions by sector (OLS).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Labor</strong></td>
<td>0.05</td>
<td>0.06</td>
<td>0.153***</td>
<td>0.141***</td>
<td>0.045*</td>
<td>0.168*</td>
</tr>
<tr>
<td></td>
<td>[0.041]</td>
<td>[0.046]</td>
<td>[0.043]</td>
<td>[0.037]</td>
<td>[0.021]</td>
<td>[0.062]</td>
</tr>
<tr>
<td><strong>Capital</strong></td>
<td>0.215***</td>
<td>0.102***</td>
<td>0.161***</td>
<td>0.175***</td>
<td>0.108***</td>
<td>0.076***</td>
</tr>
<tr>
<td></td>
<td>[0.037]</td>
<td>[0.024]</td>
<td>[0.047]</td>
<td>[0.016]</td>
<td>[0.022]</td>
<td>[0.006]</td>
</tr>
<tr>
<td><strong>Materials</strong></td>
<td>0.748***</td>
<td>0.846***</td>
<td>0.743***</td>
<td>0.707***</td>
<td>0.840***</td>
<td>0.757***</td>
</tr>
<tr>
<td></td>
<td>[0.044]</td>
<td>[0.015]</td>
<td>[0.044]</td>
<td>[0.033]</td>
<td>[0.023]</td>
<td>[0.030]</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>0.982***</td>
<td>0.910***</td>
<td>0.724***</td>
<td>0.955***</td>
<td>0.797***</td>
<td>0.967***</td>
</tr>
<tr>
<td></td>
<td>[0.175]</td>
<td>[0.123]</td>
<td>[0.093]</td>
<td>[0.135]</td>
<td>[0.058]</td>
<td>[0.099]</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>2091</td>
<td>2152</td>
<td>1935</td>
<td>429</td>
<td>762</td>
<td>1466</td>
</tr>
<tr>
<td><strong>Countries</strong></td>
<td>20</td>
<td>21</td>
<td>17</td>
<td>14</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td><strong>$R^2$</strong></td>
<td>0.94</td>
<td>0.95</td>
<td>0.94</td>
<td>0.88</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td><strong>Coef. Sum</strong></td>
<td>1.01</td>
<td>1.01</td>
<td>1.06</td>
<td>1.02</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>CRS (F-Test)</strong></td>
<td>0.17</td>
<td>0.03</td>
<td>7.88**</td>
<td>1.03</td>
<td>0.17</td>
<td>0.00</td>
</tr>
</tbody>
</table>

1. Dependent variable is net sales. All variables are in logarithms.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%). Coef. Sum indicates the sum of the coefficients on labor, capital, and raw materials. CRS reports the F-statistic for the null hypothesis that the three coefficients sum to unity.
### Table 3a: Estimated production functions by sector (Levinsohn-Petrin).

<table>
<thead>
<tr>
<th>Sector</th>
<th>Textiles</th>
<th>Leather</th>
<th>Garments</th>
<th>Beverages</th>
<th>Metals/Machinery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>0.050***</td>
<td>0.200***</td>
<td>0.179***</td>
<td>0.180**</td>
<td>0.129***</td>
</tr>
<tr>
<td></td>
<td>[0.018]</td>
<td>[0.042]</td>
<td>[0.019]</td>
<td>[0.070]</td>
<td>[0.021]</td>
</tr>
<tr>
<td>Capital</td>
<td>0.096</td>
<td>0.267</td>
<td>0.114*</td>
<td>0.33</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>[0.113]</td>
<td>[0.220]</td>
<td>[0.059]</td>
<td>[0.283]</td>
<td>[0.043]</td>
</tr>
<tr>
<td>Materials</td>
<td>0.324</td>
<td>0.496**</td>
<td>0.334***</td>
<td>0.122</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>[0.241]</td>
<td>[0.207]</td>
<td>[0.096]</td>
<td>[0.343]</td>
<td>[0.162]</td>
</tr>
<tr>
<td>Observations</td>
<td>3143</td>
<td>935</td>
<td>4309</td>
<td>231</td>
<td>3165</td>
</tr>
<tr>
<td>Countries</td>
<td>27</td>
<td>19</td>
<td>25</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Coef. Sum</td>
<td>0.47</td>
<td>0.96</td>
<td>0.63</td>
<td>0.63</td>
<td>0.36</td>
</tr>
<tr>
<td>CRS ($\chi^2$)</td>
<td>4.17**</td>
<td>0.03</td>
<td>17.88***</td>
<td>0.69</td>
<td>14.28***</td>
</tr>
</tbody>
</table>

1. Dependent variable is net sales. All variables are in logarithms.

2. Bootstrapped standard errors are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%). Coef. Sum indicates the sum of the coefficients on labor, capital, and raw materials. CRS reports the $\chi^2$-statistic for the null hypothesis that the three coefficients sum to unity.

### Table 3b: Estimated production functions by sector (Levinsohn-Petrin).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>0.041***</td>
<td>0.083***</td>
<td>0.161***</td>
<td>0.128***</td>
</tr>
<tr>
<td></td>
<td>[0.019]</td>
<td>[0.021]</td>
<td>[0.024]</td>
<td>[0.033]</td>
</tr>
<tr>
<td>Capital</td>
<td>0.642***</td>
<td>0.034</td>
<td>0.102*</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>[0.128]</td>
<td>[0.033]</td>
<td>[0.055]</td>
<td>[0.124]</td>
</tr>
<tr>
<td>Materials</td>
<td>0.935***</td>
<td>0.554**</td>
<td>0.323*</td>
<td>0.806***</td>
</tr>
<tr>
<td></td>
<td>[0.208]</td>
<td>[0.224]</td>
<td>[0.173]</td>
<td>[0.239]</td>
</tr>
<tr>
<td>Observations</td>
<td>2091</td>
<td>2152</td>
<td>1935</td>
<td>1466</td>
</tr>
<tr>
<td>Countries</td>
<td>20</td>
<td>21</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>Sum of Coefs.</td>
<td>1.62</td>
<td>0.67</td>
<td>0.59</td>
<td>1.01</td>
</tr>
<tr>
<td>CRS ($\chi^2$)</td>
<td>7.33***</td>
<td>2.19</td>
<td>7.04***</td>
<td>0</td>
</tr>
</tbody>
</table>

1. Dependent variable is net sales. All variables are in logarithms.

2. Bootstrapped standard errors are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%). Coef. Sum indicates the sum of the coefficients on labor, capital, and raw materials. CRS reports the $\chi^2$-statistic for the null hypothesis that the three coefficients sum to unity.
Table 4: OLS regression results for number of export markets vs. firm productivity.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Alternative</th>
<th>Base</th>
<th>Alternative</th>
<th>Base</th>
<th>Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tfp^{OLS}$</td>
<td>0.019*</td>
<td>0.005*</td>
<td>[0.010]</td>
<td>[0.003]</td>
<td>0.039</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.033]</td>
<td>[0.014]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$tfp^{LP}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales/Empl.</td>
<td>0.012</td>
<td>0.019**</td>
<td>[0.009]</td>
<td>[0.009]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.629***</td>
<td>0.287***</td>
<td>0.537***</td>
<td>0.236***</td>
<td>0.510***</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td>[0.010]</td>
<td>[0.003]</td>
<td>[0.100]</td>
<td>[0.043]</td>
<td>[0.071]</td>
<td>[0.054]</td>
</tr>
<tr>
<td>Observations</td>
<td>1602</td>
<td>3564</td>
<td>1202</td>
<td>2632</td>
<td>2312</td>
<td>7969</td>
</tr>
<tr>
<td>Countries</td>
<td>13</td>
<td>28</td>
<td>13</td>
<td>27</td>
<td>17</td>
<td>62</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.36</td>
<td>0.63</td>
<td>0.41</td>
<td>0.66</td>
<td>0.36</td>
<td>0.65</td>
</tr>
</tbody>
</table>

1. Dependent variable is the number of export markets entered. All models contain fixed effects by country-sector. All variables are in logarithms.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%).

Table 5: Fixed effects Tobit regression results for number of export markets vs. firm productivity.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Alternative</th>
<th>Base</th>
<th>Alternative</th>
<th>Base</th>
<th>Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tfp^{OLS}$</td>
<td>0.02</td>
<td>0.003</td>
<td>[0.029]</td>
<td>[0.036]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$tfp^{LP}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.091***</td>
<td>0.125***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.022]</td>
<td>[0.027]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales/Empl.</td>
<td>0.023**</td>
<td>0.096***</td>
<td>[0.011]</td>
<td>[0.010]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.746***</td>
<td>-5.14</td>
<td>1.060***</td>
<td>-4.993</td>
<td>0.573***</td>
<td>-5.395</td>
</tr>
<tr>
<td></td>
<td>[0.124]</td>
<td>[0.000]</td>
<td>[0.133]</td>
<td>[0.000]</td>
<td>[0.126]</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Observations</td>
<td>1602</td>
<td>3564</td>
<td>1202</td>
<td>2632</td>
<td>2312</td>
<td>7969</td>
</tr>
</tbody>
</table>

1. Dependent variable is the number of export markets entered. All models contain fixed effects by country-sector. All variables are in logarithms.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%).
Table 6: Random effects Tobit regression results for number of export markets vs. firm productivity.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Alternative</th>
<th>Base</th>
<th>Alternative</th>
<th>Base</th>
<th>Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>tfp\textsuperscript{OLS}</td>
<td>0.027</td>
<td>-0.038</td>
<td>[0.027]</td>
<td>[0.038]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tfp\textsuperscript{LP}</td>
<td>0.053</td>
<td>0.054</td>
<td>[0.039]</td>
<td>[0.046]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales/Empl.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.004</td>
<td>0.093***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.013]</td>
<td>[0.021]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.232</td>
<td>-0.454**</td>
<td>0.22</td>
<td>-0.716**</td>
<td>0.264</td>
<td>-1.548***</td>
</tr>
<tr>
<td></td>
<td>[0.155]</td>
<td>[0.209]</td>
<td>[0.151]</td>
<td>[0.284]</td>
<td>[0.171]</td>
<td>[0.232]</td>
</tr>
<tr>
<td>Observations</td>
<td>1602</td>
<td>3564</td>
<td>1202</td>
<td>2632</td>
<td>2312</td>
<td>7969</td>
</tr>
</tbody>
</table>

1. Dependent variable is the number of export markets entered. All models contain random effects by country-sector. All variables are in logarithms.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%).
Table 7: Baseline regression results and robustness checks for trade flow size and sample composition.

| Variable          | Base Entry Cost | Base Export Cost | Base Distance | Base Tariffs | Base GDP | Base GDPPC | Base TFP | Base Sales/Emp. | Base Cons. | No UMI Entry Cost | No UMI Export Cost | No UMI Distance | No UMI Tariffs | No UMI GDP | No UMI GDPPC | No UMI TFP | No UMI Sales/Emp. | No UMI Cons. | No LMI Entry Cost | No LMI Export Cost | No LMI Distance | No LMI Tariffs | No LMI GDP | No LMI GDPPC | No LMI TFP | No LMI Sales/Emp. | No LMI Cons. |
|-------------------|----------------|------------------|--------------|-------------|---------|-----------|---------|----------------|-----------|------------------|------------------|----------------|-------------|-------------|---------|-------------|---------|----------------|------------|------------------|------------------|----------------|-------------|---------|-------------|---------|----------------|------------|
|                   |                |                  |              |             |         |           |         |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| Entry Cost        | -0.06          | -0.038           | -0.039       | -0.068      | -0.087  | -0.099*   | -0.12  |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   | [0.045]        | [0.075]          | [0.044]      | [0.058]     | [0.073] | [0.053]   | [0.082] |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| Export Cost       | -0.163**       | 0.047            | -0.148*      | -0.228**    | -0.238**| -0.098    | -0.173 |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   | [0.073]        | [0.117]          | [0.083]      | [0.091]     | [0.096] | [0.101]   | [0.160] |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| Distance          | -0.316*        | -0.092           | -0.299       | -0.293      | -0.143  | 0.046     | 0.039  |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   | [0.181]        | [0.272]          | [0.201]      | [0.251]     | [0.314] | [0.275]   | [0.621] |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| Tariffs           | -0.269         | 0.195            | 0.43         | -1.356**    | -1.664* | -0.361    | -1.324* |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   | [0.321]        | [1.803]          | [0.826]      | [0.687]     | [1.001] | [0.431]   | [0.795] |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| GDP               | 0.348***       | 0.336***         | 0.306***     | 0.494***    | 0.585*** | 0.363***  | 0.436*** |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   | [0.020]        | [0.029]          | [0.020]      | [0.028]     | [0.032] | [0.027]   | [0.040] |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| GDPPC             | 0.155***       | 0.222***         | 0.270***     | 0.107       | -0.057  | -0.025    | -0.082  |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   | [0.047]        | [0.071]          | [0.085]      | [0.082]     | [0.212] | [0.212]   | [0.212] |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| TFP               |                | 0.027            |              |            |         |          |        |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   |                | [0.017]          |              |            |         |          |        |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| Sales/Emp.        |                |                  |              | 0.017*     |         |          |        |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   |                |                  |              | [0.017]    |         |          |        |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| Cons.             | -2.049         | -3.721           | 0.667        | -7.961***   | -13.213***| -5.403***| -4.909  |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
|                   | [1.824]        | [2.805]          | [2.005]      | [2.593]     | [3.264] | [2.520]   | [6.125] |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |
| Obs.              | 11808          | 682              | 2372         | 11808      | 11808   | 8736      | 4128    |                |           |                  |                  |                |             |             |         |               |         |                |           |                  |                |                |             |         |               |         |                |           |

1. Dependent variable is the number of export markets entered. All independent variables are in logarithms. All models are estimated by Poisson, and contain fixed effects by HS 2 digit sector. Columns four and five exclude 6-digit trade flows under $100,000 and $1m respectively. The first five columns exclude high income countries. The last two exclude in addition upper-middle and lower-middle income countries.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%).
### Table 8: Regression results using instrumental variables.

<table>
<thead>
<tr>
<th>Lags</th>
<th>Instrumental Variables</th>
<th>Instrumental Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2nd Stage</td>
<td>Entry Cost</td>
</tr>
<tr>
<td>Entry Cost</td>
<td>-0.099**</td>
<td>-0.392***</td>
</tr>
<tr>
<td></td>
<td>[0.044]</td>
<td>[0.090]</td>
</tr>
<tr>
<td>Export Cost</td>
<td>-0.193***</td>
<td>-0.632***</td>
</tr>
<tr>
<td></td>
<td>[0.069]</td>
<td>[0.123]</td>
</tr>
<tr>
<td>Distance</td>
<td>-0.473**</td>
<td>-0.691***</td>
</tr>
<tr>
<td></td>
<td>[0.195]</td>
<td>[0.192]</td>
</tr>
<tr>
<td>Tariffs</td>
<td>-0.185</td>
<td>-0.278</td>
</tr>
<tr>
<td></td>
<td>[0.288]</td>
<td>[0.312]</td>
</tr>
<tr>
<td>GDP00</td>
<td>0.364***</td>
<td>0.295***</td>
</tr>
<tr>
<td></td>
<td>[0.020]</td>
<td>[0.022]</td>
</tr>
<tr>
<td>GDPPC00</td>
<td>0.090**</td>
<td>0.186***</td>
</tr>
<tr>
<td></td>
<td>[0.039]</td>
<td>[0.044]</td>
</tr>
<tr>
<td>Lang_Eng.</td>
<td>0.044</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>[0.210]</td>
<td>[0.132]</td>
</tr>
<tr>
<td>Lang_Fra.</td>
<td>0.680***</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>[0.208]</td>
<td>[0.139]</td>
</tr>
<tr>
<td>Lang_Esp.</td>
<td>0.974***</td>
<td>0.174</td>
</tr>
<tr>
<td></td>
<td>[0.242]</td>
<td>[0.166]</td>
</tr>
<tr>
<td>Lang_Prt.</td>
<td>0.821</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>[0.573]</td>
<td>[0.236]</td>
</tr>
<tr>
<td>Lang_Rus.</td>
<td>-0.515***</td>
<td>0.964***</td>
</tr>
<tr>
<td></td>
<td>[0.150]</td>
<td>[0.079]</td>
</tr>
<tr>
<td>Col_UK</td>
<td>0.155</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>[0.242]</td>
<td>[0.125]</td>
</tr>
<tr>
<td>Col_Fra.</td>
<td>0.562**</td>
<td>-0.106</td>
</tr>
<tr>
<td></td>
<td>[0.238]</td>
<td>[0.148]</td>
</tr>
<tr>
<td>Col_Esp.</td>
<td>1.037***</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>[0.295]</td>
<td>[0.183]</td>
</tr>
<tr>
<td>Col_Prt.</td>
<td>0.754</td>
<td>-0.077</td>
</tr>
<tr>
<td></td>
<td>[0.574]</td>
<td>[0.242]</td>
</tr>
<tr>
<td>Col_Rus.</td>
<td>-0.892***</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>[0.241]</td>
<td>[0.184]</td>
</tr>
<tr>
<td>Cons.</td>
<td>0.155</td>
<td>7.975***</td>
</tr>
<tr>
<td></td>
<td>[1.966]</td>
<td>[2.489]</td>
</tr>
<tr>
<td>Obs.</td>
<td>12576</td>
<td>11808</td>
</tr>
<tr>
<td>Instr. F/χ²</td>
<td>24.93***</td>
<td>19.72***</td>
</tr>
<tr>
<td></td>
<td>[8.803]</td>
<td>[5.101]</td>
</tr>
<tr>
<td>R²</td>
<td>0.32</td>
<td>0.21</td>
</tr>
</tbody>
</table>
1. Dependent variable for the main regression is the number of export markets entered. All independent variables are in logarithms. All models are estimated by Poisson, and contain fixed effects by HS 2 digit sector.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%). In the first-stage regressions, Instr. F tests the null hypothesis that the coefficients on the instruments are jointly equal to zero. In the second-stage regressions, Instr. $\chi^2$ tests the null hypothesis that the coefficients on the first stage residuals are jointly equal to zero.
### Table 9: Regression results in first differences (relative to the USA).

<table>
<thead>
<tr>
<th></th>
<th>Poisson 2nd Stage</th>
<th>IV-Poisson 2nd Stage</th>
<th>IV-TSLS 2nd Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entry Cost</td>
<td>Export Cost</td>
<td>Entry Cost</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry Cost</td>
<td>-0.095**</td>
<td>-0.398***</td>
<td>-0.274***</td>
</tr>
<tr>
<td></td>
<td>[0.048]</td>
<td>[0.092]</td>
<td>[0.083]</td>
</tr>
<tr>
<td>Export Cost</td>
<td>-0.194***</td>
<td>-0.634***</td>
<td>-0.410*</td>
</tr>
<tr>
<td></td>
<td>[0.069]</td>
<td>[0.123]</td>
<td>[0.242]</td>
</tr>
<tr>
<td>Distance</td>
<td>-0.448**</td>
<td>-0.670***</td>
<td>-0.609</td>
</tr>
<tr>
<td></td>
<td>[0.198]</td>
<td>[0.193]</td>
<td>[0.193]</td>
</tr>
<tr>
<td>Tariffs</td>
<td>0.317***</td>
<td>0.257**</td>
<td>0.471***</td>
</tr>
<tr>
<td></td>
<td>[0.104]</td>
<td>[0.113]</td>
<td>[0.163]</td>
</tr>
<tr>
<td>Tariffs</td>
<td>-0.090**</td>
<td>-0.061***</td>
<td>0.377***</td>
</tr>
<tr>
<td></td>
<td>[0.021]</td>
<td>[0.041]</td>
<td>[0.028]</td>
</tr>
<tr>
<td>GDP00</td>
<td>0.365***</td>
<td>0.300***</td>
<td>-0.010**</td>
</tr>
<tr>
<td></td>
<td>[0.042]</td>
<td>[0.045]</td>
<td>[0.045]</td>
</tr>
<tr>
<td>Lang_Eng.</td>
<td>0.044</td>
<td>0.073</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td>[0.209]</td>
<td>[0.131]</td>
<td>[0.212]</td>
</tr>
<tr>
<td>Lang_Fra.</td>
<td>0.678***</td>
<td>0.035</td>
<td>0.794***</td>
</tr>
<tr>
<td></td>
<td>[0.027]</td>
<td>[0.139]</td>
<td>[0.235]</td>
</tr>
<tr>
<td>Lang_Esp.</td>
<td>0.974***</td>
<td>0.174</td>
<td>1.191***</td>
</tr>
<tr>
<td></td>
<td>[0.241]</td>
<td>[0.165]</td>
<td>[0.339]</td>
</tr>
<tr>
<td>Lang_Prt.</td>
<td>0.821</td>
<td>0.04</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>[0.571]</td>
<td>[0.235]</td>
<td>[0.531]</td>
</tr>
<tr>
<td>Lang_Rus.</td>
<td>-0.517***</td>
<td>0.965***</td>
<td>-0.447***</td>
</tr>
<tr>
<td></td>
<td>[0.149]</td>
<td>[0.079]</td>
<td>[0.155]</td>
</tr>
<tr>
<td>Cons.</td>
<td>1.325***</td>
<td>1.904***</td>
<td>0.758*</td>
</tr>
<tr>
<td></td>
<td>[0.160]</td>
<td>[0.187]</td>
<td>[0.413]</td>
</tr>
<tr>
<td>Obs.</td>
<td>11808</td>
<td>11808</td>
<td>11808</td>
</tr>
<tr>
<td>Instr. $F/\chi^2$</td>
<td>24.95***</td>
<td>19.9***</td>
<td>33.86***</td>
</tr>
<tr>
<td>Overid $\chi^2$</td>
<td></td>
<td></td>
<td>6.097</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.32</td>
<td>0.21</td>
<td>0.5024</td>
</tr>
</tbody>
</table>

1. Dependent variable for the main regression is the number of export markets entered relative to the USA. All independent variables are in logarithms, and the dependent variable is also in logarithms in model 3. The first two models are estimated by Poisson, the third by TSLS. All models contain fixed effects by HS 2 digit sector.

2. Robust standard errors adjusted for clustering by country are in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%). In the first-stage regressions, Instr. F tests the null hypothesis that the coefficients on the instruments are jointly equal to zero. In the second-stage regressions, Instr. $\chi^2$ tests the null hypothesis that the coefficients on the
first stage residuals are jointly equal to zero. Overid $\chi^2$ tests the null hypothesis that the instruments are uncorrelated with the residuals from the second-stage regression.
Chapter III

Product Standards, Harmonization, and Trade: Evidence from the Extensive Margin
III  Product Standards, Harmonization, and Trade: Evidence from the Extensive Margin

Abstract: I use a new database of EU product standards in the textiles, clothing, and footwear sectors to present the first empirical evidence that international standards harmonization is associated with increased partner country export variety. A 10 percentage point increase in the proportion of internationally harmonized standards is associated with a 0.2% increase in partner country export variety, whereas a 10% increase in the total number of standards is associated with a nearly 6% decrease in product variety. Although small, the harmonization elasticity is statistically significant, and proves highly robust to sample changes and instrumental variables estimation using instruments motivated by political economy considerations. Moreover, it is found to be around 50% higher for low income countries, which suggests that they may be particularly constrained in adapting products to meet multiple standards. Numerical simulations show that these findings are consistent with a heterogeneous firms model of trade in which harmonization is beneficial at the extensive margin provided that any increases in compliance costs are not too large.

JEL codes: F13; F15.

Keywords: International trade; Product variety; Standards; Heterogeneous firms.

A  Introduction

By increasing the fixed product adaptation costs that firms must pay in order to access foreign markets, "regulatory protectionism" (Baldwin, 2000) has the potential to impact trade at the extensive (or "new product") margin: higher fixed costs discourage export market entry, and reduce the range of product varieties exported.\footnote{I am grateful to Silja Baller, Bernard Hoekman, and John S. Wilson for their comments and suggestions in relation to this Chapter.} This effect could be particularly strong in poor countries,
where lack of access to information, technology, managerial capacity and finance can impede the ability of firms to adapt production processes quickly and adequately to meet product standards in rich country markets.

It has been suggested in the literature (Collier and Venables, 2007) that trade preferences might be one way of promoting export diversification, i.e. an expansion in the range of product varieties exported by developing countries. But such schemes focus almost exclusively on tariffs, and neglect the costs imposed by non-tariff measures (including product standards), not to mention the important role played by supply side constraints (see Hoekman, 2007, for a review). The costs imposed by NTMs are significant, however: Kee et al. (2006) estimate that they add 70% to the level of restrictiveness imposed by tariffs alone. Complementary policies to address these costs would therefore be an important part of any renewed focus on preferences as a possible means of promoting an increase in developing country export variety.

One complementary policy that deserves further attention is international harmonization of product standards. This paper provides empirical support for the view that international harmonization can limit the scope for regulatory protectionism, and thereby attenuate the possible negative impacts of product standards on developing country exporters. Its main novelty is its focus on the extensive margin of trade. The intuition behind this positive impact of harmonization on the extensive margin is simple: by providing access to all harmonizing markets upon compliance with a single standard—and payment of a single fixed market entry cost—harmonization not only reduces compliance costs for harmonizing countries, but also eliminates the multiple costs that would otherwise face exporters in non-harmonizing countries. Numerical simulations using a three-country heterogeneous firms model show that the extensive margin impacts of harmonization should be positive for both harmonizing and non-harmonizing countries, even if the harmonized standard results in a moderate increase in compliance costs.
I find strong support for these predictions using a new World Bank database of EU product standards in the textiles, clothing, and footwear sectors (Czubala et al., 2007). These standards are issued by the European Normalization Committee (CEN). Although compliance is voluntary as a matter of law (as is the case for the standards considered by Swann et al., 1996, and Moenius, 2004), there is potential for these standards to have significant economic impacts: as at the end of 2006, CEN had issued 12,357 standards and approved documents and had another 3,510 in preparation. Results in this paper should therefore be seen as complementing previous work (see below) that focuses on EU-wide harmonization of mandatory standards, through EC Directives.

Standards data are related to export variety of non-EU countries using measures built up from highly detailed, 8-digit mirror (import) data. Results show that the total count of standards in a given sector is negatively associated with the range of varieties exported by partner countries: the elasticity of product variety with respect to the number of standards is 0.6 in the preferred specification. A higher proportion of internationally harmonized standards in the total–proxied by the number of standards which are identical or equivalent to ISO standards–is found to be positively associated with export variety. Moreover, I find evidence that the impact of international harmonization is stronger for low income exporters: the elasticity of export variety with respect to the proportion of harmonized standards is 0.03, compared with 0.02 for the full sample. Although quantitatively small, these elasticities are statistically significant at the 1% level.

To be sure that I am identifying genuinely exogenous variation in the standards measures, I adopt a novel instrumental variables strategy based on the two-tier political economy framework within which trade policy and standards function in the EU. The argument is that the EU-wide size of each sector (proxied by total value added and number of employees) is an indicator of the industry’s resources with which to lobby governments. Since harmonization can reduce the costs faced by overseas competitors, incumbents have an incentive to lobby against it, as a way of
raising their rivals’ costs (see e.g., Salop and Scheffman, 1983). In addition, Herfindahl indices of employment and value added are used as indicators of the propensity of individual EU member states to overcome their collective action problem, and lobby jointly against harmonization in order to protect "their" industry. These measures have explanatory power for the potentially endogenous variables, and standard statistical tests show that they are exogenous to partner country export variety. Results prove to be highly robust when the model is re-estimated by two stage least squares using these data as instruments.

These results make two main contributions to the literature. First, they expand the scope of recent empirical work on trade growth at the extensive margin by examining the effects of harmonization. Although the need to adapt products and production processes to meet foreign standards is often used as a motivation for including fixed market entry costs in models of international trade, this appears to be the first paper to make an empirical link between standards harmonization and extensive margin growth. Hummels and Klenow (2005) show that bigger, richer countries tend to export a wider range of products, as do those which face lower international transport costs. The positive extensive margin impacts of tariff liberalization and preferences are examined by Kehoe and Ruhl (2003), and Feenstra and Kee (2007).

This paper also expands the existing literature on the trade impacts of product standards both through a focus on extensive margin growth, and by an explicit examination of the partner country impacts of international harmonization. Moenius (2004) uses a gravity model to show that bilaterally shared standards—and in some cases specific national standards—can promote trade. However, he does not differentiate between standards which are internationally (not just bilaterally) harmonized and those which are not. Swann et al. (1996) do make such a distinction, finding evidence that UK national standards are associated with higher levels of both exports and imports, but that the corresponding effects are much weaker in the case of internationally harmonized standards.
Their results are difficult to interpret, however, since their empirical model does not include factors such as market size (GDP) and trade barriers (tariffs).

To my knowledge, the only recent empirical papers to deal with third-country impacts of harmonization are Chen and Mattoo (2004) and Baller (2007). Chen and Mattoo (2004) use a sample selection gravity model (cf. Helpman et al., Forthcoming) to examine the impacts of EU Harmonization Directives and Mutual Recognition Agreements on intra- and extra-European trade. These effects are captured using dummy variables. Baller (2007) adopts the same approach using data on both EU and ASEAN harmonization and mutual recognition agreements. Empirical results generally suggest that harmonization boosts trade among harmonizing countries, as well as imports from third countries. The present paper builds on and extends these results by focusing on product variety, which is not the object of analysis in either of these two previous papers.

The plan of the paper is as follows. The next section provides some theoretical motivation for the empirical hypotheses to be tested later on, using a heterogeneous firms model of trade. I provide an explicit definition of harmonization within this framework, and use simulations to examine its impacts on export product variety in harmonizing and non-harmonizing countries. Section 3 describes the dataset, a new World Bank database of EU product standards, as well as measures of export variety covering up to 200 countries for the period 1995-2003. The empirical model is presented in section 4, along with estimation results, robustness checks, and details of instrumental variables estimates. Section 5 concludes, and offers some suggestions for further research in this area.

B Theoretical Motivation

This section develops a simple theoretical framework that is used to produce predictions which the remainder of the paper tests empirically. Since the model is essentially a three-country version
of Melitz (2003), it is presented in outline only.\textsuperscript{55} The key feature of this model is the inclusion of fixed market entry costs that are allowed to vary bilaterally. I motivate the existence of such costs by relating them to the need to modify products and production processes in order to meet foreign standards.\textsuperscript{56} The available firm-level evidence suggests that foreign standards can indeed impose substantial fixed costs of compliance: Maskus et al. (2005) report an average of $425,000 per firm, or 4.7\% of value added, based on a survey of over 600 firms in 16 developing countries. For firms in the textiles and clothing sectors, the mean investment cost required to comply with technical requirements is 2.7\% of firm sales. The range is very wide, however, running from 0.01\% to 44.1\%. This suggests that the firm-level impacts of standards can vary drastically, thereby making a heterogeneous firms approach particularly attractive.\textsuperscript{57}

After setting out a baseline scenario without harmonization, I modify the model to take account of product standard harmonization covering two of the three countries only. Harmonization lowers the fixed costs of exporting from one harmonizing country to another such that they are at the same level as the costs of domestic market entry. This reflects the fact that under harmonization, domestic producers in both harmonizing countries produce goods satisfying the same standard, and they can therefore export them inside the harmonization zone without paying any further adaptation costs. From the perspective of the third (non-harmonizing) country, harmonization means that it is only necessary to pay one fixed cost (not two) in order to access both foreign markets.

Using numerical simulations, I show that harmonization thus defined raises the export marginal

\textsuperscript{55}The presentation here is closest to the version of Melitz (2003) that appears in Helpman et al. (2004) or Baldwin and Harrigan (2007). Baller (2007) also applies a heterogeneous firms framework to the analysis of product standards.

\textsuperscript{56}The model does not directly address the distinction drawn in the policy literature between mandatory and voluntary standards. The fixed costs modeled here can be seen as related to standards that are either a commercial or legal necessity for export sales. Note that I do not consider the role that standards can play on the consumption side (Fischer and Serra, 2000; Ganslandt and Markusen, 2001), nor reductions in information costs for exporters (Moenius, 2004).

\textsuperscript{57}There is also an emerging body of case study evidence on this point. World Bank (2005) shows that management capacity and strategic decisions are an important determinant of how developing country firms react to changes in foreign product standards. The logic of the Melitz (2003) model presented here suggests that the cost increases associated with such changes will be felt most acutely by relatively unproductive firms, which may drop out of markets as a result.
cost cutoff in the two harmonizing countries vis-à-vis each other, and also in the non-harmonizing country. Foreign harmonization therefore makes it easier for producers in the third country to overcome the cost hurdles associated with exporting, which in turn results in increased export variety due to the association of each firm with a distinct product variety. Although the level of compliance costs imposed by the harmonized standard as compared with pre-existing national standards has the potential to limit or even reverse this outcome, simulations show that the third-country impact is positive even when compliance with the harmonized standard is moderately more burdensome.

1 Consumption Block

The world consists of three regions: Home (H), Foreign (F), and the Rest of the World (R). Labor is the only factor of production, and the regions are endowed with $L$ units each. Each region has two productive sectors. One produces a freely traded homogeneous good under constant returns to scale with one unit of labor required for one unit of output. Wages are therefore equal to unity in equilibrium. The other sector produces a continuum of differentiated goods under increasing returns to scale and costly trade. Absolute specialization in any sector is excluded.

Identical consumers in all markets maximize the two-tier utility function (III.1). Their expenditure shares are $\beta$ for the differentiated sector and $(1 - \beta)$ for the homogeneous sector. The elasticity of substitution in the differentiated sector is $\sigma$, across the set of varieties $V$.

$$U = q^{(1-\beta)} \left( \int_{\nu \in V} x(\nu)^{1-\frac{1}{\sigma}} d\nu \right)^{\beta \left( \frac{\varepsilon}{\sigma - 1} \right)}$$  (III.1)
As is well known, a typical demand function \( x \) is:

\[
x \left[ p(\nu) \right] = \frac{\beta E \left[ p(\nu) \right]^{-\sigma}}{\int_{\nu \in V^i} p(\nu)^{1-\sigma} \, d\nu} \equiv d^i \left[ p(\nu) \right]^{-\sigma}
\]

where \( V^i \) is the set of varieties available in country \( i \in \{H, F, R\} \), \( E = L \) is total expenditure by that country’s consumers, and the summary parameter \( d^i \) is a demand shifter.

2 Production Block

As usual in Dixit-Stiglitz models, producers in the differentiated goods sector engage in constant markup pricing such that \( p(\nu) = \frac{\sigma}{\sigma - 1} c \). Firms in each country face a fixed startup cost \( f^i_d \) that must be paid in order to enter the domestic market. This cost reflects the investment required to establish a production process that manufactures goods which accord with local product standards. A typical firm’s domestic market profit function is therefore:

\[
\pi^i_d = d^i \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} - d^i \left( \frac{\sigma}{\sigma - 1} c \right)^{-\sigma} c - f^i_d \equiv \frac{d^i}{\sigma} \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} - f^i_d
\]

(III.3)

Setting this expression equal to zero and solving for \( c \) establishes three maximum marginal costs \( c^i_d \) above which it is not possible to profitably supply the domestic market in each country.

\[
c^i_d = \left( \frac{\sigma f^i_d}{d^i} \right)^{-\frac{1}{\sigma}} \left( \frac{\sigma - 1}{\sigma} \right), \quad (i \in \{H, F, R\})
\]

(III.4)

Firms wishing to export face an additional layer of costs over and above the cost of domestic market entry: they must pay a fixed cost to adapt their production process so as to produce goods
Development Perspectives on Trade Growth at the Extensive Margin

that comply with foreign product standards.\textsuperscript{58} Product standards are unique to each country, so fixed product adaptation costs must be paid cumulatively (i.e. one fixed cost per market entered). I use $f_{i j}^x$ to indicate the fixed market entry costs that must be paid when exporting from $i$ to $j$, and assume that compliance with foreign standards is in all cases more costly than compliance with domestic ones (i.e., $f_{i j}^x > f_d^i, f_d^j$). Thus, the additional profits from exporting are:

$$\pi_{i j}^x = d^l \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} - d^l \left( \frac{\sigma}{\sigma - 1} c \right)^{-\sigma} c - f_{i j}^x \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} - f_{i j}^x$$ (III.5)

and the maximum marginal cost $c_{i j}^x$ above which it is not possible to profitably export is:

$$c_{i j}^x = \left( \frac{\sigma f_{i j}^x}{d^l} \right)^{\frac{1}{1-\sigma}} \left( \frac{\sigma - 1}{\sigma} \right), \quad (i \in \{H, F, R\}, i \neq j)$$ (III.6)

I introduce marginal cost heterogeneity by assuming that $c$ follows an identical Pareto distribution with support $[0, \bar{c}]$ in all three countries. This setup implies a cumulative distribution function $G(c) = \left( \frac{c}{\bar{c}} \right)^k$ and probability density function $g(c) = \frac{dG(c)}{dc} = \frac{k \bar{c}^{k-1}}{c^k}$, where $k > \sigma - 1$ is a "shape" parameter that indexes firm heterogeneity.\textsuperscript{59}

A free entry condition closes the production block of the model. The expected profits from domestic and export market sales are equated with the fixed cost of entering the marginal cost

\textsuperscript{58}To sharpen the focus on the role of fixed adaptation or retooling costs, variable (iceberg) trade costs are set equal to unity. This assumption could be relaxed, for instance to study the impact of per unit conformity assessment and certification costs. However, the basic insights of the model would not change.

\textsuperscript{59}Assuming $k > \sigma - 1$ ensures convergence of the integrals in the free entry conditions below.
"lottery" $f_e$ (identical in all countries):

\[
\int_0^{c_{ij}} \left( \frac{d_i}{\sigma} \left( \frac{\sigma}{\sigma - 1} - f_d^i \right) \right) dG(c) + \int_0^{c_{ij}} \left( \frac{d_j}{\sigma} \left( \frac{\sigma}{\sigma - 1} - f_d^j \right) \right) dG(c) + \\
\int_0^{c_{ik}} \left( \frac{d_k}{\sigma} \left( \frac{\sigma}{\sigma - 1} - f_d^k \right) \right) dG(c) = f_e, \quad (i \in \{H, F, R\}; i \neq j \neq k) \quad (\text{III.7})
\]

The above framework gives three free entry conditions (III.7), three cutoff expressions for domestic market entry (III.4), and six additional cutoff expressions for export market entry (III.6). The model has the same number of unknowns, namely three demand shifters $d_i$, three domestic market cost cutoffs $c_{ij}^d$, and six export cost cutoffs $c_{ij}^x$. Equilibrium will see firms in each country self-selecting into four groups based on their marginal cost draws: high cost firms will exit immediately without producing for any of the three markets, those with slightly lower costs will produce for the domestic market only, and those with relatively low costs will in addition export to one or both other markets. The higher the export marginal cost cutoff for each bilateral trading relationship, the greater the proportion of active domestic firms that will be able to enter that export market. So for a given mass of firms, a higher threshold is associated with greater export variety since each firm produces a distinct variety of the differentiated good.

### 3 Modeling Product Standard Harmonization

International harmonization of $F$’s product standards–i.e., bringing them into line with those prevailing in the rest of the world ($R$)–means that all firms now have access to two markets upon payment of a single fixed cost.\(^6\) With the same standard in place in both markets, goods manu-

---

\(^6\)The model treats harmonization as a discrete, exogenous policy shock. For an examination of harmonization in a political economy setting, see Gandal and Shy (2001).

SHEPHERD, Benjamin - Thèse IEP de Paris - 2008

136
factured in \(F\) or \(R\) can be sold freely in the other country without any need for further adaptation. The fixed costs of exporting are eliminated between these two markets, leaving only the fixed costs of domestic market entry (identical in the two countries). The profit function for sales within the harmonization zone is therefore

\[
\pi^i = \frac{d^i}{\sigma} \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} + \frac{d^j}{\sigma} \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} - \frac{f_d^i}{\text{fixed costs}}, \quad (i, j \in \{F, R\}, i \neq j) \tag{III.8}
\]

which gives a combined domestic and export market cutoff for each harmonizing country:

\[
c_d^j = \left[ \frac{f_d^i}{\frac{1}{\sigma} \left( \frac{\sigma}{\sigma - 1} \right)^{1-\sigma} [d^i + d^j]} \right]^{\frac{1}{1-\sigma}}, \quad (i, j \in \{F, R\}, i \neq j) \tag{III.9}
\]

The export conditions from \(F\) and \(R\) to \(H\) remain unchanged from the pre-harmonization model, but those in the opposite direction require modification. After harmonization, producers in \(H\) still face a fixed cost of exporting to \(F\) or \(R\) due to the need to adapt their product to the harmonized standard (originally \(R\)’s standard). However, there is now only one additional standard they need meet in order to sell in both \(F\) and \(R\), as opposed to the two separate standards that initially prevailed. \(H\)’s two export market profit expressions therefore collapse into a single one:

\[
\pi^H = \frac{d^F}{\sigma} \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} + \frac{d^R}{\sigma} \left( \frac{\sigma}{\sigma - 1} c \right)^{1-\sigma} - \frac{f_d^H}{\text{fixed costs}} \tag{III.10}
\]
Development Perspectives on Trade Growth at the Extensive Margin

which leads to a new cutoff for $H$ covering exports to both $F$ and $R$:

$$c^H_x = \left[ \frac{f^H_x}{\frac{1}{\sigma} \left( \frac{\sigma}{\sigma-1} \right)^{1-\sigma} \left( d^F + d^R \right)} \right]^{\frac{1}{1-\sigma}} \quad (\text{III.11})$$

These changes to the profit conditions require corresponding changes to the free entry conditions for all countries. For $H$, the condition is now

$$\int_0^{c^H_x} \left( \frac{d^H}{\sigma} \left( \frac{\sigma}{\sigma-1} c \right)^{1-\sigma} - f^H_d \right) dG(c) + \int_0^{c^H_x} \left( \frac{d^F}{\sigma} \left( \frac{\sigma}{\sigma-1} c \right)^{1-\sigma} + \frac{d^R}{\sigma} \left( \frac{\sigma}{\sigma-1} c \right)^{1-\sigma} - f^H_x \right) dG(c) = f_e \quad (\text{III.12})$$

while for $F$ and $R$ the condition takes the following form:

$$\int_0^{c^F_j} \left( \frac{d^i}{\sigma} \left( \frac{\sigma}{\sigma-1} c \right)^{1-\sigma} + \frac{d^j}{\sigma} \left( \frac{\sigma}{\sigma-1} c \right)^{1-\sigma} - f^i_d \right) dG(c) + \int_0^{c^{iH}_x} \left( \frac{d^H}{\sigma} \left( \frac{\sigma}{\sigma-1} c \right)^{1-\sigma} - f^i_x \right) dG(c) = f_e, \quad (i, j \in \{F, R\}; i \neq j) \quad (\text{III.13})$$

The system to be solved now has nine equations: the three modified free entry conditions (III.12) and (III.13) together with the unchanged domestic cutoff for $H$ (III.4), the two unchanged export cutoffs for $F$ and $R$ to $H$ (III.6), and the three modified cutoffs (III.9) and (III.11). The number of unknowns is also equal to nine: three demand shifters $d^i$, three domestic market cutoffs $f^i_d$, and three export market cutoffs $c^H_x$, $c^{FH}_x$, and $c^{RH}_x$. 

SHEPHERD, Benjamin - Thèse IEP de Paris - 2008 138
4 Simulation Results

To gauge the impact of product standard harmonization in $F$ and $R$ on the extensive margin of trade, I compare numerical simulations of the pre- and post-harmonization models over a range of reasonable parameter values. This requires making some assumptions as to the relative and absolute values of the various fixed cost parameters. To fix ideas, I focus on the case of "harmonizing up": i.e., the initial standard in $R$ is more costly to comply with than the initial standards in $H$ and $F$, so bringing $F$’s standards into line with $R$’s entails an increase in restrictiveness. Intuitively, harmonizing up should be less favorable to product variety growth in third countries than "harmonizing down" to a less costly standard in $R$. These results can therefore be interpreted as putting a lower bound on harmonization’s impact.

Prior to harmonization, $f_{xH} = f_{xF} = 0.15$ (export costs into $H$ and $F$) and $f_{dH} = f_{dF} = 0.1$ (costs of domestic market entry in $H$ and $F$). To consider different levels of compliance costs in $R$, $f_{xR}$ (export costs into $R$) varies over the range 0.15 to 0.2, while $f_{dR}$ varies over the range 0.1 to 0.15. It is important to allow both of these parameters to vary, in order to reflect the fact that more costly standards in $R$ impact both foreign and domestic producers. Post harmonization, $f_{xH}$ remains unchanged but $f_{xH}$ (the new export threshold for firms in $H$) varies over the range 0.15 to 0.2, while $f_{dR} = f_{dF}$ varies over the range 0.1 to 0.15. For the values of invariant parameters, I follow Bernard et al. (2007).

Figure 1 shows simulated equilibrium export cutoffs for all three countries. Simulation results are expressed as percentage changes following harmonization. The first simulation point represents the limiting case in which $F$ and $R$ impose different but equally costly standards prior to harmonization. Moving from left to right across the figure shows changes in the impact of harmonization.

---

$^{61}$See Baldwin (2000) for examples of the many ways in which these parameters can be varied to capture aspects of different types of harmonization.

$^{62}$Like those authors, I assume: $\sigma = 3.8$, $k = 3.4$, $L = 1000$, $\beta = 0.5$, $\bar{c} = 0.2$, and $f^e = 2$. 

SHEPHERD, Benjamin - Thèse IEP de Paris - 2008
as the additional cost burden associated with $R$’s initial standard becomes larger.

At all points in the parameter space used here, harmonization raises the export cutoffs in all countries. This is the case even though the harmonized standard is moderately more costly to comply with than the unharmonized one in $F$ (the difference is around 33% at the upper boundary). The intuition behind this is that reduction in the multiplicity of fixed market entry costs dominates a moderate increase in compliance costs. However, a very burdensome standard in $R$ could entail a rise in compliance costs for $F$ that would dominate the reduction in multiplicity and produce a negative impact on $H$.

The largest harmonization effects are felt in the two harmonizing countries $F$ (dotted and dashed line) and $R$ (crossed line). The percentage jump in $F$’s effective export threshold to $R$ declines only slowly with increasing compliance costs under harmonization, but for $R$ this effect is much more rapid. This difference reflects the fact that from $R$’s point of view, harmonization leads to the imposition of a more burdensome standard in $F$; however, from $F$’s point of view, the standard prevailing in $R$ does not change.

Although the impact of harmonization on $H$’s export cutoff is positive vis-à-vis both foreign markets, the effect varies substantially by destination. The effective rise in the export threshold to $R$ remains approximately constant over the full range of compliance costs considered here. This again reflects the fact that $R$’s standard does not change due to harmonization. On the other hand, $F$’s standard does change, and this results in an export cutoff increase that declines sharply in the costs of compliance under harmonization: with equally burdensome standards in $F$ and $R$, harmonization lifts $H$’s export threshold by just under 13%, but when $R$’s compliance costs are 33% higher than $F$’s, the rise is limited to 4.6%.
C Data and Stylized Facts

The remainder of the paper tests the hypothesis that international harmonization is associated with an increase in partner country export variety in a way consistent with the mechanics set out above. This section presents two new data sources that play an important role in the empirical work: the World Bank’s EU Standards Database, and measures of export variety covering up to 200 countries based on highly-detailed, 8-digit mirror (import) data from Eurostat.

1 The EU Standards Database

Measuring the extent of standardization in EU product markets is not an easy business. Each member state sets both voluntary and mandatory standards on a national level, while centralized EU bodies also have the power to set standards with transnational application. Swann et al. (1996) and Moenius (2004) examine the trade impacts of voluntary national standards, while Chen and Mattoo (2004) and Baller (2007) focus on transnational mandatory standards (EC Directives). Only Czubala et al. (2007) look at the role played by transnational voluntary standards, such as those issued by the European Committee for Standardization (CEN).

CEN is a transnational association established in 1961 by national standards bodies from across Europe. Its standards must be adopted by all EU countries, and override any conflicting or inconsistent national standards. In addition to its work complementing EU Harmonization Directives, CEN is also active in independently developing standards in consultation with industry and national bodies. As noted above, CEN’s output to date is substantial: 12,357 standards and approved documents, with 3,510 more in preparation. By contrast, the European Commission has issued less than two dozen Harmonization Directives under its "New Approach". (See Pelkmans, 1987, for a review of the New Approach.)

For a general review of these mechanisms, see EC (2000).
The World Bank’s EU Standards Database (EUSDB) provides the first catalogue of CEN European standards in the agriculture, textiles, clothing, and footwear sectors, with mapping to a standard trade classification (HS 2000). This paper focuses exclusively on the three manufactured goods sectors.\textsuperscript{64} From a development point of view, these sectors are particularly important since they are associated with the early stages of industrialization in many countries.

For a full overview of EUSDB’s methodology, see Czubala et al. (2007). The general approach is similar to that of Swann et al. (1996) and Moenius (2000, 2004), although those authors both consider national standards in EU member states rather than CEN’s transnational standards.\textsuperscript{65} The primary information source for EUSDB is Perinorm (www.perinorm.com), a bibliographic database maintained jointly by the British, French, and German national standards bodies. It contains over 1.1 million records from 22 (mostly OECD) countries. Each record corresponds to a single national, regional, or international standard. For each standard, EUSDB contains data including the dates of entry into force and withdrawal, and a 1-0 dummy variable indicating whether or not it is "identical" or "equivalent" to an ISO standards. This variable is used as a proxy for de facto international harmonization. All information is cross-checked against CEN’s own on-line standards catalogue, before being manually mapped to the Harmonized System product classification (http://www.cen.eu/catweb/cwsen.htm).

For each 2- and 4-digit HS code, EUSDB provides a count of the number of CEN standards in force in a given year over the sample period (1995-2003).\textsuperscript{66} It also counts the number of those standards that entered into force prior to 1995 provided they were still in force at some point during the sample period. A standard is considered to be in force for a given year if it came into force before or during that year. If it is withdrawn at some point during the year, it is still assumed to be in force for the entire year. Amendments to existing standards are counted as additional standards.

\textsuperscript{64}I exclude agriculture because the measure of international harmonization recorded by EUSDB (equivalence with an ISO standard) is arguably less relevant to that sector. Standards promulgated by organizations such as the Codex Alimentarius are likely to be of greater importance.

\textsuperscript{65}An alternative approach is taken in recent papers by Fontagné et al. (2005) and Disdier et al. (Forthcoming). They use WTO notifications under the SPS and TBT Agreements to build databases of mandatory national standards. This is a promising approach, but one which currently suffers from the inconsistent reporting behavior of WTO Members.

\textsuperscript{66}Counts include standards that entered into force prior to 1995 provided they were still in force at some point during the sample period. A standard is considered to be in force for a given year if it came into force before or during that year. If it is withdrawn at some point during the year, it is still assumed to be in force for the entire year. Amendments to existing standards are counted as additional standards.
standards that are treated as being internationally harmonized using the above definition. Table 1 presents basic descriptive statistics, which disclose a number of notable features (see Czubala et al., 2007, for a complete discussion). All three sectors have undergone rapid growth in terms of the total number of standards in force. However, the bulk of standards remain concentrated in the textiles sector (84%). Although the proportion of internationally harmonized standards has generally risen, the pattern across sectors is by no means uniform. In the context of instrumental variables estimates below, some of the possible reasons for this non-uniformity will become apparent.

2 Measuring Export Variety

To examine the impact of standards and international harmonization on the extensive margin of trade, I construct new measures of export variety covering up to 200 countries. I follow the recent empirical literature on product variety in trade (e.g., Hummels and Klenow, 2005; and Broda and Weinstein, 2006), in building on the theory-consistent measure of variety developed by Feenstra (1994). I use the version of his measure set out by Feenstra and Kee (2006):

\[
\Lambda_{xst} = \frac{\sum_{l \in V_{s,t}^{w}} p_{l}^{w} d_{l}^{w}}{\sum_{l \in V_{s}^{w}} p_{l}^{w} d_{l}^{w}}\tag{III.14}
\]

The denominator is the total value of world exports in a particular sector, summing across all product varieties within that sector. Thus, \( V_{s}^{w} \) is the full set of varieties exported in sector \( s \), taking account of all exporting countries and all time periods. Average world trade values by product variety across all years \( (p_{l}^{w} d_{l}^{w}) \) are used to create the sum. While the denominator is invariant by exporter and time, the numerator is not. It consists of the sum of world average trade values in product varieties shipped by exporter \( x \) at time \( t \). The use of world average trade values ensures that variation in the numerator—and in \( \Lambda_{xst} \) itself—is due only to changes in \( x \)’s variety set. This measure
therefore has the important advantage of allowing consistent comparisons of product variety to be made across years and countries.

To implement this approach empirically, I use 8-digit import data from the European Union for the years 1995-2003. In line with availability of standards data, I calculate $\Lambda$ for three sectors: textiles (HS chapters 50-60), clothing (chapters 61-63), and footwear (chapter 64). Prior to calculation, I exclude from the dataset all observations relating to internal trade amongst EU-15 members, as well as product codes without verbal description which correspond to residual categories covering confidential or otherwise unclassified flows. For the world average trade value $\bar{p}^w_i q^w_i$, I take the average over the sample period of import values for the EU-15 (treated as a single entity).

Table 2 provides some basic descriptive statistics for this variety measure, broken down by sector and year. (Full results are available on request.) The median variety measure in the clothing and footwear sectors ($\Lambda_{med} = 0.2$ to 0.3) is noticeably higher than in textiles ($\Lambda_{med} \leq 0.1$). However, the range in each case is very wide, running from just a little above zero to 0.8 or 0.9. The fact that the median is so low within this range suggests that most countries export a relatively modest range of varieties in these three sectors, but that a few countries export a very wide range.

In terms of the rank ordering of countries by variety, results are broadly sensible: China, Turkey, India, and a number of countries in Central and Eastern Europe appear at the top of the list for clothing and footwear, while highly industrialized countries like Switzerland and the United States arrive in the lead for the more capital intensive textiles sector. The presence of the United States and Switzerland amongst the leading countries in clothing and footwear suggests that the trade data from Eurostat may be picking up some amount of re-exports or processing trade. I cannot be sure as to the exact extent of this issue, but I come back to it below in the context of implementing robustness checks for the empirical model.

67These data are freely available through the Eurostat website (http://fd.comext.eurostat.cec.eu.int/xtweb/).
D Empirical Model and Estimation Results

The theoretical model suggests that international harmonization of EU standards in the textiles, clothing, and footwear sectors should be positively associated with partner country export variety, provided that any additional compliance costs imposed by harmonized standards are not too high.\textsuperscript{68} To test this hypothesis empirically, I use EUSDB to calculate the number of internationally harmonized standards \( (iso_{st}) \) as a percentage of the total number of EU standards \( (stds_{st}) \) for each sector-year. I use this measure \( \frac{iso_{st}}{stds_{st}} \) as a proxy for the degree of international harmonization. To take account of the overall cost burden imposed by EU standards, I include \( \ln \frac{stds_{st}}{iso_{st}} \) as a separate regressor.

In addition to these factors, the theoretical model and past empirical work suggest that bigger markets (proxied by exporter GDP), more developed exporter technology (proxied by per capita GDP), and lower variable trade costs (tariffs \( \tau_{x_{st}} \)) should also impact product variety positively.\textsuperscript{69} I use the estimating equation (III.15) to test these predictions, taking advantage of the availability of panel data to include fixed effects that control for unobserved heterogeneity in the exporter, sector, and year dimensions.\textsuperscript{70} These fixed effects absorb influences such as EU GDP (which varies by year only), sectoral expenditure shares (on the assumption that they are reasonably stable over time), exporter resource endowments, and the type of international transport costs usually proxied by great circle distance. To take account of the possible influence of quota arrangements under the Agreement on Textiles and Clothing (ATC) and associated bilateral agreements between the EU and some non-WTO members, I include a dummy variable \( (quota) \) equal to unity only for

\textsuperscript{68}Due to lack of data, I do not consider intersectoral linkages or the possibility of cumulation of standards as goods pass from one processing stage to another.

\textsuperscript{69}I source GDP data (aggregate and per capita) in constant 2000 US dollars from the World Development Indicators. Simple average tariff rates (effectively applied) come from the UNCTAD-TRAiNS database accessed via WITS.

\textsuperscript{70}I do not convert the \( iso_{st}/stds_{st} \) term to logarithms, since it is already a percentage. \( \beta_2/100 \) is therefore the elasticity of export variety with respect to international harmonization.
countries subject to such quotas. Two other dummies (atc2 and atc3) are equal to unity for years 1998 onwards and 2002 onwards, in order to capture the effects of quota liberalization under phases 2 and 3 of ATC implementation. I also interact these dummies with quota to allow for differences in impact across quota and non-quota countries.

\[
\ln (\Lambda_{xst}) = \beta_1 \ln (stds_{st}) + \beta_2 \frac{isos_{st}}{stds_{st}} + \beta_3 \ln (1 + \tau_{xst}) + \beta_4 \ln (gdpt_{xt}) + \beta_5 \ln (gdppc_{xt}) + \\
\beta_6 atc2_{st} + \beta_7 atc3_{st} + \beta_8 quota_{xs} + \beta_9 quota_{xs} \times atc2_{st} + \beta_{10} quota_{xs} \times atc3_{st} + \varepsilon_{xst} \quad (III.15)
\]

Baseline results using OLS appear in the first column of Table 3. The overall fit of the model is good, with \( R^2 = 0.68 \). All of the substantive coefficients carry the expected signs: GDP and per capita GDP are both positive, while tariffs are negative. This is in line with previous work, although these coefficients are not statistically significant in this case. Countries subject to ATC quotas tend to export a narrower range of varieties, and this effect is statistically significant at the 1% level. The impacts from ATC liberalization are as expected in terms of sign (positive for quota countries, negative for the others), but are not statistically significant.

In line with the model’s predictions, the total standards count enters with a negative coefficient, while the extent of international harmonization is strongly positive. Both are statistically significant at the 1% level. In terms of magnitude, a 1% increase in the total number of EU standards is associated with a 0.8% decrease in export variety, while an increase of one percentage point in

---

71 Countries considered to be subject to quotas over the 1995-2003 period are: Argentina, Belarus, Bosnia and Herzegovina, Brazil, China, Hong Kong China, India, Indonesia, Macao China, Malaysia, North Korea, Pakistan, Peru, the Philippines, Serbia and Montenegro, Singapore, South Korea, Sri Lanka, Taiwan China, Thailand, and Vietnam. (Source: WTO Trade Policy Reviews of the European Union, various years.)

72 I assume that ATC quotas apply only to the textiles and clothing sectors as defined here, and not to footwear. This is basically consistent with the product list in the Annex to the ATC, which has extensive coverage in HS chapters 50-63, but lists only three 6-digit product lines in Chapter 64.

73 Since the two standards measures vary in the sector-year dimension, the robust standard errors are adjusted for clustering at that level.
the proportion of those standards which are internationally harmonized is associated with a 0.02% increase in export product variety. While this latter effect is quite small, I would argue that it could nonetheless be significant in an economic sense. This is particularly true for the sectors under consideration here, which are of considerable importance to developing countries in the early stages of industrialization. These results highlight, however, that in terms of assessing the impact on exporting countries, harmonization needs to be considered within the context of the overall level of compliance costs involved.

1 Instrumental Variables Results

It is important to be sure that these results are not biased due to the possible endogeneity of product standards (and tariffs) with respect to export variety.\textsuperscript{74} It is commonplace to view endogenous tariffs as the outcome of a lobbying process in which a government balances support from lobbies seeking protection against the national welfare costs of imposing that protection (e.g., Grossman and Helpman, 1994). Since harmonization can, like tariff liberalization, lower the cost barriers facing foreign firms, local incumbents have an interest in lobbying against it. For the moment, the political economy of this process remains to be examined in detail in the literature.\textsuperscript{75} I draw on the general framework that has emerged from the endogenous tariffs literature to identify some likely determinants of endogenous standards, which I then use as instruments.

In the EU context, the type of firm-to-government lobbying envisaged by Grossman and Helpman (1994) is accompanied by an additional layer of government-to-government lobbying within centralized EU bodies.\textsuperscript{76} This is due to the fact that both trade policy and the EU standards under

\textsuperscript{74}Moennis (2004) addresses endogeneity by using five year lags of his standards variables as instruments for their current levels. Chen and Mattoo (2004) use harmonization of closely related sectors as an instrument for harmonization of a given sector. In both cases, instrumental variables estimation does not substantially change the results.

\textsuperscript{75}Essaji (2005) is the only empirical study dealing with the political economy of standards in a trade setting.

\textsuperscript{76}Balahoing and Francois (2006) report that after controlling for the size of an industry, its "nationality" remains an important determinant of the level of protection it receives through the EU common external tariff.
consideration here are not decided unilaterally by national governments, but by European bodies (the Commission and CEN). To capture the lobbying potential of industries on an EU-wide basis, I use total sector value added and employment. The first of these measures proxies a sector’s political "muscle", i.e. the potential lobbying resources at its disposal. The second measure proxies a different dimension of sectoral politics, namely the direct voting power of an industry’s workers. In addition, I capture the lobbying potential of national governments using Herfindahl indices of value added and employment across EU member states. These measures proxy the extent to which each sector is geographically concentrated in particular countries, and thus the extent to which governments have a perceived economic or political interest in protecting "their" industry.

Treating all independent variables in (III.15) except standards and tariffs as exogenous, I re-estimate the equation using two stage least squares (TSLS). Second stage results are in the second column of Table 3. Rejection of the null by a Hausman test (1%) indicates that endogeneity is indeed a problem in these data, but correcting for it leaves the basic conclusions unchanged: the coefficient on the total number of standards remains negative, while the percentage of harmonized standards still enters the equation positively. The magnitudes of both coefficients change slightly—becoming larger for the percentage of harmonized standards, and smaller (in absolute value) for the total number of standards—but remain statistically significant at the 5% level (total number of standards) and the 1% level (proportion of harmonized standards). The tariff coefficient increases

77 An alternative approach could be to instrument using data on antidumping actions taken by each industry, using the dataset compiled by Bown (2007). This is a promising avenue for future research, but it is not pursued here for two reasons. First, the instruments used in this paper are arguably more likely to be exogenous to partner country export variety, and to satisfy the second stage exclusion restriction. Second, it is desirable for the final model to be overidentified, but it would be difficult to achieve this using antidumping data since only a small number of independent data series would be available.

78 These data come from Eurostat annual enterprise statistics on industry and construction, freely downloadable from http://epp.eurostat.ec.europa.eu. Series codes are V12150 (value added at factor cost) and V16110 (number of persons employed). I construct the Herfindahl indices as the sum of the squared country shares for each indicator. For the textiles, clothing, and footwear sectors, I use NACE industry codes DB17, DB18, and DC193 respectively.

79 For this and the test of instrument validity below, I use the regression based tests set out in Wooldridge (2002, pp. 118-124).
markedly in absolute value terms, but remains statistically insignificant. Amongst the remaining variables, only per capita GDP undergoes an (unexpected) sign change, but it remains statistically insignificant.

How valid is this choice of instruments? Unconditional correlation coefficients between the endogenous explanatory variables and the instruments range from 0.14 to 0.68 for $stds_{st}$, 0.1 to 0.29 for $\frac{iso_{st}}{stds_{st}}$, and 0.05 to 0.12 for $\ln(1 + \tau_{xst})$. First stage estimates in columns 3-5 of Table 3 suggest that when used jointly, the instruments do a relatively good job of explaining the variation in standards and tariffs. $R^2$s are high: over 0.7 for tariffs, and over 0.9 for both standards variables. In all three columns, a Wald test of the joint significance of the four instruments rejects the null hypothesis at the 1% level.

The pattern of signs in the first stage regressions is broadly but not perfectly consistent with the political economy story set out above: larger sectoral value added is associated with less international harmonization and higher tariffs, as is greater geographical concentration in particular member states. Only the coefficients in the tariff equation are individually statistically significant, however. A higher number of employees is also associated with less harmonization, but the geographical dispersion of employment does not carry the expected sign in either case. Results in relation to the total number of standards are mixed: higher employment and greater geographical concentration of employment are both associated with a higher standard count, but the two value added variables carry unexpected negative signs. While these results suggest that it would clearly be desirable to flesh out in more detail the political economy mechanisms behind standards and harmonization, they nonetheless support the relevance of the instruments in this case.

The second point of view from which the validity of these instruments can be tested is their exogeneity with respect to export variety. To do this, I exploit the fact that the model is overidentified (four instruments and three endogenous explanatory variables). A test of the validity of the
overidentifying restriction does not reject the null hypothesis at the 10% level (see Table 3), thereby suggesting that the instruments are indeed exogenous. Taking these results together with the first stage regressions, I conclude that the TSLS estimates appropriately correct for endogeneity bias as it affects standards and tariffs, but that the empirical extent of that bias is relatively limited in this particular case.

2 Robustness Checks: Quotas, Income Levels, and Regional or Preferential Agreements

Thus far, the estimation sample has covered all exporting countries for which data are available (with the exception of the EU-15). The presence in the dataset of a wide variety of very different countries can be used to examine the potential for factors that are omitted in (III.15) to influence the results. For instance, it could be argued that the dummy variables in (III.15) are inadequate to take account of the impacts of ATC quotas on constrained countries. I therefore eliminate those countries from the sample and re-run the TSLS version of (III.15). Results (column 1 of Table 4) are not greatly changed, although the elasticities for the total number of standards and the extent of harmonization are both larger in absolute value.

Next, I focus on countries at relatively low income levels so as to indirectly gauge the extent to which low levels of development might negatively impact firms’ ability to make the investments necessary to comply with foreign standards. It is conceivable, for instance, that lack of access to technology, skills, or finance could effectively raise the costs of compliance with multiple foreign standards for firms in developing countries, thereby making foreign standards more burdensome and harmonization more beneficial. Columns 2-4 of Table 4 present the results from using narrower country samples: in addition to excluding ATC quota countries, I progressively exclude high income countries, upper middle-income countries, and lower middle-income countries.80

80Focusing on low income countries also alleviates the possible difficulties with re-exports and processing trade that were mentioned earlier in relation to variety measures for some high income countries.
country groups are based on World Bank classifications.) The basic thrust of the results remains unchanged. The coefficient on the total number of standards is negative in all cases, and does not change too much in magnitude. It loses statistical significance, however, in the last two columns. The situation is more interesting for the extent of harmonization: it generally increases in magnitude as income falls. The estimated coefficient using low income countries only is 3.0, which is over 50% higher than the estimate obtained using the full sample. The same progression can also be seen for tariffs, although comparison is made more difficult by the lack of precision with which this parameter is estimated. While these results should be interpreted with caution due to the small sample size in column 4, the data are nonetheless suggestive of development-based differences in compliance costs.

As a final check, I consider the possibility that the omission of data on regional and preferential trade agreements could lead to biased estimates in (III.15). It could be argued, for instance, that such agreements indicate a degree of economic and/or political proximity to the EU that could be associated either with the use of similar product standards domestically, or the existence of EU aid or technical support aimed at assisting the compliance process. I therefore re-estimate the model excluding (in addition to ATC quota countries) those countries that have a regional trade agreement with the EU, and the African, Caribbean, and Pacific group of countries that benefit from preferential treatment under the Lomé Convention and the Cotonou Agreement.  

Results are in Table 4 columns 5-6. Estimated coefficients for the two variables of primary interest are in line with those found previously: both are statistically significant at the 5% level, with magnitudes slightly larger in absolute value than those reported in Table 3 column 2. The remaining coefficients are not statistically significant, but generally have signs and magnitudes that are consistent with earlier estimates. The only exception is EU tariffs, which enter column 5 with

---

an unexpected positive sign, and the two ATC liberalization dummies, which enter column 6 with unexpected negative signs. However, given the lack of statistical significance of these coefficients, no strong conclusions are drawn from them at this stage.\textsuperscript{82}

\section*{E Conclusions, Policy Implications, and Future Research}

This paper has provided the first direct empirical evidence that while product standards overall impact negatively on partner country export variety, international harmonization can act as an important mitigating factor. Based on instrumental variables estimates, I conclude that a 10\% increase in the total number of standards leads to a 6\% decrease in partner country export variety, but that a 10 percentage point increase in the proportion of internationally harmonized standards leads to a 0.2\% increase in export variety. I find evidence that this latter effect is around 50\% stronger for low income countries. The empirical results prove to be extremely robust to changes in country sample, and the use of instrumental variables based on a simple political economy framework. They are consistent with a view of the world in which product standards impose fixed costs of adaptation, of which harmonization tends to reduce the multiplicity.

These are significant findings from a development point of view, given the importance of the textiles, clothing, and footwear sectors to economies in the early stages of industrialization. Based on a heterogeneous firms framework, these results would tend to suggest that harmonization can be an effective way of promoting foreign market access for firms with lower productivity than incumbent exporters, since it induces an upwards shift in the export marginal cost cutoff. International harmonization could therefore be expected to encourage exports by small and medium enterprises in developing countries--a prediction that future work using firm level data could test. Importing

\textsuperscript{82}An additional issue, not pursued here, is rules of origin. While such measures can undoubtedly impose additional cost burdens on exporters (e.g., Cadot and de Melo, 2007), they are in effect held constant within this sample because only one importing market (the EU) is included.
countries looking to provide an impulsion to non-traditional exports from developing countries could perhaps use international standards harmonization as a complement to more generous tariff preferences and more open rules of origin.

An alternative way of interpreting the results presented here is in terms of export diversification, an important policy issue for many developing countries. By equating variety growth and diversification, a case can be made that international standards harmonization could be one way in which the large, rich country import markets could help support export diversification in developing countries. These results therefore complement recent work on diversification, which has highlighted the importance of policies such as trade facilitation and the rationalization of barriers to domestic market entry within developing countries (Dennis and Shepherd, 2007).

The main obstacle to future empirical work in the area of product standards and their trade effects remains limited data availability. While the World Bank’s EU Standards Database provides information on the textiles, clothing, and footwear sectors, there is clearly a need to expand on this. One direction for possible expansion would be the addition of further countries, including data on national standards in EU-15 member states. At the current time, the data do not permit an assessment of the impact of product standard harmonization on geographical export diversification; however, the relatively weak (but statistically significant) elasticity found here suggests that geographical diversification, rather than product variety, might be the more important channel for trade growth in this case. It is to be hoped that future work will investigate this possibility.

It would of course also be desirable to increase the sectoral reach of standards data. In particular, it would be useful to include products of interest to middle income developing countries, such as electronic goods. Building on previous efforts to exploit the Perinorm bibliographic database (e.g., Swann et al., 1996; Moenius, 2004) would likely prove very beneficial to empirical work in this area.
TSLS estimates suggest that the size of a sector and its geographical distribution across EU member states are important determinants of the total number of standards and the degree of international harmonization. This constitutes important preliminary support for a political economy view of trade-related standards and harmonization, in which these measures are influenced by the kinds of factors more commonly associated with endogenous tariffs. However, that argument is only briefly sketched out here, and results are not uniformly consistent with it. To my knowledge, only Essaji (2005) addresses the political economy of product standards and trade from an empirical perspective. A future expansion of his approach to a multi-country framework, along with incorporation of Perinorm data, would provide important insights into the mechanisms at work. Such research would have important policy implications since it could help lay the foundations for a rigorous identification of those standards which are unduly tainted by protectionist pressures.

F References


and Heterogeneous Firms”, *Review of Economic Studies*, 74, 31-66.


Disdier, Anne-Célia, Lionel Fontagné, and Mondher Mimouni, Forthcoming, "The Impact of Regulations on Agricultural Trade: Evidence from SPS and TBT Agreements", *American Journal of Agricultural Economics*.


G  Tables

Table 1: Descriptive statistics for the EU Standards Database (EUSDB).

<table>
<thead>
<tr>
<th></th>
<th>Clothing(^1)</th>
<th>Footwear(^2)</th>
<th>Textiles(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% ISO</td>
<td>Total</td>
</tr>
<tr>
<td>1995</td>
<td>15</td>
<td>6.67</td>
<td>5</td>
</tr>
<tr>
<td>1996</td>
<td>18</td>
<td>5.56</td>
<td>10</td>
</tr>
<tr>
<td>1997</td>
<td>20</td>
<td>5.00</td>
<td>15</td>
</tr>
<tr>
<td>1998</td>
<td>28</td>
<td>17.86</td>
<td>18</td>
</tr>
<tr>
<td>1999</td>
<td>32</td>
<td>18.75</td>
<td>31</td>
</tr>
<tr>
<td>2000</td>
<td>28</td>
<td>21.43</td>
<td>39</td>
</tr>
<tr>
<td>2001</td>
<td>32</td>
<td>21.88</td>
<td>57</td>
</tr>
<tr>
<td>2002</td>
<td>35</td>
<td>22.86</td>
<td>60</td>
</tr>
<tr>
<td>2003</td>
<td>41</td>
<td>19.51</td>
<td>67</td>
</tr>
</tbody>
</table>

1. HS chapters 61-63.
2. HS chapter 64.
3. HS chapters 50-60.
4. Total refers to the total number of CEN European standards in force for each sector-year. % ISO is the proportion of the total number of standards that are internationally harmonized (i.e., identical or equivalent to an ISO standard).

Table 2: Descriptive statistics for product variety (\(\Lambda_{xxx}\)).

<table>
<thead>
<tr>
<th></th>
<th>Clothing(^1)</th>
<th>Footwear(^2)</th>
<th>Textiles(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Max.</td>
<td>Countries</td>
</tr>
<tr>
<td>1995</td>
<td>0.22</td>
<td>0.94</td>
<td>185</td>
</tr>
<tr>
<td>1996</td>
<td>0.21</td>
<td>0.95</td>
<td>193</td>
</tr>
<tr>
<td>1997</td>
<td>0.28</td>
<td>0.95</td>
<td>185</td>
</tr>
<tr>
<td>1998</td>
<td>0.24</td>
<td>0.95</td>
<td>187</td>
</tr>
<tr>
<td>1999</td>
<td>0.26</td>
<td>0.96</td>
<td>186</td>
</tr>
<tr>
<td>2000</td>
<td>0.26</td>
<td>0.96</td>
<td>191</td>
</tr>
<tr>
<td>2001</td>
<td>0.28</td>
<td>0.96</td>
<td>203</td>
</tr>
<tr>
<td>2002</td>
<td>0.28</td>
<td>0.96</td>
<td>200</td>
</tr>
<tr>
<td>2003</td>
<td>0.27</td>
<td>0.96</td>
<td>204</td>
</tr>
</tbody>
</table>

5. \(\Lambda_{xxx}\) is calculated as described in the text, using 8-digit Eurostat import data.
Table 3: Regression results.

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>TSLS</th>
<th>TSLS First Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ln (A)</td>
<td>ln (A)</td>
<td>ln(Std)</td>
</tr>
<tr>
<td>ln(Std)</td>
<td>-0.763***</td>
<td>-0.587**</td>
<td>[0.152]</td>
</tr>
<tr>
<td>ISO/Std</td>
<td>1.535***</td>
<td>1.816***</td>
<td>[0.334]</td>
</tr>
<tr>
<td>ln(GDP)</td>
<td>0.056</td>
<td>0.885</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>[0.500]</td>
<td>[1.848]</td>
<td>[0.013]</td>
</tr>
<tr>
<td>ln(GDPPC)</td>
<td>0.379</td>
<td>-0.394</td>
<td>0.036*</td>
</tr>
<tr>
<td></td>
<td>[0.470]</td>
<td>[1.665]</td>
<td>[0.021]</td>
</tr>
<tr>
<td>ln(1+t)</td>
<td>-0.019</td>
<td>-0.617</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.039]</td>
<td>[1.156]</td>
<td></td>
</tr>
<tr>
<td>ATC2</td>
<td>-0.112</td>
<td>0.028</td>
<td>-1.082***</td>
</tr>
<tr>
<td></td>
<td>[0.115]</td>
<td>[0.214]</td>
<td>[0.194]</td>
</tr>
<tr>
<td>ATC3</td>
<td>-0.095</td>
<td>0.052</td>
<td>-0.204</td>
</tr>
<tr>
<td></td>
<td>[0.087]</td>
<td>[0.130]</td>
<td>[0.124]</td>
</tr>
<tr>
<td>Quota</td>
<td>-0.347***</td>
<td>-0.349***</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>[0.060]</td>
<td>[0.082]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>Quota*ATC2</td>
<td>0.006</td>
<td>-0.04</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>[0.096]</td>
<td>[0.150]</td>
<td>[0.002]</td>
</tr>
<tr>
<td>Quota*ATC3</td>
<td>0.017</td>
<td>0.061</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>[0.176]</td>
<td>[0.172]</td>
<td>[0.002]</td>
</tr>
<tr>
<td></td>
<td>[10.690]</td>
<td>[20.457]</td>
<td>[18.325]</td>
</tr>
<tr>
<td>ln(VA)</td>
<td>-1.837**</td>
<td>-0.189</td>
<td>0.360**</td>
</tr>
<tr>
<td></td>
<td>[0.697]</td>
<td>[0.277]</td>
<td>[0.168]</td>
</tr>
<tr>
<td>ln(Emp)</td>
<td>4.329***</td>
<td>-0.553</td>
<td>-0.064</td>
</tr>
<tr>
<td></td>
<td>[1.267]</td>
<td>[0.520]</td>
<td>[0.351]</td>
</tr>
<tr>
<td>ln(Herf VA)</td>
<td>-4.653***</td>
<td>-0.341</td>
<td>0.820**</td>
</tr>
<tr>
<td></td>
<td>[0.825]</td>
<td>[0.345]</td>
<td>[0.311]</td>
</tr>
<tr>
<td>ln(Herf Emp)</td>
<td>6.710***</td>
<td>1.051</td>
<td>-0.109</td>
</tr>
<tr>
<td></td>
<td>[2.051]</td>
<td>[0.770]</td>
<td>[0.420]</td>
</tr>
<tr>
<td>Obs.</td>
<td>3607</td>
<td>3607</td>
<td>3607</td>
</tr>
<tr>
<td>R2</td>
<td>0.68</td>
<td>0.65</td>
<td>0.99</td>
</tr>
<tr>
<td>Exogeneity (F)</td>
<td>14.59***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overidentification (X²)</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H0: Instr. = 0 (F)</td>
<td>13.78***</td>
<td>11.63***</td>
<td>6.23***</td>
</tr>
</tbody>
</table>

1. All models include fixed effects by exporter, sector, and year. Robust standard errors adjusted for clustering by sector-year appear in square brackets under the coefficient estimates. Statistical significance
is indicated using * (10%), ** (5%), and *** (1%).

2. Exogeneity and overidentification tests are the regression-based tests outlined in Wooldridge (2002, pp. 118-124). The F-test is of the null hypothesis that all instruments are jointly zero.
Table 4: Robustness checks.

<table>
<thead>
<tr>
<th></th>
<th>No Quota</th>
<th>No High</th>
<th>No U-Mid</th>
<th>No L-Mid</th>
<th>No RTA</th>
<th>No ACP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Stds)</td>
<td>-0.638**</td>
<td>-0.642*</td>
<td>-0.424</td>
<td>-0.698</td>
<td>-0.728**</td>
<td>-0.621**</td>
</tr>
<tr>
<td></td>
<td>[0.290]</td>
<td>[0.328]</td>
<td>[0.368]</td>
<td>[0.485]</td>
<td>[0.341]</td>
<td>[0.277]</td>
</tr>
<tr>
<td>ISO/Stds</td>
<td>2.054***</td>
<td>2.256***</td>
<td>2.148***</td>
<td>3.021***</td>
<td>2.406***</td>
<td>1.948***</td>
</tr>
<tr>
<td></td>
<td>[0.441]</td>
<td>[0.548]</td>
<td>[0.671]</td>
<td>[0.986]</td>
<td>[0.500]</td>
<td>[0.510]</td>
</tr>
<tr>
<td>Ln(GDP)</td>
<td>0.463</td>
<td>1.155</td>
<td>0.973</td>
<td>-3.467</td>
<td>0.057</td>
<td>0.826</td>
</tr>
<tr>
<td></td>
<td>[2.278]</td>
<td>[2.969]</td>
<td>[2.640]</td>
<td>[3.494]</td>
<td>[1.066]</td>
<td>[0.949]</td>
</tr>
<tr>
<td>Ln(GDPPC)</td>
<td>-0.049</td>
<td>-0.679</td>
<td>-0.322</td>
<td>3.913</td>
<td>0.315</td>
<td>-0.220</td>
</tr>
<tr>
<td></td>
<td>[1.991]</td>
<td>[2.746]</td>
<td>[2.523]</td>
<td>[3.645]</td>
<td>[0.837]</td>
<td>[1.424]</td>
</tr>
<tr>
<td>Ln(1+t)</td>
<td>-0.152</td>
<td>-0.708</td>
<td>-1.857</td>
<td>-5.721</td>
<td>0.302</td>
<td>-0.280</td>
</tr>
<tr>
<td></td>
<td>[1.224]</td>
<td>[1.452]</td>
<td>[1.625]</td>
<td>[3.942]</td>
<td>[1.619]</td>
<td>[1.395]</td>
</tr>
<tr>
<td>ATC2</td>
<td>0.059</td>
<td>0.108</td>
<td>0.319</td>
<td>0.272</td>
<td>0.116</td>
<td>-0.029</td>
</tr>
<tr>
<td></td>
<td>[0.200]</td>
<td>[0.246]</td>
<td>[0.215]</td>
<td>[0.347]</td>
<td>[0.225]</td>
<td>[0.252]</td>
</tr>
<tr>
<td>ATC3</td>
<td>0.017</td>
<td>0.085</td>
<td>0.274</td>
<td>0.662</td>
<td>0.015</td>
<td>-0.067</td>
</tr>
<tr>
<td></td>
<td>[0.133]</td>
<td>[0.211]</td>
<td>[0.285]</td>
<td>[0.396]</td>
<td>[0.144]</td>
<td>[0.126]</td>
</tr>
<tr>
<td></td>
<td>[26.209]</td>
<td>[33.497]</td>
<td>[53.348]</td>
<td>[45.206]</td>
<td>[20.707]</td>
<td>[19.590]</td>
</tr>
<tr>
<td>Obs.</td>
<td>3265</td>
<td>2696</td>
<td>2005</td>
<td>992</td>
<td>2894</td>
<td>1548</td>
</tr>
<tr>
<td>R2</td>
<td>0.66</td>
<td>0.59</td>
<td>0.34</td>
<td>0.63</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Exog. ($F$)</td>
<td>14.92***</td>
<td>17.38***</td>
<td>19.12***</td>
<td>9.01***</td>
<td>17.21***</td>
<td>37.61***</td>
</tr>
<tr>
<td>Overid. ($\chi^2_1$)</td>
<td>0.98</td>
<td>1.348</td>
<td>0.2</td>
<td>0.69</td>
<td>0.00</td>
<td>0.868</td>
</tr>
</tbody>
</table>

1. All models include fixed effects by exporter, sector, and year. Estimation is by TSLS (first stage results available on request). Robust standard errors adjusted for clustering by sector-year appear in square brackets under the coefficient estimates. Statistical significance is indicated using * (10%), ** (5%), and *** (1%).

2. Exogeneity and overidentification tests are the regression-based tests outlined in Wooldridge (2002, pp. 118-124).

3. Column 1 excludes countries subject to quotas under the ATC, columns 2-4 exclude (in addition) high income, upper middle income, and lower middle income countries. Column 5 excludes countries having a regional trade agreement with the EU, while column 6 excludes (in addition) all members of the African, Caribbean, and Pacific group of countries.
H  Figures

Figure 1: Simulated change in export marginal cost cutoffs (%) following harmonization.

1. Following Bernard et al. (2007), I use the following invariant parameters: $\sigma = 3.8$, $k = 3.4$, $L = 1000$, $\beta = 0.5$, $\bar{c} = 0.2$, and $f^e = 2$.

2. Prior to harmonization, $f^i_{xH} = f^i_{xF} = 0.15$ and $f^H_{dx} = f^F_{dx} = 0.1$. $f^HR_x = f^FR_x$ varies over the range 0.15 to 0.2, while $f^R_d$ varies over the range 0.1 to 0.15.

3. Post harmonization, $f^i_{xH}$ remains unchanged but $f^i_{xH}$ varies over the range 0.15 to 0.2, while $f^R_d = f^F_d$ varies over the range 0.1 to 0.15.

4. The solid line shows the percentage change in the export cutoff for $H$ to $F$ following harmonization. The dotted line is for $H$ to $R$, the dashed line is for $F$ and $R$ to $H$, the dotted and dashed line is for $F$ to $R$, and the crossed line is for $R$ to $F$. 
General Conclusion
General Conclusion

This dissertation has taken some initial steps towards a better understanding of the development policy implications of recent advances in trade theory involving heterogeneous firms and trade growth at the extensive margin. Chapter I showed that this framework provides a useful way of conceptualizing export (product) diversification, an issue which has long been of concern to developing countries. In Chapter II, the framework was adapted to more closely examine the question of export market (geographical) diversification. Both chapters highlighted the importance of trade costs and barriers to domestic market entry in promoting or impeding diversification in the product and market dimensions. Finally, Chapter III examined the extensive margin impacts of product standards. It showed that the reduction in cost multiplicity brought about by international harmonization can promote partner country export growth at the extensive margin, even when harmonization entails moderate increases in compliance costs.

The purpose of this final section is twofold. It first presents some more general policy implications flowing from the results presented in the three substantive chapters. They are put forward with the aim of contributing to broader discussions on trade and development policy. They are therefore more speculative in nature than those presented in the concluding sections of the individual chapters. Finally, the second part of this section examines some possible directions that could usefully be explored by future research in this area.

Policy Implications

In broad terms, what are the policy implications flowing from the findings presented in this dissertation? First, the dissertation’s focus on firm heterogeneity and its links to trade growth at the extensive margin suggests that it is important to keep dynamics in mind when designing development policies that take full advantage of the opportunities provided by international trade. Export
growth does not only take the form of higher levels of exports of the same goods to the same markets, but can involve both new products and new trading partners. Growth at the extensive margin is particularly important because it entails new entry by domestic firms into international markets. Since it is generally the largest and most productive firms that are the first to enter world markets, trade expansion at the extensive margin implies the possibility of greater export involvement on the part of small and medium firms. Measures to expand trade at the extensive margin could therefore produce substantial effects in terms of broader private sector development efforts.

The second major set of policy implications associated with this dissertation’s findings relates to the role of "behind-the-border" issues in the context of current trade and development policy. All three chapters reinforce the potential importance from a trade point of view of policies that might at first glance appear to fall under the heading of "domestic regulations". This is the case for measures such as product standards, and regulations governing market entry. While there have long been discussions as to the potential trade impacts of product standards, the importance of market entry regulations is only now coming to be recognized. Consideration of the impact that such a fundamental set of rules can have on trade outcomes is consistent with a broader movement in the literature towards a focus on the importance of institutions: see e.g., Acemoglu et al. (2001) in the growth context, and Anderson and Marcouiller (2002), Levchenko (2007), and Nunn (2007) in the trade context.

From a policy point of view, the moral of this story is that in an increasingly globalized world, the distinction between trade policies and domestic regulations is becoming harder and harder to sustain. Ultimately, there could be benefits both analytically and in terms of the policy process in allowing the two areas to coalesce to an even greater degree than at present. Such an approach would imply a focus on the costs and benefits of individual regulatory measures, but without the traditional limitation of a closed economy framework. Regulatory impact assessment would, in
other words, take account of both domestic and trade effects.

The case of regulations governing market entry is a useful example. There is surely a strong case for developing countries to take a close look at the economic logic behind, and effects of, these rules. For instance, Klapper et al. (2007) show that higher business registration costs tend to hold back entrepreneurial activity. As the three chapters of this dissertation emphasize, however, it is not just such "domestic" outcomes that are in question: high entry costs can also impact the economy more broadly, including at its interface with the world market. The suggestion made here is that in assessing these types of regulations, developing countries should take account of trade impacts in addition to the initial impacts on domestic economic activity.

Developed countries could, of course, adopt a similar approach. In light of the results presented in Chapter III, product standards is one area in which it could be important to expand the scope of regulatory impact assessments beyond the border. Such an approach could help support a genuine mainstreaming of development issues in the broader regulatory reform context. It could also aid significantly in efforts to ensure coherence between overseas development objectives, and more general economic and regulatory policies.

The third major policy implication of this dissertation’s findings flows from the previous point. In line with a number of other recent contributions to the literature, the results presented here suggest that it may be damaging for discussions on trade and development policy to be reduced to a question of tariff preferences. (See also Ozden and Reinhardt, 2005.) Clearly, maintenance of high or escalating tariff levels is unlikely to do much to help developing country trade performance. But it does not follow that lowering or removing such tariffs on a preferential basis is all that is required.

While preference schemes can certainly have a useful role to play, they do little if anything to address a number of important constraints on developing country export growth. One set of such constraints relates to non-tariff barriers, such as the product standards analyzed in Chapter
III. Restrictive rules of origin also limit the ability of developing countries to take advantage of preferences, by increasing access costs: see e.g., Anson et al. (2005). But it is not only developed country policies that are important in this context. Developing country exporters also experience serious supply-side constraints that make it difficult for them to respond appropriately to market access opportunities overseas. As Dollar et al. (2006) and Francois and Manchin (2007) show, many factors contribute to these constraints. Both hard (physical) and soft (regulatory or institutional) infrastructure are important. Chapters I and II reinforce the point that it is not just physical infrastructure that matters, since export costs related to port and customs charges, documentation costs, and internal transport, are found to impact the extensive margin of trade.

Bringing these points together suggests that if preference schemes are to be effective in increasing developing countries’ level of integration with world markets, then they will need to be accompanied by a whole raft of complementary policies. Some such policies can be implemented by the donor countries themselves, but others will require action by preference beneficiaries. In terms of loosening the binding constraints on developing country exporters, it seems likely that these complementary policies may be at least as important as tariff preferences themselves. The results in all three chapters of the dissertation are broadly consistent with this view: while trade facilitation measures, entry regulations, and product standards are consistently found to impact significantly on extensive margin trade growth, the evidence regarding partner country tariffs is much more mixed. Although it would be inappropriate to conclude that foreign tariffs do not impact extensive margin growth at all, these results would nonetheless caution policymakers against relying too heavily on preferences.
Directions for Future Research

The findings presented in this dissertation represent a useful starting point in terms of understanding the implications of firm heterogeneity and extensive margin trade growth for developing countries. Much remains to be done, however. The purpose of this final section of the dissertation is to identify a number of directions in which future research could usefully extend and expand upon the work presented here.

Endogeneity of Trade Costs

Trade costs are clearly an important determinant of the way in which export and import activity is distributed over the globe. It is striking, however, that a large proportion of empirical trade research treats these costs as exogenous to trade flows. The most surprising example is that gravity models including contemporaneous applied tariffs are rarely estimated by instrumental variables. This is so despite the existence of a large and sophisticated political economy literature on endogenous tariffs: see e.g., Grossman and Helpman (1994), and Goldberg and Maggi (1999); Rodrik (1995) provides a review. Trade facilitation indicators (e.g., Wilson et al., 2005) are not usually instrumented either, even though there is again good reason to suspect that at least some of these measures—such as infrastructure quality, and corruption—could be endogenous to trade.

There are some notable exceptions to this tendency, which demonstrate the importance of finding appropriate instruments for international trade costs. Baier and Bergstrand (2007), for instance, find that instrumenting for membership in free trade agreements leads to a five fold increase in their estimated trade impacts. In the area of trade facilitation, Djankov et al. (2006) use the number of administrative formalities involved in importing as an instrument for import time, in order to avoid bias due to congestion effects.

Clearly, it is important in the context of policy-relevant international trade research to deal
Development Perspectives on Trade Growth at the Extensive Margin

squarely with these types of endogeneity issues. In this dissertation, geographical and historical (institutional) data have been used as instruments for certain trade cost factors, as have a range of political economy variables. There remains much to do in this area, however, and it is likely to prove fruitful in the future. A better understanding of the political economy of non-tariff measures would be a useful input into this process.

Data

As in many other areas of development-related work, the lack of consistent, high-quality data can be a significant constraint affecting applied research on developing countries’ trade relations. The last few years have, in fairness, seen some important improvements in this situation. The Market Access Map (MAcMap) database provides a fully bilateral breakdown of applied tariffs rates at the 6-digit HS level, including those linked to preference schemes (Laborde et al., Forthcoming). The World Bank’s Doing Business database provides detailed information on the costs of exporting and importing in nearly 180 countries, as well as the time taken for each operation and the number of administrative formalities involved (Djankov et al., 2006). Finally, the World Bank’s Enterprise Surveys provide consistent firm level data, including a rich variety of trade-related information, across 104 countries.

Despite these advances, non-tariff barriers remain a black spot. A number of interesting and useful research questions simply cannot be answered until better and more detailed data on NTBs become available. For instance, Chapter III has shown that international harmonization of product standards can increase partner country export variety. However, the analysis relied exclusively on product standard data from the European Union, due to the lack of comparable information for other countries. It would also be interesting to look at the market diversification effects of international harmonization: by eliminating cost multiplicity, it should be easier for countries to enter
multiple export markets. But this question cannot be investigated empirically without information on the extent of international harmonization in individual markets—and apart from the data used in Chapter III, or the data on EU harmonization directives used by Chen and Mattoo (2004), such information simply does not exist.

**Technological Change and Productivity Growth**

As noted in the Introduction, the Melitz (2003) framework is potentially of huge relevance to developing countries, because it provides a convincing link between trade openness, technological change, and productivity growth. The process at the heart of the model is a Schumpeterian one: productivity (technology) differs across firms, and openness leads to a reallocation of resources from low- to high-productivity firms. There is strong empirical support for the essential features of this mechanism. It is considerably clearer and more convincing than the largely inconclusive "openness and growth" literature of the 1980s and 1990s.

As yet, however, there is relatively little literature examining the importance of this openness-productivity growth linkage in an empirical sense (cf. the theoretical model presented by Baldwin and Robert-Nicoud, 2008). Feenstra and Kee (2006) develop a Melitz (2003)-type model in order to investigate the impact of export variety growth on total factor productivity growth. They find that the association is positive and significant: average growth in export variety of 4.6% per annum is associated with a productivity gain of 4.5% during the 1980s and 1990s.

Although useful in explaining changes in productivity within countries over time, the mechanism investigated by Feenstra and Kee (2006) accounts for only a small proportion of the productivity differences observed between countries. There is thus considerable scope for extending this type of work in order to better understand the role that openness and trade can play in fostering productivity growth. One interesting angle to explore might be a combination of the Melitz (2003) and
Ethier (1982) approaches. Fugazza and Robert-Nicoud (2006) take the first steps towards constructing such a model, but embed it in a specific context of South-South and North-South interactions that makes it difficult to bring out the more general point suggested here. Moreover, there have been as yet no empirical tests of this approach along the lines of those conducted by Feenstra and Kee (2006) for the more standard Melitz (2003) framework.

The interest of this research direction is that it would produce a heterogeneous firms model in which the availability and price of a range of intermediate inputs is a determinant of export performance in final goods sectors. Such a framework might support the existence of a link from both export and import diversity to productivity growth. From an empirical point of view, it would be of real interest in the development context to know which of these two effects might dominate.

Given the centrality of productivity growth in the development process, it is to be hoped that future work will pay attention to the policy measures that can help promote development through the kinds of mechanisms discussed here. This dissertation can be seen as a first, preliminary step towards such an understanding.

References


