Heterogeneity in Development

Manuscript committee:

Prof. Peter van Bergeijk
Prof. Jaap Bikker
Prof. Erwin Bulte
Prof. Wolter Hassink

ISBN number: 978-94-91870-04-0
Tjalling C. Koopmans Dissertation Series
USE 017

Heterogeneity in Development

Heterogeniteit in Ontwikkeling
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus,
prof. dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
maandag 9 december 2013 des middags te 14:30 uur

door

Han-Hsin Chang

geboren op 22 februari 1986 te Taipei, Taiwan

Promotor:
Prof. dr. C. van Marrewijk

Co-promotor:
Dr. M.C. Schramm

This book is dedicated to my loving family

Acknowledgement

Starting a life away from home and abroad is not entirely new to me. However, working it out for the first time alone and in a non-English speaking country was definitely something new. Looking back on this journey I started out six years ago, it surprises me how various turning points would have made things different. Without the support of many people I met along this journey, the completion of this thesis would not have been possible. It is these people I would like to thank now.

First and foremost, I would like to thank my supervisors, Charles van Marrewijk and Marc Schramm for giving me the opportunity to start out my research career at Utrecht University School of Economics. Charles, I remember vividly how I felt intimidated by you when we first worked together. After each meeting, starting from trying to meet up with your expectations, I gradually learned to plan my research and to progress accordingly, in which you always responded with a very positive attitude. I appreciated that you gave me the time to think and struggle as a natural part of the research training. I am also very grateful of your full support in becoming a visiting scholar at the University of Adelaide. The trip not only granted me the chance to pay my first visit to Australia, the country I dreamed of going to since I was a child, but also expanded my field of interest in economics. Charles, I thank you for all the challenges and opportunities you allowed me to experience along this journey. I am glad to have you as my mentor, who gave me guidance, support and confidence as I try out my wings. I would also like to thank my co-promoter, Marc Schramm. Once in a while, you would read my papers very thoroughly and ask me very detailed questions and propose ideas beyond the many things I have considered or looked into. I thank you for your comments and suggestions for my research.

I also thank my reading committee, Prof. Peter van Bergeijk, Prof. Jaap Bikker, Prof. Erwin Bulte and Prof. Wolter Hassink, for carefully reading and giving constructive comments for this dissertation. I also have to thank all my colleagues who had a direct influence on this piece; including my co-author, Charles van Marrewijk, who is of course related to all chapters in this dissertation. Chapter 4 is a collaboration work with Randy Stinger and Wendy Umberger from the University of Adelaide; Marc Schramm in Chapter 5 and Jaap Bikker, Joras Ferwerda, Loek Groot and Mark Kattenberg in collaboration of working on Chapter 6 of this dissertation. I also thank seminar participants at USE and ETSG 2011, NARSC 2011 and APTS 2012 and 2013 conference participants for all the comments and suggestions given to the papers collected into this dissertation.

I would like to pay my gratitude to the teachers that had such a great influence in supporting me all this way. Prof. Christopher Findley, I thank you for taking time out of your busy schedule to make sure my visit to the University of Adelaide was fruitful. By meeting the professors and discussing with you the possible avenue for a new research topic, I found new passion in a different area of study. I would not have such unexpected happy encountering if I had not met you. Brigitte, I thank you for your encouragement. Your words: "always shoot for the brightest stars" will forever be in my mind giving me the strength to take challenges. Wolter, I thank you for having so much trust in me and giving me the opportunity to teach. I learned a lot from the experience.

In my daily life, I have been blessed by working with many amazing people and having the chance to share parts of your life over the years. Saraï, Britta, Suzanne, Ye, Maryam and Krista, my sweet office-mates, I am happy to have shared the office with you making our office one of the most popular places among the PhD rooms. I will always remember the soup days in the cold winter and the cheerful sound of "coffee?, tea?" offers echoing in the room. Also, the after-work runs buddies, Saraï, Suzanne and Anna, who reminded me to have a good work-life balance. I thank my fellow friends, Shiwei, Lu, Bastian, Ioana, Bahar, Sarah, Malka, Emre, Joyce, Jasper, Peter, Luuk, Thomas, Daniel, Marcel, Secil, Swantje, Ian and Dan for bringing so much more laughter to the sometimes plain research life. I also express sincere thanks to my friends in Adelaide: Faqin, Xiaobo, Hang, Kai, Jiang, Di, Dou, Kofi, Sharmina, Jaga and Kate. Thank you for the unconditional welcome and discussion of my early work. I thank you all, my friends, for making Adelaide a memorable part of my PhD and making Utrecht a beautiful place in my heart.

Also I would like to thank my dearest friends, Mark Kattenberg, Bon and Yifen. Mark, I am grateful for knowing you and having you as a companion early on in this journey. I thank you for spending time and effort in reading and improving my manuscripts. You have giving me a lot more than encouragement and constructive advice that made the completion of this dissertation possible. But even more importantly, I am thankful for your friendship. I will always remember the hours we spend studying together during the research master years; and all the coffee breaks spent with you, for which we talked about our research progress and ideas, and things that are closest and matter so much to our life. I am thankful to meet you.

邦, 謝謝你時常任我在最低落的時候找你, 傾聽我, 與我談心直到我整理完我紊亂的思緒. 幾回的驚喜總讓我感受到滿滿的友情與來自台灣的溫暖. 宜雰, 我很高興能夠在荷蘭認識你, 與你分享旅行與生活的所見所聞, 討論許許多多我們對於社會的在乎與關懷.

Words cannot express my gratitude to my family for their unconditional love and caring. 謝謝阿公阿嬤總是關心我在國外的飲食起居,老要擔心我論文進度,怕我壓力太大,沒有好好吃好好睡. 謝謝儒,總用溫暖的孩子氣的聲音紓解我的壓力; 也謝謝致,在我不在家的時間對家裡的照顧. 謝謝 Mommy 總在我回家時煮滿桌我最愛的 讓我出國之後魂牽夢縈的菜,讓我感受你滿滿的愛.謝謝爸爸媽媽這些年的支持讓我在國外闖蕩,儘管多數時候不至讓你們太過擔心,但免不了的是掛心. 謝謝你們總相信我,讓我實現自己,讓我知道無論發生事,我都有溫暖的家可以回,有你們做我最重要的依靠.

Last, I thank my boyfriend, soon to be husband, Rhymer Young for his love and support throughout these years. I thank you for being such a good listener, and always being the first one who I mumble my messy and puzzling thoughts to. I am so happy to be able to share everything with you. I am grateful for your support in so many decisions, the decisions in particular that also challenge our relationship. You would encourage me however hard it is for you to say and put even more patience and love to make us strong. I am thankful of having you. You know me so well and love me so dearly, however far apart we are.

Han-Hsin Chang
Utrecht, July 2013

Table of Contents

List of Tables

Chapter 2

Chapter 4

Chapter 5

Chapter 6

List of Figures

Chapter 1 Introduction

1.1 Background[1]

Globalization, trade liberalization and the lowering of trade barriers have generally led to increased opportunities for producers to access foreign markets. The relationship between international trade and development has been widely recognized, while the heterogeneous actions taken by economic agents and the impact on them resulting from further trade liberalization has only been recognized and studied more intensively in recent years. Popularized by Thomas Friedman's book, "The World is Flat", the idea is that foreign competition seems to come falling from the sky as if there exist no barrier to trade. It depicts vividly the increasing foreign competition faced by previously unaffected sectors and domestic producers.

The traditional (classical) trade theories predict that trade liberalization will result in an increasing specialization at the sectoral level due to the comparative advantage that arises from the differences in technology or factor endowment (Ohlin; Ricardo). Studies have shown that in any industry, producers differ widely in size and productivity. Heterogeneity among producers may be obvious, but the documentation of such differences in a more stylized manner at the micro-level started only in the past two decades. It is now recognized that analyses and policy can no longer continue to treat developing countries, firms and farmers as homogeneous entities and economic agents, as traditional trade theory assumes.

The focus is no longer on identifying the winning (booming) and losing (diminishing) sectors, but on stressing that not all producers in the sector with less comparative advantage will be wiped out on facing intensified competition. Instead we see that some producers struggle while other producers were able to seize the opportunity presented by the enlarged market and to become exporters or foreign investors. In short, the falling trade tariffs and the advance in ICT technologies over the past decades further decreases the cost for economic

[1] We point out some key contributors in the field without referencing them here in detail as references can be found in the chapters that follows.

agents to integrate into all sectors in the world economy. This provides both opportunities and challenges to all producers.

1.2 Gaps in the Literature, questions and contributions

Empirical heterogeneity

Several studies have linked stylized facts revealing the *on average* heterogeneity between producers. Most notably, producers differ in their productivity and size. A corollary to that observation is that exporters are different from non-exporters. They tend to be bigger, more capital intensive, and more productive, thus paying higher wages. Motivated by the empirical evidence, Melitz provided the cornerstone theoretical model for explaining the reallocation of resources among heterogeneous producers which is induced by trade liberalization through a selection effect. The model provides a strong prediction that trade liberalization induces an aggregate growth in productivity. It captures the ranking order of average productivity between domestic, exporting and FDI firms. The heterogeneity models following Melitz (Helpman, Yeaple, Aw and Lee) are characterized by productivity *thresholds* and *cutoff values* in distinguishing producers by their international orientation. The first productivity threshold – the viability cutoff – indicates the minimum productivity level a producer must reach to generate non-negative profits. A second threshold – the export cutoff – indicates the minimum productivity level needed for profitable exports.

Current research has so far paid little attention to illustrating the actual productivity distribution, and the research that listed some stylized facts on producer heterogeneity was based mostly on the data from an individual developed country. This is to some extent because analyzing the degree of heterogeneity among producers requires access to reliable micro-level data, which is hard to come by for most developing countries. Not to mention providing a coherent analysis for several countries in one study.

Against this background, Chapter 2 aims to fill in these gaps through providing coherent evidence of the productivity heterogeneity for 15 developing countries. The chapter is probably the first attempt to put so much emphasis on illustrating the productivity distribution and making a cross country comparison using developing country data. The first set of questions we will answer concerns the empirical productivity distribution differentials between countries in different development stage and between producers that differs in terms of international orientation. More specifically,

> *What does the empirical productivity distribution of viable firms looks like? How do we characterize the productivity distributions of producers in different countries and producers of different international orientation?*

An innovative treatment was used to make a consistent comparison of the productivity distribution in countries at different stages of development. The treatment also allowed a comparison between firms of different degrees of internationalization. The existing studies in international trade usually focus on manufacturing and less on the services sectors. Another contribution made is providing a comparison of the productivity distribution of producers in the services sectors side by side with the productivity distribution of producers in the manufacturing sectors.

Theoretical Heterogeneity

After the sketch of the empirical productivity distribution and establishment of the contradiction between the empirical and theoretical models, the second question which should be answered with an improvement in the model is:

> *How do productivity and fixed cost heterogeneity affect producer profitability and how does trade liberalization affect overall aggregate welfare through a selection effect?*

The simple yet innovative ingredient added to the Melitz model is introducing fixed cost heterogeneity. Chapter 3 presents the model. The results provide strong implications for the assumptions we made about the underlying distribution of producers' productivity and fixed costs. We shall also show by simulation that we will arrive at different predictions for the productivity distribution of viable firms and exporters in equilibrium.

Agricultural Heterogeneity

While the studies in international trade usually focus on manufacturing and less on services sectors, producers in the agricultural sector are even less mentioned in the trade literature. Trade in agriculture has a long history, but it was dominated by trade in commodities such as wheat, rice, salt, sugar, tea and coffee beans. The dynamics that followed the increase in per capita income, urbanization and changes in lifestyle as more consumers in developing countries shift their consumption from staple food towards high value fruits and vegetables, has resulted in rapid modernization of the agri-food sector since the 1980s. This was stimulated further with policy changes that drive privatization, trade and foreign investment in these countries. The increasing demand from exporting and initially mostly foreign invested supermarkets for fresh fruits and vegetables (high-value agricultural), provides opportunities for the population lagging most behind economically to participate in the growing modernized agricultural sector in the developing countries. The FAO reports that around 48 percent of the total economically active population in developing countries is employed in the agricultural sector and nearly 55 percent of the total population lives in the rural area. With fewer intermediate transactions between producers and food consumers, Neven and Reardon and many others believed that the dynamics along the value chain for fresh fruits and vegetable can have more direct impact on its producers.

Various studies have identified the supermarket as a critical channel giving agricultural producers, producers of high value agricultural products, such as fresh fruits and vegetables in particular, the chance to improve their welfare in the same way as exporters did. Both the exporters and the supermarket buyer values product quality and usually sets private standards for procurement. This differentiates them from the traditional market. The demand for quality often requires the producers to make additional investments. In recent years, a growing body of literature has been looking into two main questions: (i) the determinants of supermarket participation and (ii) the welfare implications of the potential exclusion of small scale producers from entering the modern retail channel. Although a great deal of effort has been devoted to improving the research methodology to remedy the constraints of the data, the results so far remain inconclusive, particularly because the studies are usually context-specific and without a theoretical model to support the underpinning relationship.

Various elements, including productivity, fixed cost and profitability, are put together to answer the question we propose below and to provide a causal-effect interpretation in a new context. That is, reflecting the heterogeneous actions that producers make in the agricultural sectors and the effect of their heterogeneous actions. We ask:

> *Does productivity affect producers' investment decisions? Does this in turn affect their profitability and welfare?*

Other than providing strong evidence to answer the question proposed in the literature, there are two important contributions that we made in chapter 4. First, a model is provided to explain the possible self-selection mechanism behind producers engaging in different investment levels and how this affects their profitability and so in general contributes to producers' household income. Second, another heterogeneous attribute which is completely ignored by the current literature is the awareness in the marketing channel to which the producers are supplying. We examine how awareness can explain the welfare and profitability differentials between the producers.

Geographical Heterogeneity

In the 1990s, theorists developed a new approach to understanding why some regions attract a disproportionate share of economic activities. The reallocation of commodities (trade), labor (migration) and firms (capital flows) across space driving the agglomeration of economic activities are studied under the heading "New Economic Geography (NEG)", or "Geographical Economics". The observed patters of agglomeration is explained by the benefit of the co-location of producers and consumers through scale economies. Despite theoretical advancement in the literature, for a long time there has been no agreed upon model specification to estimate and the literature continued to develop in various directions.

Head and Mayer tried to weave together the diverse strands of empirical work and put them into several thematic categories back in 2004. Four of the main propositions proposed are: (i) a large market potential raises local factor prices; (ii) a large market potential induces factor inflows; (iii) reduction in trade costs induces agglomeration; and (iv) the sensitivity of spatial distribution to shocks. Over the past decade since the review, empirical estimation of the NEG model continues to flourish. We provide an update of the progress made along these four propositions in Chapter 5. The contribution is to deliver an organized comparison of the work done over the last decade and point out the commonalities as well as the disagreements found in these literature.

Distance plays a crucial role in determining trade and the agglomeration of economic activities across space in all of the studies reviewed in Chapter 5. Since shipping of goods as well as migration of labor is costly, distance is a deterring factor for agglomeration. Based on Newton's universal law of gravity, the gravity model of trade suggests that the attraction between two places depends on the mass (population or economic size) of these two places and the inverse of the mutual distance. Chapter 6 revisits the gravity model and applies it to estimating the flow of illegal capital (money laundering) around the globe. Since the 911 terrorist attacks in 2001, the US has enforced their sensitivity to illegal money flows as authorities presume such money transfers can be funds supporting international terrorist and criminal activities. Heterogeneity in terms of the size of aggregate trade flows, development level and anti-money laundering policy in each country made some money laundering destinations more attractive than others. We use a special dataset that documents the illicit flows of money disguised in trade and test how the bilateral money laundering flows can be explain by the gravity model.

Summing up

Identifying and characterizing the producers who benefit from trade liberalization has far reaching implications. On the one hand, it takes us a step further from simply predicting that the trade liberalization has an impact at the micro-producer level instead of the sectoral level. On the other hand, a more accurate illustration of the heterogeneous producers supports the assumptions made and confirms the predictions made from theoretical models. As so many resources have been invested into decreasing barriers to trade and attracting foreign investment in the developing countries, more effort should be invested into understanding the uneven impact on the producers experiencing this fast transition of further integration into the global economy. In doing so, we can again justify the assumption that trade enhances welfare and we shall have a better knowledge of how to help producers through the transition.

1.3 Outline of the thesis

This dissertation consists of five chapters. The first chapter provides empirical evidence that captures the heterogeneous aspects of producers in the manufacturing and services sectors (chapter 2) in the developing countries. Against the empirical evidence from chapter 2, a theoretical model that reconciles the empirical and theoretical literature is provided by introducing fixed cost heterogeneity into the Melitz model (chapter 3). Chapter 4 reveals the producer heterogeneity in the agricultural sector and shows how it may affect producers' investment decisions, which will later affect their welfare when they are facing the opportunity to integrate into the international value chain. Chapter 5 is based on an invited paper contributing to the forthcoming Handbook of Research Methods and Applications in

Economic Geography edited by Karlsson and Andersson. The chapter provides a thorough overview of the recent studies in New Economics Geography, showing how heterogeneity in space is dealt with in the literature. Chapter 6 shows an application of the gravity model in international money laundering. The chapter was originally written as a project paper for the Public Economics course the author took during the master program, which was later published in *Applied Economics* 2013. Chapter 2 is published in a special issue of the *Journal of International Trade and Economic Development* in 2013.

To set the stage for exposing how producers in developing countries are heterogeneous, chapter 2 analyzes the differences in productivity in 15 developing countries in Latin America and for four types of firms differing particularly in their international orientation. These types are National Domestic, National Exporter, Foreign Domestic, and Foreign Exporter. The aim is to (i) provide an international comparison of the heterogeneity in the firms productivity and to depict the relationship between the development level and the average firm productivity; (ii) illustrate the empirical productivity distribution of firms in developing countries; (iii) explore the productivity differentials between the national domestic firms that do not export, the nationally owned exporters, the foreign-owned firms that do not export and the foreign-owned exporters. The most robust output among the results is the existence of a development productivity premium at the firm level, which is particularly strong for the manufacturing sector. The other important finding is the lack of productivity thresholds for viability, export activity, or multinational activity, which is suggested by the current literature. This stylized fact in particular motivates the next chapter which reconciles the discrepancy between the empirical findings shown in chapter 2 and the theoretical model in the existing literature.

Having shown the lack of a minimum productivity cutoff threshold for firm viability and lack of a clear productivity threshold between producers that differ in terms of international orientation, chapter 3 presents a model that reconciles the discrepancy between the theoretical and empirical depiction of the productivity distribution. Based on empirical evidence on fixed cost heterogeneity from Jørgensen and Schröder, we present a model for which producers are not only heterogeneous in terms of productivity but also in terms of fixed cost. By doing so, the model successfully depicts the productivity distributions of active firms in the market that resembles the empirical findings, for which a great range of productivity distributions of exporters and domestic firms overlap. At the same time, we retain all the positive attributes from the model setting as proposed by Melitz in his 2003 paper. The model continues to support that trade liberalization improves the welfare through stronger selection and increasing product varieties for the domestic consumers.

Chapter 4 explores the heterogeneous actions and the effect on the welfare impact on agricultural producers resulting from privatization, trade and foreign direct investment which initiated a rapid modernization of the food retail sector in Indonesia. There are two main questions often posed following the swift rising market penetration of supermarkets in developing countries but which had so far no clear answers. First, "are farmers with particular characteristics excluded from supplying the modern food retail channel?". Second, "do farmers benefit from supplying the supermarket food retail channel?" In addition, we highlight how the awareness in the marketing channel which the farmers are supplying would have a significant effect on their welfare. This is an interesting point which is often neglected. Our results show that the initial productivity differentials between producers significantly affect the investment decisions of the producers, which then affects their welfare. We also find that information asymmetry, that is being unaware of modern channel participation, significantly reduces the ability of producers to benefit fully from their participation in the modern food retail channel.

In Chapter 5, the latest empirical studies in Geographical Economics / New Economic Geography models are reviewed comprehensively. Four main issues are addressed in this chapter, namely: (i) how market access affects factor mobility, (ii) how market access affects factor prices, (iii) how reductions in trade costs affect core-periphery dynamics, and (iv) the shock sensitivity of the spatial distribution of economic activity. In general, the overview finds strong empirical support for the main theoretical implications of the geographical economics literature. We argue that future work needs to incorporate urban aspects in the geographical economics models, allow for heterogeneity, and focus more attention on service sectors and networks.

Having seen how geography, and distance in particular, still plays a crucial role in determining trade and the agglomeration of economic activities across space, chapter 6 applies the well-known gravity model in estimating the international flow of illegal money around the globe. Several attempts have been made in the literature to measure money laundering. However, the adequacy of these models is difficult to assess because money laundering takes place secretly and hence goes unobserved. An exception is Trade Based

Money Laundering which is a special form of trade abuse that has only recently been discovered. Trade Based Money Laundering (TBML) refers to criminal proceeds that are transferred around the world using fake invoices that under- or overvalue imports and exports. This chapter is a first test on the well-known prototype models proposed by Walker and Unger for predicting illicit money laundering flows and applying traditional gravity models familiar in international trade theory. The test is based on the dataset of Zdanowicz of TBML flows from the US to 199 countries and rejects the prototype specifications that Walker and Unger proposed. The traditional gravity model presented explains TBML flows worldwide in a plausible manner. An important determinant is trade in which TBML is hidden. Furthermore, the model suggests that criminals use TBML in order to escape the stricter anti-money laundering regulations of financial markets.

Finally, Chapter 7 recapitulates the main findings and policy implications of the research, which is followed by a brief discussion of the limitations and some suggestions for future research.

Acknowledgements

Financial support by the GAP heterogeneity project (Utrecht University School of Economics, Dutch Ministry of Economic Affairs, Agriculture and Innovation, Ecorys consultancy, and CBS Statistics Netherlands), is gratefully acknowledged.

Chapter 2 Firm Heterogeneity and Development: Evidence from Latin American countries[2]

2.1 Introduction

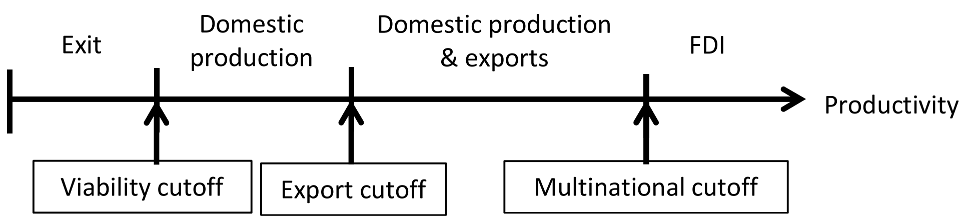
During the past decade, a new field of research analyzed the impact of trade liberalization on overall productivity growth in a country by modeling firms as heterogeneous entities that differ in terms of productivity. The workhorse model developed by Melitz (2003) suggests that aggregate productivity will increase as a result of falling trade cost. Selection effects and resource reallocation across plants of different productivity levels are the main mechanisms that induce overall productivity growth. The model predicts that the least productive firms exit the market when trade cost falls, while the most productive non-exporting firms expand production and start to export. At the same time, the existing exporters will expand their sales in the foreign market as the marginal export costs decrease (Bernard et al, 2003; Helpman et al., 2004). Moreover, the effectiveness of the resource reallocation across plants depends on the international trade involvement of a country. The model provides important new insights and partially reconciles theory with the stylized facts of international trade by allowing firm to differentiate with respect to their cost structures (Schmitt and Yu, 2001; Jean, 2002; Bernard et al, 2007; Greenaway and Kneller, 2007; Yeaple, 2005, 2009)

The heterogeneity models are characterized by productivity thresholds and cutoff values in distinguishing firms by international orientation. The first productivity threshold – the viability cutoff – indicates the minimum productivity level a firm must reach to generate non-negative profits. Firms not meeting this threshold are forced to leave the market. A second threshold – the export cutoff – indicates the minimum productivity level needed for profitable export activities. Firms must thus be sufficiently productive to become exporters. A third threshold the multinational cutoff – indicates the productivity level needed to engage in foreign direct investment (FDI) activities. Only the most productive firms are able to engage

[2] This chapter is also published in *The Journal of International Trade and Development*: 22(1) page.11-52, 2013. It is a joint work co-author with Charles van Marrewijk, the promoter of this dissertation.

in multinational activities (Helpman et al., 2004; Yeaple, 2005; Aw and Lee, 2008). As indicated in Figure 2.1, there is thus a sorting of firms into four distinct categories: (*i*) the least efficient firms exit the market, (*ii*) the low productivity firms produce for the domestic market only, (*iii*) the higher productivity firms produce for the domestic market and export their products, and (*iv*) the most efficient firms become multinationals.

Figure 2.1 Productivity thresholds and firm sorting



We analyze the empirical connections between a firm's productivity level and a firm's international organization. Our paper contributes to the heterogeneity literature in four ways.

First, we analyze a *range of developing* countries with a similar social and cultural background, rather than focusing on one, usually high income, country. The majority of studies examine the export decision of firms in developed countries[3] and only a few investigate developing countries.[4] Using the World Bank Enterprise Survey (WBES) data we investigate 15 Latin American countries.[5]

Second, we take the heterogeneity of firms seriously by analyzing differences in *distributions* rather than focusing on differences in averages. In contrast, some empirical studies still compare average differences in firm performance among sub-groups, such as exporters versus non-exporters or domestic firms versus foreign-owned firms (Aw, Chung and Roberts, 2000; Tomiura, 2007). In doing so, not only the information from the micro-firm level data is

---

[3] Belgium: Mayer and Ottaviano, 2008; Germany: Wagner and Bernard, 1997, Wagner and Vogel, 2010; Sweden: Greenaway et al, 2005; USA: Bernard and Jensen, 1999, 2004; Bernard, Jensen, Redding and Schott, 2007. See also Wagner (2007) for an extensive survey of the empirical research on firm heterogeneity.
[4] Chile: Alvarez, R. and R.A. López, 2005; Colombia: Robert and Tybout, 1997; Colombia, Mexico, and Morocco: Clerides, et al., 1998; Indonesia: Sjöholm , 2003; Sub-Saharan Africa: Van Biesebroeck, 2005.
[5] Research on Chile by Blyde and Iberti (2010, p.13) suggests that: "there is scope for increasing aggregate productivity in developing countries via the reallocation of resources from low to high productivity plants." We examine whether there is indeed scope for such a resource reallocation mechanism.

overlooked, but also the most important messages from the firm heterogeneity models are neglected. We complement our analysis with detailed regression analysis.

Third, our data consists of both manufacturing sectors and *services* sectors. This allows us to determine if the conclusions regarding firm type and productivity derived for manufacturing sectors also hold for services sectors. The answer will be affirmative regarding the foreign-ownership productivity premium but negative regarding the export productivity premium.

Fourth, and most importantly, we analyze in detail the productivity performance of foreign-owned firms relative to other types of firms by investigating *all four types* of firms in the market orientation – ownership dimensions, rather than lumping all foreign-owned firms together. As depicted in Figure 2.1, the literature so far analyzes only three types of firms, namely nationally-owned domestic firms, nationally-owned exporting firms, and foreign-owned firms. Since firm typology is along two dimensions, namely export orientation and ownership, we can in principle identify four types of firms (see Table 2.1): National Domestic firms, National Exporters, Foreign Domestic firms, and Foreign Exporters. The empirical literature so far finds (see below) that exporting firms are more productive than domestic firms while foreign-owned firms are more productive than national firms. One would therefore expect that National Domestic firms are the least productive and Foreign Exporters are the most productive firms. Whether this is true and whether National Exporters are more or less productive than Foreign Domestic firms are open questions.

Table 2.1 Firm typology

	Ownership	
Market orientation	National	Foreign
Domestic	ND – National Domestic	FD – Foreign Domestic
Export	NE – National Exporter	FE – Foreign Exporter

Regarding the fourth contribution we note that the exporters' superior performance (in terms of productivity, size, length of survival and wage paid) is well-known and robust (Handoussa et al., 1986; Chen and Tang, 1987; Tybout and Westbrook, 1995; Aw and Hwang, 1995; Aw and Betra, 1998, 1999; Bernard and Jensen, 1999; Tybout, 2000). The impact of foreign-ownership is less independently identified in this research scope. A foreign-owned firm is different from an exporting firm. A foreign-owned firm is selected by foreign profit seeking investors[6], while exporting activities are initiated by the firms themselves through self-selection (Aw, Chen, and Roberts 2001; Clerides, Lach and Tybout, 1998; Bernard and Jensen, 1999). It is well documented that foreign investment not only brings financial support but also advanced technology. Both lead to higher productivity due to higher capital intensity or R&D investment in these firms (see e.g. Haddad and Harrison, 1993 and Sinha, 1993). We distinguish between all logically possible types of firms as given in Table 2.1 and examine their productivity differences.

Section 2.2 discusses the data and methods we use. Section 2.3 characterizes the data along three dimensions (development, organization structure, and size). Section 2.4 provides formal distribution tests and regression estimates to control for simultaneous effects. Section 2.5 provides robustness checks when we use a different sector classification, when we base our productivity measure on value added per worker rather than sales per worker, and when we control for capital intensity. Section 2.5 also analyzes the impact of foreign ownership intensity. Section 2.6 concludes. Throughout the paper we summarize our findings in stylized facts.

Table 2.2 Number of observations across countries and sectors

#	Country	Count	%	#	Sectors	Count	%
1	Argentina	1,063	9.73		Manufacturing	7,202	66
2	Bolivia	613	5.61	1	Food	1,727	15.80
3	Chile	1,017	9.30	2	Garments	1,166	10.67
4	Colombia	1,000	9.15	3	Textiles	725	6.63
5	Ecuador	658	6.02	4	Machinery & Equipment	451	4.13
6	El Salvador	693	6.34	5	Chemicals	1,056	9.66
7	Guatemala	522	4.78	6	Electronics	89	0.81
8	Honduras	436	3.99	7	Non-Metallic minerals	348	3.18
9	Mexico	1,480	13.54	8	Other Manufacturing	1,640	15.00
10	Nicaragua	478	4.37		Services	3728	34
11	Panama	604	5.53	9	Retail	1,561	14.28
12	Paraguay	613	5.61	10	Information Technology	494	4.52
13	Peru	632	5.78	11	Other Services	964	8.82
14	Uruguay	621	5.68	12	Construction	638	5.84
15	Venezuela	500	4.57	13	Wholesale	71	0.65
	Total	10,930	100		Total	10,930	100

[6] "Hence cross-sectional studies may suffer from simultaneity bias because MNCs are attracted to profitable sectors, and negative spillover effect may occur in the short run because MNCs siphon off domestic demand and/or bid away high quality labor when they set up shop in the host country (Aitken and Harrison forthcoming)." Tybout (2000) P.37.

2.2 Data and methods

We use data provided by the World Bank Enterprise survey (WBES). The 15 Latin American countries studied here are countries sampled in the 2006 survey, consisting of: Argentina, Bolivia, Chile, Colombia, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, and Venezuela. The core survey consists of 10,930 micro firm-level observations. Firms were selected using stratified random sampling. The three strata used for each country are: region, sector, and establishment size. The stratification sampling ensures that the database consists of observations from different subdivisions of the firm population. According to this stratification methodology, the larger the country and the greater the sector, the more firms will be sampled, see the World Bank website www.enterprisesurveys.org/Methodology for details. Table 2.2 gives the number of observations by country and sector. There are eight manufacturing sectors and five services sectors. There are 66% manufacturing firms and 34% services firms. Food, garments, other manufacturing and retail are the biggest sectors in the Latin American countries under study. We focus attention on 11 sectors only and put less emphasis on the electronics and wholesales sectors, which both have fewer than 100 observations. Moreover, the wholesale sector is only separately recorded in Panama, while other countries recorded them in the retail sector.

Table 2.3 Country development classification, year 2006

Country	Country code	GDP per capita	GDP per capita PPP	Development level
Mexico	MEX	6,414	13,070	
Chile	CHL	5,870	12,599	
Argentina	ARG	8,699	11,623	Middle-High
Venezuela	VEN	5,401	10,721	
Uruguay	URY	7,522	10,075	
Panama	PAN	4,737	9,799	
Colombia	COL	2,789	7,589	
Ecuador	ECU	1,664	7,055	Middle-Middle
Peru	PER	2,502	6,731	
El Salvador	SAL	2,515	5,902	
Guatemala	GTM	1,811	4,178	
Paraguay	PAR	1,392	3,990	
Bolivia	BOL	1,145	3,857	Middle-Low
Honduras	HND	1,353	3,419	
Nicaragua	NIC	865	2,383	

Source: World Development Indicator (WDI); GDP per capita in constant 2000 US $; GDP per capita PPP in constant 2005 international $.

Ten of the fifteen Latin American countries in our study have an income level above the world average for middle income countries ($4,940 GDP per capita in PPP), but much lower than the world average for high income countries ($33,184 GDP per capita in PPP). The other five countries have an income level below the world average for middle income countries, but still higher than the world average for low income countries ($945 GDP per capita in PPP).[7] We categorize the countries in the sample relative to each other and the above averages into three groups of middle-high (MH), middle-middle (MM), and middle-low (ML) income countries, see Table 2.3. The classification is the same whether we correct for purchasing power parity or not. The income threshold used to distinguish the three development groups are 5,000 and 10,000 for GDP per capita in PPP terms and 2,000 and 5,000 for GDP per capita in constant dollar terms.

2.2.1 Productivity

The most important variable under study is productivity. Without a direct measure of productivity, we compute the sales per worker as our productivity measure, which is also used in other research (Wagner and Vogel, 2010). A more comprehensive productivity measure such as total factor productivity (TFP) is not use here because the time dimension required for computing TFP is lacking. Bartelsman and Doms (2000) point out that heterogeneity in output per worker productivity is accompanied by similar heterogeneity in total factor productivity. As part of our robustness check in section 2.5 we also compute value-added per worker as an alternative productivity measure for the manufacturing subset of the data (as this is not available for services firms). The analytic results are very similar. In calculating productivity we used *total worker employed,* which is sum of the permanent and temporary worker in the data (Section 2.5 of the paper provides a robustness check using only permanent workers). The *sales* value originally recorded in the local currency units (LCU) was converted to international currency, the US dollar, using the official exchange rate of the sample year (period average; WDI 2006). Among the Latin American countries, Ecuador, El Salvador and Panama's sales value remained the same either because the dollar is used in the country or the local currency is fixed (pegged) at parity with the US dollar. All firms with sales data are included (90% of the surveyed firms). The number of observations

[7] The 2010 World Bank country classification is based on three GNI per capita thresholds using the Atlas method: $1005, $3975 and $12,275; leading to the following income groups: low income, lower middle income, upper middle income, and high income.

decreased to 9,835 as the sales value is lacking due to "respond refusal" (498 obs.) or "don't know" (587 obs.) and some missing values for the aggregated total labor (31 obs.).

$$(1) \qquad \theta_{ijk} = \frac{\ln(\overline{\theta}_{ijk}) - \min_{i,k} \ln(\overline{\theta}_{ijk})}{\max_{i,k} \ln(\overline{\theta}_{ijk}) - \min_{i,k} \ln(\overline{\theta}_{ijk})}, \quad where \quad \overline{\theta}_{ijk} = \frac{sales_{ijk}}{labor_{ijk}}$$

The scale and shape of productivity measures differ per sector, see Mayer and Ottaviano (2007, Table 7). The remainder of this paper therefore uses *normalized productivity* θ_{ijk} for firm i in sector j and country k, as given in equation (1). This measure scales the log of productivity to a scale from zero to one for each sector. This allows for a comparison of normalized productivities between countries and different sectors. Firms with productivity that appear above or below four standard deviations from the mean in each sector are considered as outliers. The minimum and maximum productivity in each sector are thus the productivity values four standard deviations from the mean. Around 1.7% observations are dropped as outliers, ranging from 0.24% for machinery and equipment to 3.75% for garments.

Table 2.4 Normalized productivity by country

Development	Country	country code	mean	variance	min	median	max
Middle-High	Mexico	MEX	0.428	0.026	*0.000*	0.411	*1.000*
	Chile	CHL	0.477	0.019	0.020	0.467	*1.000*
	Argentina	ARG	0.490	0.019	0.032	0.482	*1.000*
	Venezuela	VEN	0.399	0.021	0.027	0.390	0.967
	Uruguay	URY	0.439	0.021	0.000	0.438	0.955
Middle-Middle	Panama	PAN	0.433	0.027	*0.000*	0.414	*1.000*
	Colombia	COL	0.383	0.016	0.052	0.381	0.985
	Ecuador	ECU	0.446	0.018	0.020	0.443	0.897
	Peru	PER	0.442	0.023	*0.000*	0.449	*1.000*
	El Salvador	SAL	0.362	0.022	0.011	0.353	*1.000*
Middle-Low	Guatemala	GTM	0.345	0.020	0.009	0.344	0.828
	Paraguay	PAR	0.344	0.023	*0.000*	0.341	*1.000*
	Bolivia	BOL	0.333	0.024	*0.000*	0.321	0.901
	Honduras	HND	0.340	0.024	*0.000*	0.345	0.960
	Nicaragua	NIC	0.280	0.029	*0.000*	0.263	0.982

The summary statistics give us an idea of the differences in productivity distribution for each sector (appendix A2.2). For example, the normalized productivity for the garments sector appears gives the lowest median while the wholesales sector has the greatest variance. Table 2.4 reports the summary statistics of the normalized productivity for each country. Countries with a maximum of one have at least one superstar firm with the highest productivity in some sector (holds for: Argentina, Chile, El Salvador, Mexico, Panama, Paraguay and Peru). In contrast, for countries with minimum of zero have the least productive firm in some sector (holds for: Bolivia, Honduras, Mexico, Nicaragua, Panama Paraguay, Peru and Uruguay). Take the food sector as an example; the most productive firm is located in Panama and the least productive in Peru. Since all firms in the sector are scaled relative to the best and the worst performing firms in that sector, the variance reported suggests that Peru is also the country with the highest productivity variation in the food sector (see appendix A2.2.3).

The same normalization procedure is performed with the sectors identified differently (summary statistics in appendix A2.2.2). In the WBES data, an additional four-digit ISIC code is recorded according to the main output product that generated the largest proportion of the firms' annual sales. This additional sector classification is coded for manufacturing firms only, except for firms in Venezuela. The productivity measure normalized by this ISIC-code is later referred as the *ISIC-normalized productivity*.

2.2.2 Export status and ownership characteristics

As shown in Table 2.1 we identify four types of firms along two dimensions: export status and ownership characteristic. To operationalize our typology we use a cutoff value of 10 percent in both cases. A firm is therefore classified as an exporter if it exports at least 10 percent of its sales. Similarly a firm is classified as a foreign firm if it is at least 10 percent foreign-owned. Our cutoff percentages therefore mimic the standard International Monetary Fund Balance of Payments procedures.[8] Using this classification there are 1,562 (15%) exporters and 8,841 (85%) domestic firms. Similarly, there are 9,304 (89.4%) national and 1,099 (10.6%) foreign firms, see Table 2.5. The majority of firms (77.9%) is nationally owned and sells to the domestic market. They are followed in abundance by National Exporters (11.5%), Foreign Domestic firms (7.1%), and Foreign Exporters (3.5%). The percentages and ordering are similar for the manufacturing and services sectors separately, except for the fact that the share of National Exporters is lower than the share of Foreign Domestic firms for the services sectors (5.9% versus 10.8%), see also appendix A2.3.1. The total share of exporters (11.5%) is higher than the 4.2% for the US reported in Bernard et al.

[8] The IMF (2009) definition of "a direct investment relationship" is a shareholder who owns 10 percent or more of the ordinary shares or voting power and who resides in another country.

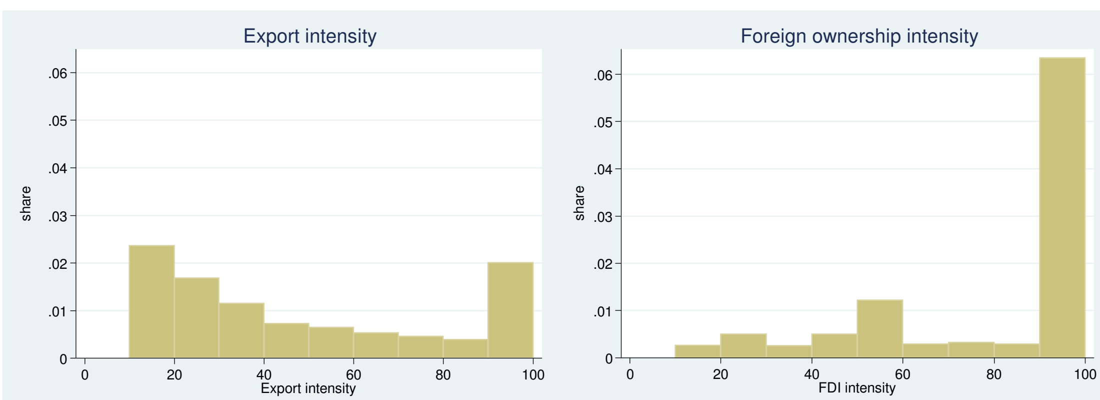
(2003) or the 4.65% exporters plus outsourcing-exporters reported for Japan in Tomiura (2007). Considering the relative size of the economy (the numbers of firms sampled in each country), Argentina, El Salvador and Guatemala are the countries with the highest percentage of exporters (over 20%); while Venezuela is the country with the lowest percentage of exporters (3%), see appendix A2.3.3.

Table 2.5 Summary of firm types in manufacturing and services sectors

Category	All firms	%	Manufacturing	%	Services	%
National Domestic	8,107	77.9	5,261	76.2	2,846	81.3
National Exporter	1,197	11.5	989	14.3	208	5.9
Foreign Domestic	734	7.1	355	5.1	379	10.8
Foreign Exporter	365	3.5	296	4.3	69	2.0
Total	10,403	100	6,901	100	3,502	100

Export and FDI intensity varies greatly between firms. Among the 1,562 exporters, about one third export 10 to 20 percent of their output, while another one third export between 20 to 59 percent of their output. The last one third export more than 60 percent of their output, see Figure 2.2 (left). The export intensity among domestic firms and foreign-owned firms is slightly different. Most nationally-owned firms export at a lower export intensity than their foreign-owned competitors, see A3.2.

Figure 2.2 Export intensity and foreign ownership intensity



The distribution of the foreign-ownership intensity is quite different, see Figure 2.2 (right). Among the 1,099 foreign-owned firms, there are 604 firms (55%) with foreign ownership above 90 percent while the remaining 45% have foreign ownership intensity ranging in between 10 to 90 percent range. This distinction turns out to be important, so we will get back to it in section 2.5 when we discuss the impact of foreign-ownership intensity. In that section we will label the 604 firms with foreign ownership above 90 per cent as Foreign-high and the remaining foreign firms as Foreign-medium.[9]

2.3 Development, size, and firm type

We first discuss the general shape of the productivity distributions and then investigate the connections with development, firm type, and size. The discussion in this section is suggestive of possible relationships. These suggestions are analyzed in the next section.

2.3.1 Productivity distributions

Figure 2.3 shows the kernel density estimates of normalized productivity of different sectors for the 15 countries combined. The left panel depicts the manufacturing sectors. All distributions are nicely bell-shaped, although the electronics sector has a fatter right-tail. The right panel depicts the services sectors. These too are nicely bell-shaped, with the exception of the wholesale sector, which is more lumpy and shifted to the right. As suggested by these figures, the electronics sector and the wholesale sector have a higher mean and variance than the other sectors, see Table A2.2.1 in the appendix. The distribution per sector for each individual country is available from the authors upon request.

The common feature of the productivity distributions in Figure 2.3 for all sectors is the absence of a threshold level for firm survival. Firms of various productivity levels remain viable in the market (as proved by their existence) at the same time, in contrast to the Melitz model. This result may be caused by other heterogeneity characteristics of firms, such as (human) capital intensity, efficient use of capital, or fixed cost heterogeneity.[10]

Stylized fact I: There is no clear productivity level threshold for firm viability (existence).

---

[9] The domestic firms could then alternatively be referred to as Foreign-low.
[10] Chapter 3 suggests that the empirical regularity of the absence of a viability threshold level can be explained easily by allowing firm heterogeneity in both fixed and marginal costs.

Figure 2.3 Productivity distributions for different sectors, kernel density estimates

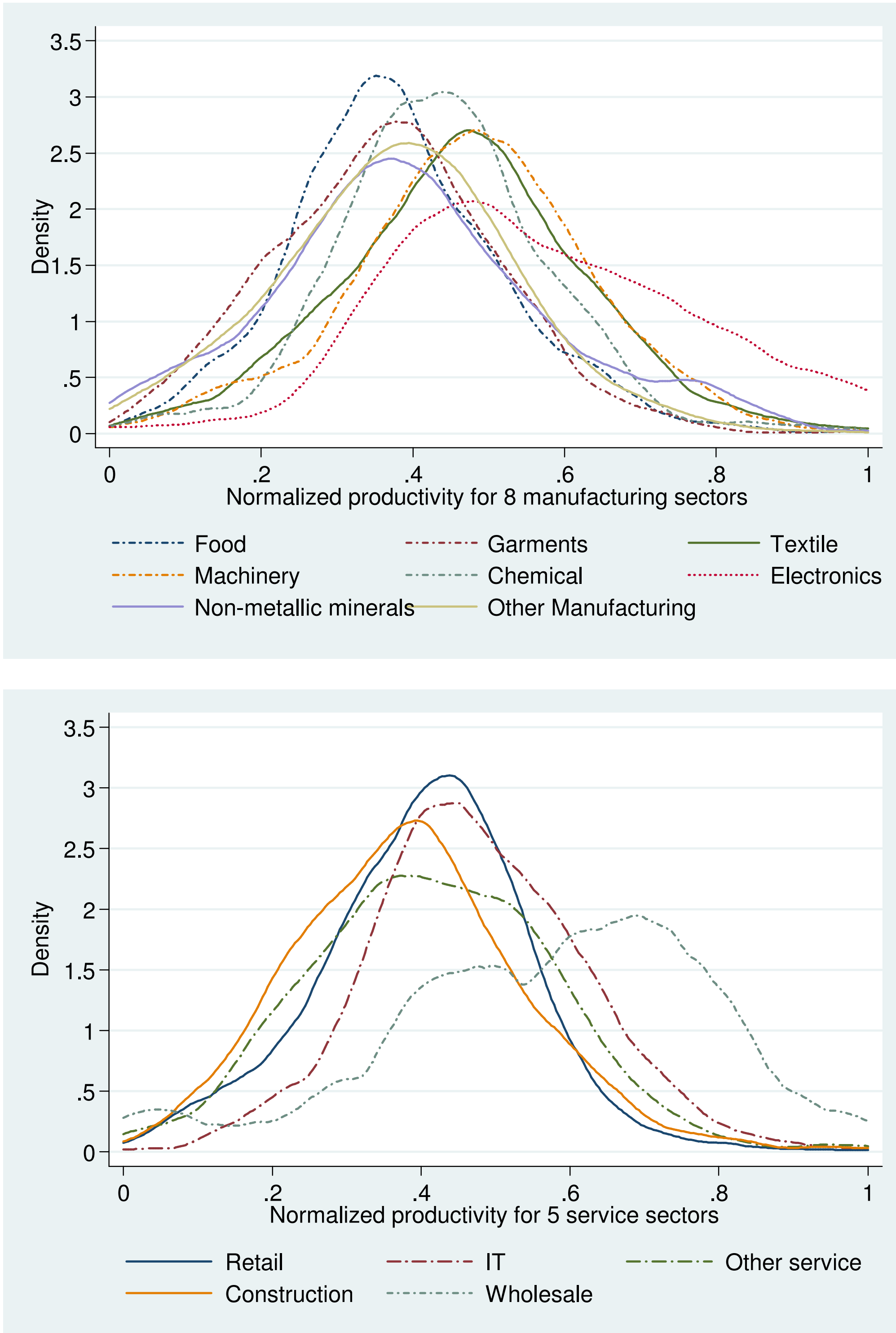
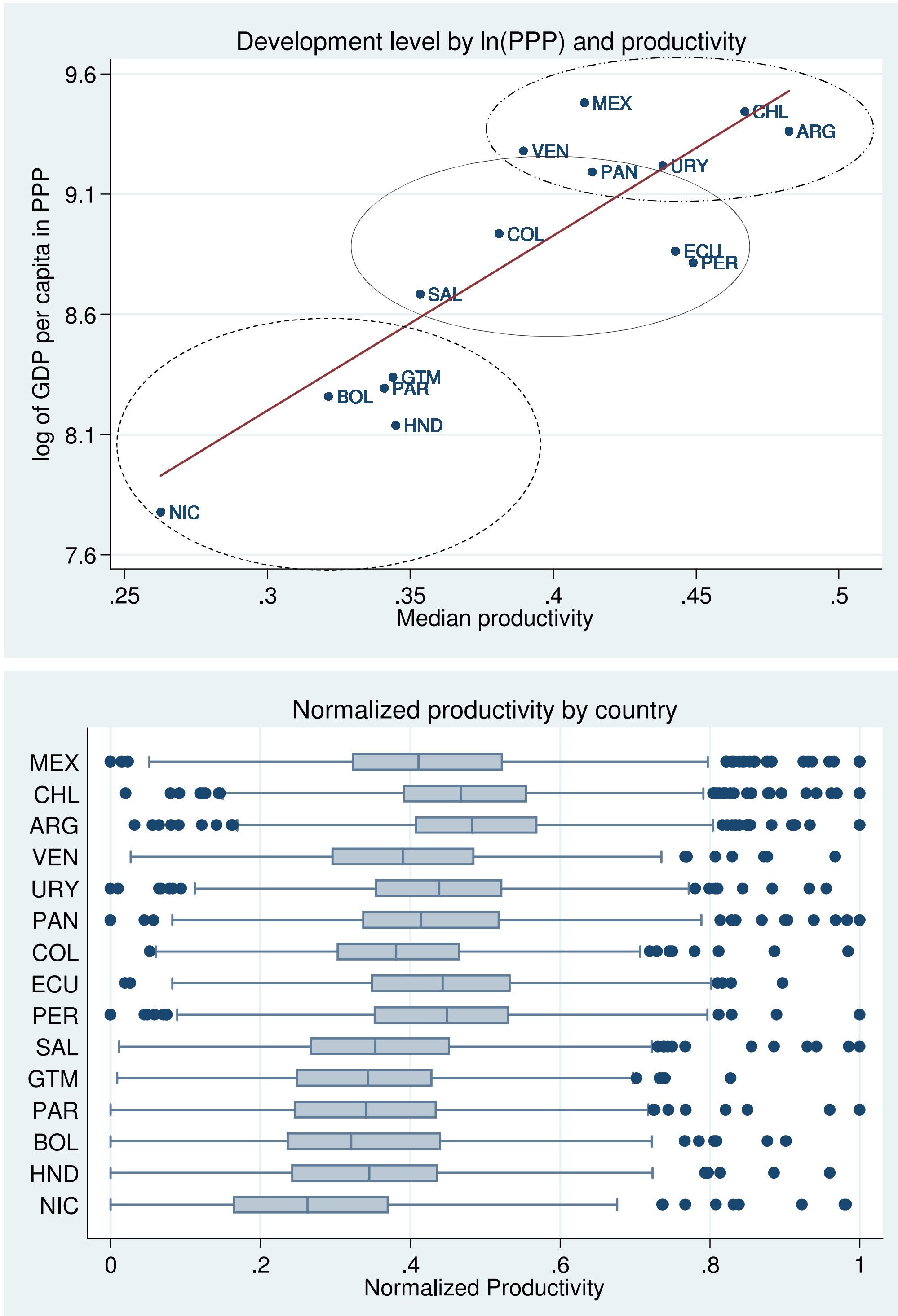


Figure 2.4 Productivity and development per country



2.3.2 Productivity and development

Figure 2.4 depicts the relationship between firm productivity and a country's development level. The left panel shows the median normalized productivity for a country (for all sectors combined) on the horizontal axis and the log of GDP per capita (PPP) on the vertical axis. A clear positive association between these variables is found (R^2=0.69), with Chile and Argentina in the upper-right corner, Nicaragua in the lower-left corner, and other countries in between.[11] This suggests that more developed countries (as measured by a country's income level) are populated with higher productivity firms. This suggestion is confirmed by the box plots in the right panel of Figure 2.4, which depict the 25th, 50th, and 75th percentile of the distribution for each country to get an idea of the range of the distribution.[12]

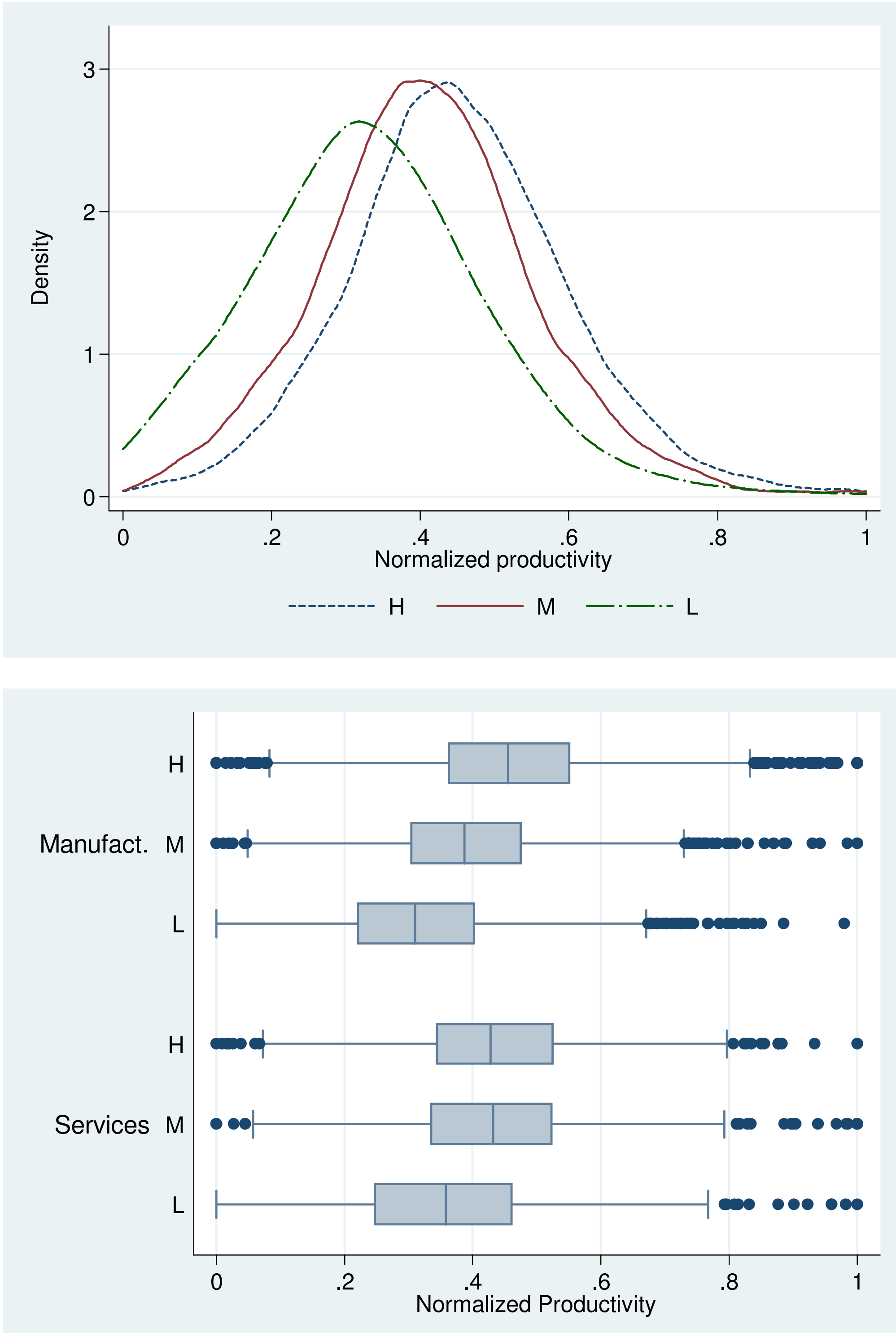
To clarify the role of development in the productivity distribution, the left panel of Figure 2.5 depicts the kernel density estimates for the three groups of development while the right panel shows the box plots for manufactures and services per group separately. The productivity distribution of the middle-low income group (dashed line) is located to the left of the middle-middle (solid line) and middle-high (dotted-dashed) income group. This relationship exists also at the sector level, where the ordering from low to medium to high development is most vivid in the garments, textile, chemicals, and other manufacturing sectors. The positive relationship between development level and productivity is slightly less distinct in the services sectors (right panel of Figure 2.5).

Suggestion I: there is a positive association between productivity and development level.

---

[11] A similar figure is generated by using mean productivity instead of median productivity; the R^2 rises to 0.76.
[12] The box plots provide a summary of the productivity distribution across countries; the median is represented by the vertical bar in the middle of each box; the upper and lower limits of the boxes represented the 25 (Q1) and 75 (Q3) quartiles of the productivity distribution; productivity values outside 1.5 times the interquartile range (difference between Q3 and Q1) are shown by dots outside the horizontal whiskers.

Figure 2.5 Productivity and development per country group



H = middle-high income countries; M = middle-middle income countries; L = middle-low income countries

2.3.3 Productivity and firm type

Next, we compare the productivity distributions of the four types of firms listed in Table 2.1 based on export status and foreign ownership. Figure 2.6 shows the kernel density estimates for all sectors and countries (left panel) and box plots for manufactures and services sectors separately (right panel). The literature suggests that (i) exporting firms are more productive than domestic firms and (ii) foreign-owned firms are more productive than national firms. Both suggestions are confirmed by the information in Figure 2.6 since:

(i.a) National Exporters tend to be more productive than National Domestic firms

(i.b) Foreign Exporters tend to be more productive than Foreign Domestic firms

(ii.a) Foreign Domestic firms tend to be more productive than National Domestic firms

(ii.b) Foreign Exporters tend to be more productive than National Exporters.

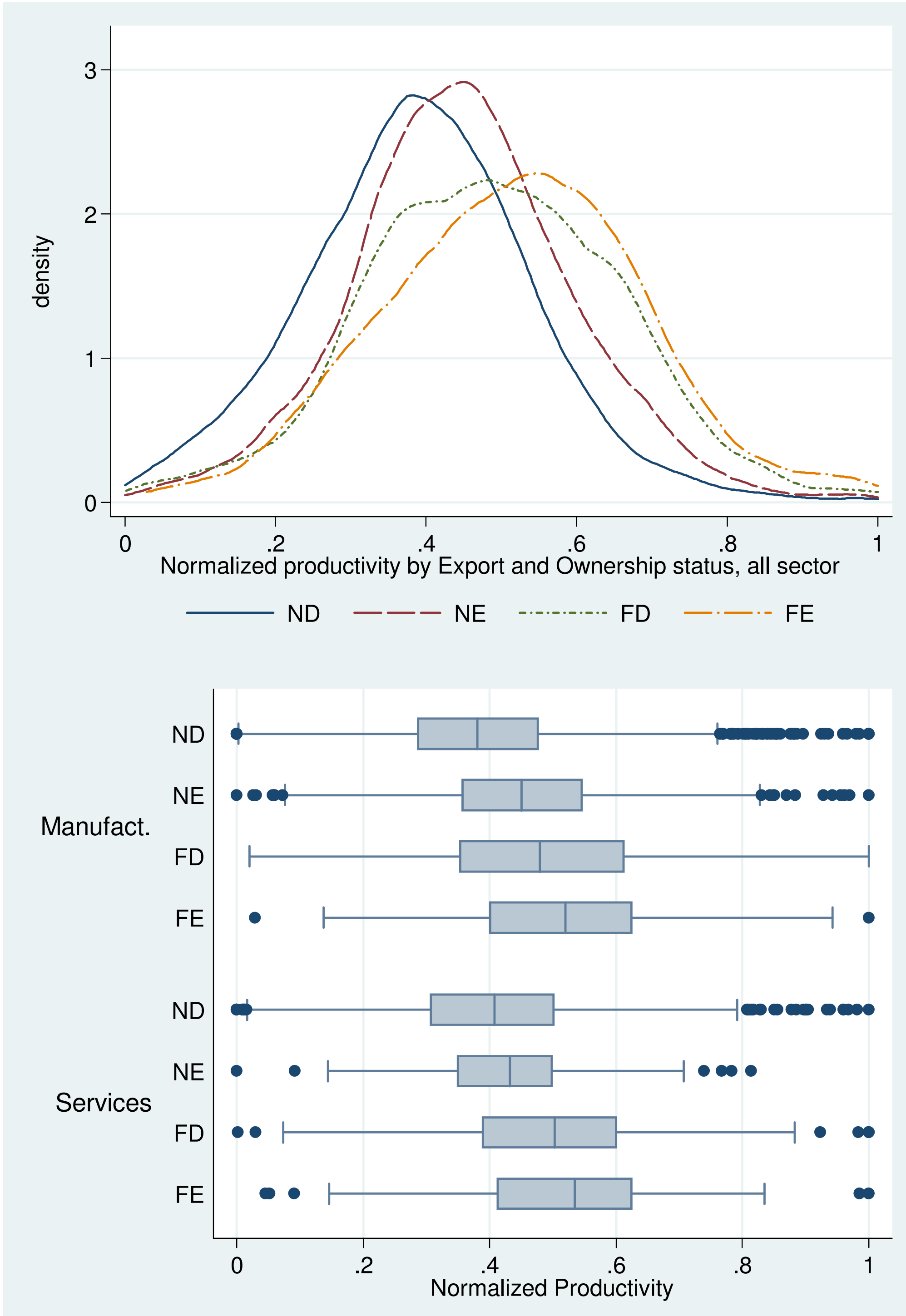
Figure 2.6 also indicates that Foreign Domestic firms tend to be more productive than National Exporters, suggesting that from a productivity perspective the foreign-ownership premium is more important than the exporter premium.

The ranking of productivity and firm types per sector largely coincide with the above observations. It is clear from the figures there and the above information that there is an enormous overlap in terms of productivity for different firm types and not the sorting of firm types in exclusive productivity ranges illustrated in Figure 2.1. Firms with the same productivity level can therefore belong to different types of firms, such that there is no productivity threshold for exporting or foreign-ownership status. This observation and the productivity ranking above brings us to a stylized fact and a suggestion.

Stylized fact II: There are no clear productivity thresholds sorting firms in different types.

Suggestion II: The productivity ranking (from low to high) appears to be: National Domestic firms, National Exporters, Foreign Domestic firms, Foreign exporters.

Figure 2.6 Normalized productivity by exporting and ownership status



ND =National Domestic; NE =National Exporter; FD =Foreign Domestic; FE =Foreign Exporter[13]

13 Note that the number of observations for the Foreign Exporter firms is small, making it more difficult to draw conclusions if this firm type is subdivided into three size classes.

2.3.4 Productivity and size

Empirical studies have shown that foreign-owned firms and exporting firms are usually larger in size (usually identified by employment level, see Bernard et al., 2007). This size regularity is also found in our data. The superior performance of exporters and foreign-owned firms shown above should therefore be studied with caution since our productivity measure does not take the scale effect of absolute firm size into consideration. Table 2.6 reports the quantitative productivity characteristics of the four types of firms. The ranking order associated with the average productivity and the 1st, 2nd and 3rd productivity quartile of the firm types in the top part of the table is in accordance with Suggestion II of subsection 2.2. A similar ranking of firm types is found in the bottom part of the table regarding firm size, with some reversals only with respect to National Exporters and Foreign Domestic firms.

Figure 2.7 illustrates the simultaneous impact of firm size and firm type on productivity by distinguishing between three different size classes within each firm type. The left panel shows that for each firm type (with the exception of Foreign Exporters) productivity tends to be higher for larger size classes. The right panel, similarly, shows that within each size class the firm type ordering is in accordance with Suggestion II, with the exception of the reversal of Foreign Domestic firms and Foreign Exporters for the largest size class.

Suggestion III: Productivity is positively related to firm size.

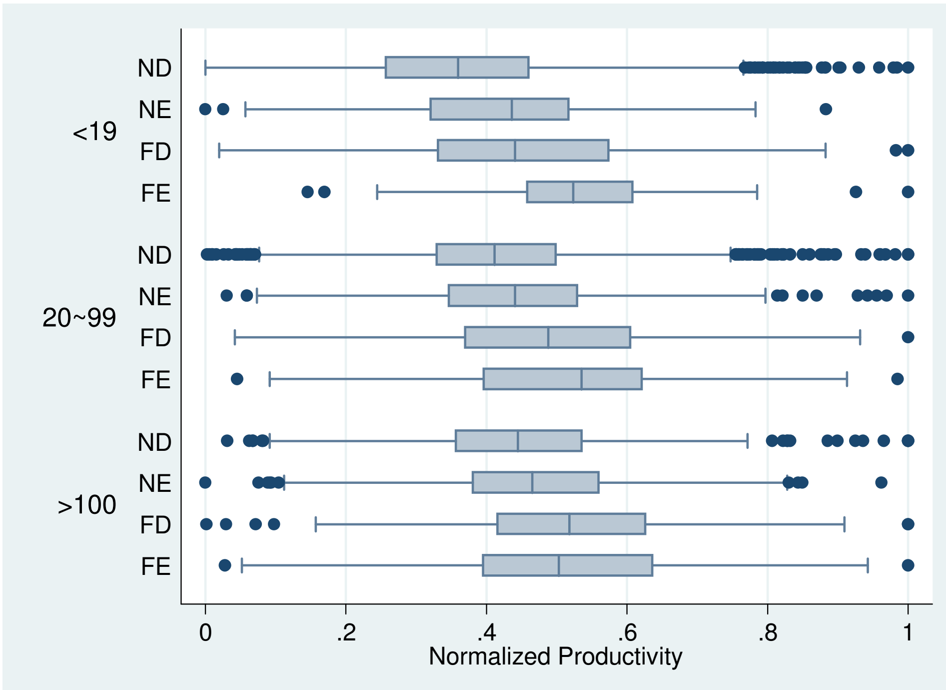
*Table 2.6 Productivity and size characteristics of firm types**

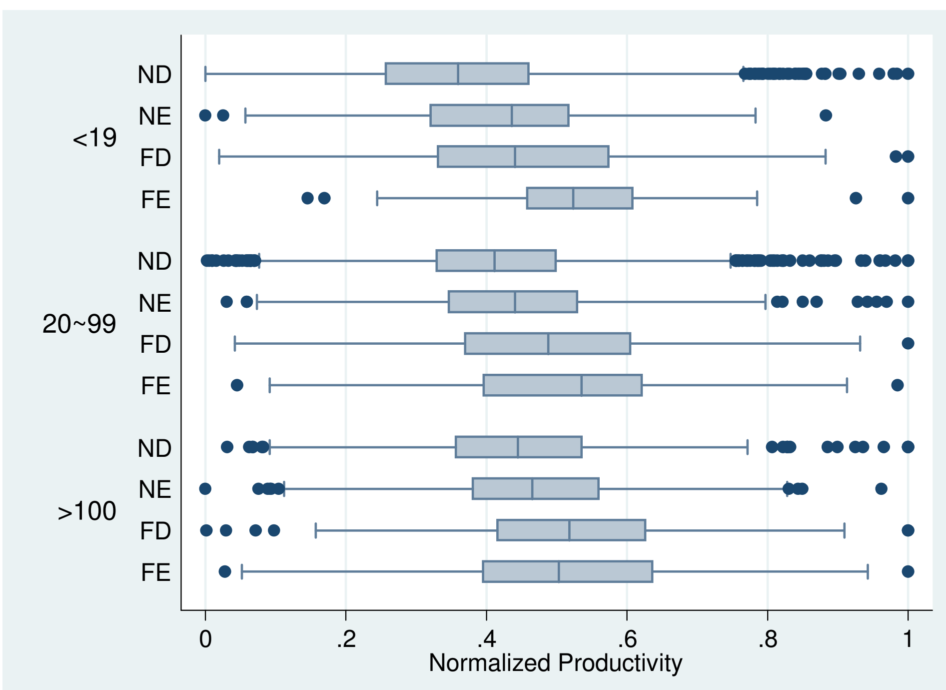
	# obs	Mean	St dev	Min	Q1	Q2	Q3	Max
Normalized productivity								
National Domestic	7,135	0.39	0.15	0.00	0.29	0.39	0.49	1.00
National Exporter	1,086	0.45	0.15	0.00	0.36	0.45	0.54	1.00
Foreign Domestic	659	0.49	0.17	0.00	0.37	0.49	0.61	1.00
Foreign Exporter	341	0.52	0.17	0.02	0.40	0.52	0.62	1.00
Size, by number of workers								
National Domestic	8,085	65.6	330.3	0	10	20	46	18,000
National Exporter	1,193	227.7	735.2	3	27	68	180	19,500
Foreign Domestic	732	*197.3*	*686.9*	4	*19*	*51*	*136*	*14,542*
Foreign Exporter	365	500.3	972.1	5	63	180	500	*9,000*

**Italic* indicates reverse ordering of the firm type compared to the productivity ordering

Clearly, no single dimension (size, firm type, or development level) can completely explain differences in productivity distributions. Since the dimensions may be correlated with one another we need to control for other firm characteristics before we can draw any definite conclusions. The next section therefore tests the suggestions above.

Figure 2.7. Firm size, firm type, and productivity





ND = National Domestic; NE = National Exporter; FD = Foreign Domestic; FE = Foreign Exporter

2.4 Distribution tests and regression analysis

The suggestions in section 2.3 focused on differences in certain distribution characteristics, such as mean or median, to argue for differences in terms of firm type, size of firm, and development level. This section starts with some formal distribution tests, that is analyzing differences in the whole distribution rather than focusing on some distribution characteristic, before continuing with a formal analysis that controls for simultaneous effects.

2.4.1 Distribution tests

The majority of distribution tests reported in this subsection are based on Kolmogorov-Smirnov (KS; Kolmogorov, 1933; Smirnov, 1939), but other tests lead to similar results. To test whether the underlying productivity *distribution* differs per country, sector, firm size, and firm type, we perform pair wise KS tests within each dimension. The p-value for the probability that the two distributions are the same is reported in appendix A2.4.1-A2.4.4. We reject the null hypothesis if the p-value is less than 0.05. Using this criterion 87.6% of the country pairs and 89.7% of the sector pairs reject the null hypothesis, indicating that the distributions generally differ per country and sector. Moreover, the null hypothesis is rejected for all firm type categories and firm size categories, indicating that distributions differ per firm type and firm size.

Table 2.7 KS test on similarity of productivity distribution for firm types (p-values)

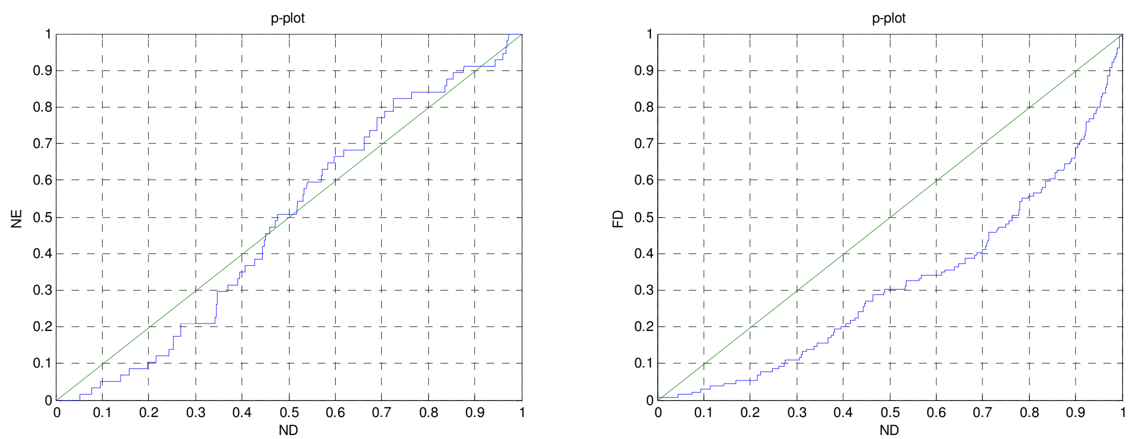
	ND v NE	FD v FE	ND v FD	NE v FE	NE v FD	ND v FE
Manufacturing sectors						
Food	0.00**	0.06*	0.00**	0.24	0.11	0.00**
Garments	0.00**	0.01**	0. 09*	0.00**	0.75	0.08*
Textile	0.02**	0.57	0.00**	0.17	0.11	0.00**
Machinery	0.00**	0.71	0.16	0.31	0.73	0.03**
Chemical	0.00**	0.05**	0.00**	0.00**	0.01**	0.00**
Electronics^{+}	0.47	0.54	0.47	0.13	0.45	0.07*
Non-metallic mineral	0.10*	0.86	0.02**	0.04**	0.05**	0.03**
Other manufacturing	0.00**	0.05**	0.00**	0.00**	0.11	0.00**
Service sectors						
Retail	0.47	0.03**	0.00**	0.00**	0.00**	0.00**
Information technology (IT)	0.19	0.98	0.00**	0.01**	0.00**	0.00**
Other Services	0.31	1.00	0.00**	0.02**	0.00**	0.02**
Construction	0.10**	0.63	0.00**	0.45	0.58	0.04**
Wholesale^{+}	0.86	0.81	0.70	0.90	0.86	0.56

$^{+}$Sectors with sample size < 100; ** significant at 5%; * significant at 10%; ND = National Domestic; NE = National Exporter; FD = Foreign Domestic; FE = Foreign Exporter

Table 2.7 reports the p-value statistics of the KS test for firm types per sector. The four firm types create six comparison pairs. In most cases the underlying distributions are different, but note that the manufacturing sectors have a larger fraction of significantly different underlying distributions than the services sectors.

Figure 2.8 visualizes the test by using the percentile-percentile (p-p) plot for the Other Services sector as an example. The left panel shows the p-p plot for National Domestic versus National Exporters. Since the plot is fairly close to the diagonal we cannot reject the hypothesis that the two draws are from the same underlying distribution (the p-value is 0.31, see Table 2.7). The right panel shows the p-p plot for National Domestic versus Foreign Domestic firms. This plot deviates quite far from the diagonal, such that we conclude that the two underlying distributions are different (the p-value is 0.00, see Table 2.7).

Figure 2.8P-P plots[14] for Other Services sector; left panel: not rejected; right panel: rejected



As an extension of the distribution test, we applied the Mann-Whitney-Wilcoxon (MWW) test to determine if the median is different for the various firm types, see Table 2.8. The results correspond closely to the KS test and confirm the ranking order of firm types listed in Suggestion II, with the insignificant difference between Foreign Domestic firms and Foreign Exporters as the only exception. We summarize our findings as follows:

---

[14] A P-P plot is a two-dimension probability plot for assessing how closely two data sets agree. This is done by plotting two cumulative distribution functions against each other. Thus, for input z the output is the pair of numbers giving the percentages that the distributions have below z: $((F_1(z), F_2(z)) = (P_1(X_1 \leq z), P_2(X_2 \leq z))$. The diagonal in the p-p plot is the comparison base that shows when the percentages of the two cumulative distribution functions are the same: $P_1(X_1 \leq z) = P_2(X_2 \leq z)$. The closer the p-p line is to the diagonal line, the more certain we are that the two samples have the same underlying distribution.

Stylized fact III: Productivity distributions differ in general per country, sector, firm type, and firm size.

Suggestion IV: the ranking order of firm types given in Suggestion II holds in general; the equality of Foreign Domestic and Foreign Exporter productivity for the services sectors is the only exception

Table 2.8 Equality test on distribution and median summary (p-value)

	ND v NE	FD v FE	ND v FD	NE v FE	NE v FD	ND v FE
Manufacturing sectors						
Distribution (KS)	0.00**	0.05**	0.00**	0.00**	0.00**	0.00**
Median (MWW)	0.04**	0.24	0.00**	0.00**	0.00**	0.00**
Service sectors						
Distribution (KS)	0.00**	0.07*	0.00**	0.00**	0.09*	0.00**
Median (MWW)	0.06*	0.21	0.00**	0.00**	0.00**	0.00**

** significant at 5%; * significant at 10%; ND = National Domestic; NE = National Exporter; FD = Foreign Domestic; FE = Foreign Exporter

2.4.2 Regressions

We now analyze more closely the exporter productivity premium and foreign-ownership productivity premium while controlling for other firm and country and sector fixed effects. We distinguish between manufacturing sectors and services sectors throughout the analysis and present the results in parallel. Based on our discussion above, we estimate equation (2).

$$\begin{aligned}(2)\qquad \theta_{ijk} &= \beta_0 + \beta_1 NE + \beta_2 FD + \beta_3 FE + \beta_5 (size_{medium}) + \\ &+ \beta_6 (size_{large}) + \beta_7 (\ln(GDP_{cap})) + \alpha(controls) + \varepsilon_{ijk}\end{aligned}$$

The four firm types are identified by dummy variables, with the National Domestic firms as default. The coefficients estimated for the National Exporters and the Foreign Domestic firms therefore reflect the export productivity premium among national firms and the foreign-ownership productivity premium among domestic firms, respectively. To analyze the export premium for foreign firms we compare the estimated coefficients for Foreign Exporters and Foreign Domestic firms. Similarly, to analyze the foreign ownership productivity premium for exporting firms we compare the estimated coefficients for Foreign Exporters and National-Exporters. We also analyze size effects (dummies for medium and large firms) and development levels (ln(GDP_{cap})), while controlling for fixed effects per country, sector, and sector-country interaction.[15]

Table 2.9 Productivity, exports, and foreign-ownership

Normalized productivity	*Manufactures*			*Services*		
	1	2	3	4	5	6
National Exporter	0.046	0.039	0.043	0.020	0.013	0.013
	(8.53)**	(7.43)**	(8.03)**	(1.79)	(1.13)	(1.17)
Foreign Domestic	0.081	0.070	0.070	0.103	0.087	0.087
	(9.80)**	(8.79)**	(8.80)**	(11.70)**	(9.96)**	(9.91)**
Foreign Exporter	0.092	0.076	0.075	0.109	0.078	0.082
	(10.12)**	(8.53)**	(8.39)**	(5.63)**	(4.15)**	(4.32)**
Size Medium	0.042	0.038	0.036	0.024	0.018	0.019
	(10.19)**	(9.60)**	(8.97)**	(3.99)**	(3.03)**	(3.20)**
Size Large	0.069	0.065	0.063	0.009	0.008	0.009
	(12.83)**	(12.36)**	(11.91)**	(1.23)	(1.05)	(1.22)
Ln(GDP/cap)	0.115	0.126	0.080	0.067	0.049	0.032
	(33.15)**	(21.66)**	(6.97)**	(11.69)**	(4.61)**	(1.12)
Constant	-0.670	-0.715	-0.174	-0.206	0.103	0.187
	(21.56)**	(13.74)**	(1.35)	(4.00)**	(1.05)	(0.71)
Observations	6146	6146	6146	3075	3075	3075
R-squared	0.23	0.29	0.32	0.09	0.17	0.20
Test if coefficients are significantly different, F-test (Prob > F)						
NE v FD	0.000**	0.001**	0.002**	0.000**	0.000**	0.000**
NE v FE	0.000**	0.000**	0.001**	0.000**	0.002**	0.001**
FD v FE	0.330	0.635	0.705	0.761	0.660	0.801

Dependent variable: normalized productivity; robust t statistics in parentheses; * significant at 5%; ** significant at 1%; the specification in columns 2, 3, 5 and 6 include sector and country fixed effects; the 3rd and 6th specification also include sector-country interaction fixed effects; NE = National Exporter; FD = Foreign Domestic; FE = Foreign Exporter.

Table 2.9 lists the basic estimates of the relationships between productivity, export premium, and foreign ownership premium. All estimates control for firm size, where small firms (less than 20 employees) are the default, medium size firms have between 20 and 99 employees, and large firms have 100 or more employees. Similarly, all estimates control for development level using the log of GDP per capita in PPP. We also included, but do not report, a dummy variable for firms that are part of a conglomerate and for a dummy variable for firms located in the capital city.[16] Columns 1 and 4 of Table 2.9 do not control for country and sector fixed

[15] As usual these controls are not reported in the tables; a fixed effect per country, for example Peru, is generated by including a dummy variable equal to 1 for Peruvian firms and 0 otherwise for all but one of the 15 countries. Similarly for sector and country-sector interaction fixed effects.
[16] The definition of the variables and their correlation is provided in Appendix A2.5.1 and A2.5.2

effects, while columns 2, 3, 5, and 6 do have country and sector fixed effects. Columns 3 and 6, in addition, include sector-country interaction fixed effects.

We can now provide a better and more detailed answer to suggestions II en IV. Starting with the *manufacturing* sectors, we can conclude that National Domestic firms are less productive than National Exporters, which in turn are less productive than Foreign Domestic firms. Although the point estimate of productivity for Foreign Exporters is higher than for Foreign Domestic firms the difference is not statistically significant, so we cannot conclude that Foreign Exporters are more productive than Foreign Domestic firms. This suggests that the foreign-ownership productivity premium is more important than the export productivity premium. Note that the same conclusions hold for all three columns (1, 2, and 3). For the *services* firms we arrive at almost the same conclusions. This time, however, there is not even an export productivity premium for national firms. Note again that the same conclusion holds for all three columns (4, 5, and 6). We can summarize our results as follows.

Stylized fact IV: There is a foreign-ownership productivity premium, both for domestic firms and for exporting firms and both for manufacturing sectors and services sectors.

Stylized fact V: There is only an export productivity premium for national manufacturing sectors. There is no export productivity premium for foreign manufacturing sectors, nor for services sectors (either national or foreign).

Briefly looking at the size of the coefficients in Table 2.9 shows that they have the expected effect in the manufacturing sectors since medium size firms are more productive than small firms and large firms are even more productive. The results are mixed for the services sector, however, where medium size firms are more productive than small firms but the effect of large firm size is insignificant. Regarding the development productivity premium the results are clear and strong for the manufacturing sectors. The effect is about half as strong for the services sectors (and is even insignificant if we include sector-country interaction effects).

Stylized fact VI: There is a clear size productivity premium and development productivity premium in the manufacturing sectors. These premia are smaller and less clear for the services sectors.

2.5 Robustness and foreign ownership intensity

The analysis in section 2.4 raises some questions which we further analyze in the next three subsections. First, are our conclusions on firm types robust if we measure productivity relative to the *permanent* labor force only (thus excluding temporary workers)? Second, are our conclusions robust if we use a *different sector* classification or *value added* per worker instead of sales per worker as the basis of our productivity measure, while controlling for capital intensity? Note that these alternatives for the second question are only available for the manufacturing sectors. Third, if the foreign-ownership productivity premium is more important than the export productivity premium, does the *intensity* of foreign-ownership matter?

2.5.1 Robustness I: permanent labor force

Our analysis above is based on measuring productivity relative to the total labor force, including both permanent and temporary workers. We think this is the most reasonable option since temporary workers also contribute to the firm's production level. By their very nature, however, the size of the temporary workforce may vary substantially in a given period. As a first robustness check, therefore, we measure a firm's productivity using only the permanent labor force. Repeating all steps in the paper above gives similar results as before, culminating in the regression results reported in Table 2.10. Note that, as expected by focusing only on a part of the labor force, the indicators for firm type in Table 2.10 are stronger than in Table 2.9 (larger coefficients). More importantly, their significance and ranking order is unchanged. All other results, on firm size and development level, are also similar as before. Our results are thus robust with respect to this alternative specification.

2.5.2 Robustness II: different sectors, value added, and capital intensity

We repeat the estimation analysis of section 2.4 for manufacturing sectors for two alternative specifications. First, we identify sectors more precisely by the main output that generates most sales according to the ISIC classification to calculate ISIC – normalized productivity measures (with ISIC – sectors fixed effects). Second, we calculate value added per worker (= (sales–intermediate input costs) / workers) as the basis for our productivity measure. This is then normalized per sector as before, see equation (2). In addition, we can now control for capital intensity using fixed costs per worker as a proxy, both for the original specification and for the two new specifications.

$$(3) \qquad \hat{\theta}_{ijk} = \frac{\ln(\tilde{\theta}_{ijk}) - \min\limits_{i,k} \ln(\tilde{\theta}_{ijk})}{\max\limits_{i,k} \ln(\tilde{\theta}_{ijk}) - \min\limits_{i,k} \ln(\tilde{\theta}_{ijk})}, \quad where \quad \tilde{\theta}_{ijk} = \frac{sales_{ijk} - int\ input\ costs_{ijk}}{labor_{ijk}}$$

Table 2.10 Productivity, exports, and foreign ownership with permanent labor

Normalized productivity	*Manufactures*			*Services*		
	1	2	3	4	5	6
National Exporter	0.053	0.045	0.048	0.021	0.018	0.018
	(9.61)**	(8.27)**	(8.74)**	(1.78)	(1.56)	(1.60)
Foreign Domestic	0.083	0.075	0.075	0.102	0.087	0.087
	(9.80)**	(9.04)**	(8.98)**	(11.18)**	(9.71)**	(9.63)**
Foreign Exporter	0.098	0.083	0.083	0.114	0.089	0.094
	(10.52)**	(9.12)**	(8.97)**	(5.63)**	(4.61)**	(4.80)**
Size Medium	0.049	0.045	0.043	0.031	0.024	0.025
	(11.54)**	(10.91)**	(10.32)**	(4.85)**	(3.98)**	(4.07)**
Size Large	0.079	0.075	0.073	0.025	0.022	0.024
	(14.29)**	(13.78)**	(13.40)**	(3.24)**	(2.85)**	(3.05)**
Ln(GDP/cap)	0.106	0.120	0.081	0.051	0.045	-0.074
	(29.26)**	(19.71)**	(6.72)**	(8.45)**	(4.05)**	(1.45)
Constant	-0.617	-0.694	-0.177	-0.086	0.187	1.406
	(19.11)**	(12.70)**	(1.32)	(1.60)	(1.83)	(3.02)**
Observations	6103	6103	6103	3052	3052	3052
R-squared	0.22	0.28	0.30	0.08	0.18	0.20
Test if coefficients are significantly different, F-test (Prob > F)						
NE v FD	0.001**	0.001**	0.005**	0.000**	0.000**	0.000**
NE v FE	0.000**	0.000**	0.000**	0.000**	0.001**	0.001**
FD v FE	0.209	0.465	0.498	0.583	0.904	0.726

Dependent variable: normalized productivity based on permanent labor input only; robust t statistics in parentheses; * significant at 5%; ** significant at 1%; the specification in columns 2, 3, 5 and 6 include sector and country fixed effects; the 3rd and 6th specification also include sector-country interaction fixed effects; NE = National Exporter; FD = Foreign Domestic; FE = Foreign Exporter

Table 2.11 summarizes the results for the various alternative specifications. All specifications control for capital city location and whether the firm is part of a conglomerate (not reported). Similarly, all columns include sector, country, and sector-country interaction fixed effects (not reported). Our general findings from section 2.4 are very robust: all estimated coefficients for firm type remain positive and significant. The same holds for firm size and development level. More capital intensive firms have, as expected, a higher productivity level. Including capital intensity, however, only has limited effect on the estimated coefficients for the other main variables, which all remain statistically significant. In addition, stylized facts

IV (on foreign-ownership productivity premium) and V (on export productivity premium) continue to hold.[17]

Table 2.11 Manufactures robustness checks: ISIC sectors, value added, and capital intensity

	1	2	3	4	5	6
Manufactures only	Normalized productivity	Normalized productivity	ISIC-Normalized productivity	ISIC-Normalized productivity	Normalized value added per worker	Normalized value added per worker
National Exporter	0.043	0.046	0.045	0.044	0.041	0.038
	(7.83)**	(7.85)**	(8.02)**	(7.57)**	(7.08)**	(6.26)**
Foreign Domestic	0.070	0.063	0.065	0.061	0.060	0.057
	(7.45)**	(5.74)**	(6.36)**	(5.58)**	(5.92)**	(5.32)**
Foreign Exporter	0.075	0.066	0.070	0.066	0.059	0.055
	(7.41)**	(5.80)**	(6.37)**	(5.76)**	(5.38)**	(5.32)**
Size Medium	0.036	0.044	0.035	0.047	0.023	0.034
	(8.99)**	(10.24)**	(8.68)**	(10.84)**	(5.61)**	(7.82)**
Size Large	0.063	0.081	0.063	0.085	0.043	0.062
	(11.47)**	(13.08)**	(11.09)**	(13.95)**	(7.49)**	(10.24)**
Ln(GDP/cap)	0.063	0.048	0.070	0.055	0.057	0.049
	(5.60)**	(3.99)**	(6.10)**	(4.60)**	(5.00)**	(3.86)**
Ln(Cap int)		0.250		0.331		0.287
		(14.12)**		(16.02)**		(14.11)**
Constant	-0.223	-0.131	-0.273	-0.195	-0.102	-0.083
	(2.23)*	(1.23)	(2.69)**	(1.84)	(1.01)	(0.73)
Observations	6146	4693	5702	4696	5171	4386
R-squared	0.32	0.35	0.33	0.37	0.27	0.30
Test if coefficients are significantly different, F-test (Prob > F)						
NE v FD	0.008**	0.161	0.075*	0.152	0.087*	0.115
NE v FE	0.002**	0.092*	0.031**	0.064*	0.112	0.122
FD v FE	0.745	0.820	0.733	0.732	0.935	0.905

Dependent variable listed above column; robust t statistics in parentheses; * significant at 5%; ** significant at 1%; all specifications control capital city, conglomerate, sector, country and country-sector interaction fixed effects; NE = National Exporter; FD = Foreign Domestic; FE = Foreign Exporter

2.5.3 Foreign ownership intensity

We have shown above that the export productivity premium is only important for national firms in the manufacturing sectors while the foreign-ownership productivity premium is important for both domestic and exporting firms in all sectors. The question then arises if the extent of foreign-ownership is important for the productivity premium. Since there is no export productivity premium for foreign-owned firms there is no need to identify them

[17] The only exception is the foreign-ownership productivity premium for exporting firms when value added is used as the basis for the productivity measure.

separately. Instead, we subdivide these firms into two roughly equally sized group based on foreign-ownership intensity, namely:

- Foreign-medium intensive firms (foreign-ownership in between 10 and 90 percent)
- Foreign-high intensive firms (foreign-ownership share above 90 percent)

Table 2.12 Productivity and foreign-ownership intensity

	Services	*Manufacturers*					
	Normalized productivity	Normalized productivity		ISIC-Normalized productivity		Normalized value added per worker	
	1	2	3	4	5	6	7
National Exporter	0.013	0.043	0.046	0.045	0.044	0.041	0.038
	(1.23)	(7.83)**	(7.85)**	(8.02)**	(7.56)**	(7.14)**	(6.31)**
Foreign-medium	0.049	0.057	0.058	0.052	0.058	0.041	0.043
	(3.73)**	(5.41)**	(4.96)**	(4.62)**	(4.95)**	(3.42)**	(3.73)**
Foreign-high	0.106	0.083	0.068	0.077	0.067	0.073	0.065
	(9.61)**	(9.18)**	(6.46)**	(7.76)**	(6.27)**	(7.93)**	(6.79)**
Size Medium	0.019	0.036	0.045	0.035	0.047	0.024	0.034
	(3.19)**	(9.04)**	(10.3)**	(8.74)**	(10.85)**	(5.66)**	(7.81)**
Size Large	0.010	0.064	0.081	0.063	0.086	0.043	0.062
	(1.33)	(11.6)**	(13.2)**	(11.20)**	(14.04)**	(7.60)**	(10.28)**
Ln(GDP/cap)	-0.094	0.063	0.048	0.070	0.055	0.057	0.049
	(2.78)**	(5.63)**	(4.00)**	(6.12)**	(4.61)**	(5.01)**	(3.87)**
Ln(Cap int.)			0.249		0.330		0.285
			(14.1)**		(15.98)**		(14.00)**
Constant	1.508	-0.223	-0.132	-0.273	-0.195	-0.104	-0.084
	(4.72)**	(2.24)*	(1.23)	(2.70)**	(1.85)	(1.03)	(0.74)
Observations	3077	6147	4693	5702	4696	5171	4386
R-squared	0.20	0.32	0.35	0.33	0.37	0.27	0.30
Joint significance F-test (Prob > F)							
NE v FM	0.028**	0.203	0.331	0.547	0.261	0.987	0.701
NE v FH	0.000**	0.000**	0.046**	0.003**	0.041**	0.002**	0.010**
FM v FH	0.000**	0.048**	0.495	0.086*	0.546	0.031**	0.116

Dependent variable listed above column; robust t statistics in parentheses; * significant at 5%; ** significant at 1%; all specifications control for capital city location and if the firm is part of a conglomerate; all specifications include sector, country and country-sector interaction fixed effects; NE = National Exporter; FM = Foreign-Medium; FH = Foreign-High

We therefore again have four types of firms. The National Domestic firms are again the default in the specifications, which is the same as in the previous section. We therefore immediately perform robustness checks on different sector classification, capital intensity, and value added for the manufacturing sectors. The results are summarized in Table 2.12, where all columns again control for capital city location and whether the firm is part of a conglomerate, while also including sector, country, and sector-country interaction fixed effects (not reported). We thus have one specification for the services sectors and six specifications for the manufacturing sectors.

The three reported firm types (National Exporter, Foreign-medium, and Foreign-high) all have statistically significantly higher productivity than the National Domestic firms. The only exception is for the National Exporters in the services sectors. This is exactly in line with Stylized fact V, which therefore continues to hold. Similarly, the estimates on the firm size productivity premium and development productivity premium are clear for the manufacturing sectors and smaller and less clear for the services sectors, which is in line with Stylized fact VI (which therefore also continues to hold).

Our main interest lies, of course, in the relative rankings of the foreign-ownership intensity productivity premium. Note, first of all, that the estimated coefficients for all seven columns listed in Table 2.12 lead to the same productivity ranking among the four types of firms (from low to high): National Domestic < National Exporters < Foreign-medium < Foreign-high. These differences are all statistically significant for the services sectors. For the manufacturing sectors the difference between National Exporters and Foreign-medium is never significant while the difference between National Exporters and Foreign-high is always significant. Finally, the difference between Foreign-medium and Foreign-high is statistically significant in all cases if we do *not* control for capital intensity, while it is *not* significant if we *do* control for capital intensity. Evidently, the higher productivity of high-intensity foreign-owned firms is largely based on higher capital intensity and not on some other inherent advantage over medium-intensity firms.

Stylized fact VII: The productivity premium of high-intensity foreign-owned firms over medium-intensity foreign-owned firms appears to be based solely on higher capital intensity

2.6 Conclusions

We analyze normalized productivity differences for various firm types for 15 developing Latin American countries. We summarize our findings in seven stylized facts, see Table 2.13. We first identify four types of firms: National Domestic, National Exporter, Foreign Domestic, and Foreign Exporter. This allows us to separately investigate the export productivity premium and the foreign-ownership productivity premium. We start by noting that, in contrast to the theoretical literature, there is considerable productivity overlap across firm types. We are therefore unable to identify either a viability threshold, an export threshold, or a multinational threshold. Future theoretical developments should try to incorporate the absence of such thresholds one way or another.

Table 2.13 Summary of stylized facts

I.	There is no clear productivity threshold for firm viability.
II.	There are no clear productivity thresholds sorting firms in different types.
III.	Productivity distributions generally differ per country, sector, firm type, and firm size.
IV.	There is a foreign-ownership productivity premium, both for domestic firms and for exporting firms and both for manufacturing sectors and services sectors.
V.	There is only an export productivity premium for national manufacturing sectors. There is no export productivity premium for foreign manufacturing sectors, nor for services sectors (either national or foreign).
VI.	There is a clear size productivity premium and development productivity premium in the manufacturing sectors. These premia are smaller and less clear for the services sectors.
VII.	The productivity premium of high-intensity foreign-owned firms over medium-intensity foreign-owned firms appears to be based solely on higher capital intensity.

We continue by noting that even after our normalization productivity distributions still differ in general across countries, sectors, firm types, and firm sizes. We incorporate these differences in our subsequent analysis. We find a clear size productivity premium and development productivity premium in the manufacturing sectors. These premium are smaller and less clear for the services sectors. We also find a clear foreign-ownership productivity premium, both for domestic firms and for exporting firms and both for manufacturing sectors and services sectors. In contrast, we only find an export productivity premium for national firms in the manufacturing sectors. The foreign-ownership productivity premium is thus more important than the export productivity premium. All these effects are robust when we use a different sector classification, when we base our productivity measure on value added per worker rather than sales per worker, when we base our productivity measure on the permanent labor force only, and when we control for capital intensity. A final question we address is whether the intensity of foreign-ownership is important for productivity by identifying medium-intensity and high-intensity firms. The results are in line with all our previous findings. In addition, it appears that the higher productivity of high-intensity firms over medium-intensity firms is based on higher capital intensity only.

Acknowledgement

We thank Hein Roelfsema for drawing our attention to the World Bank Enterprise survey data. We are grateful to two anonymous referees, Steven Brakman and Martijn Boermans for useful comments and suggestions. We thank participants at all conferences and workshops for useful comments and suggestions. The usual disclaimer applies.

2.7 References

Alvarez, R. and R.A. López (2005), "Exporting and performance: evidence from Chilean plants," Canadian Journal of Economics 38(4): 1385-1400.

Aw, B.Y. and A. R. Hwang (1995), "Productivity and the export market: a firm-level analysis*," Journal of Development Economics* 47: 313-332.

Aw, B.Y. and G. Batra (1998), "Firm size and the pattern of diversification," International Journal of Industrial Organization 16: 313-331.

Aw, B.Y. and G. Batra (1999), "Wages, firm size, and wage inequality: how much do exports matter?" In D.B. Audretsch and R. Thurik, eds., Innovation, Industry Evolution, and Employment. Cambridge, U.K.: Cambridge University Press.

Aw, B. Y., S. Chung, and M.J. Roberts (2000), "Productivity and turnover in the export market: micro-level evidence from the Republic of Korea and Taiwan (China)," The World Bank Economic Review 14(1): 65-90.

Aw, B.Y., X. Chen and M.J. Roberts (2001), "Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing," *Journal of Development Economics* 66(1): 51-86.

Aw, B.Y. and Y. Lee (2008), "Firm heterogeneity and location choice of Taiwanese multinational," *Journal of International Economics* 76(2): 402.

Bartelsman, E.J. and M. Doms (2000), "Understanding productivity: lessons from longitudinal data," Journal of Economic Literature XXXVIII (3): 569-594.

Bernard, A.B., J.B. Jensen, and R.Z. Lawrence (1995), "Exporters, jobs, and wages in U.S. manufacturing: 1976-1987," Brookings Papers on Economic Activity, Microeconomics: 67-119.

Bernard, A. B. and J. Wagner (1997), "Exports and success in German manufacturing," Review of World Economics 133: 134-157.

Bernard, A.B. and J.B. Jensen (1999), "Exceptional exporter performance: cause, effect, or both?" Journal of International Economics 47(1):1-25.

Bernard, A.B., J. Eaton, J.B. Jensen, and S. Kortum (2003), "Plants and productivity in international trade," The American Economic Review 93(4): 1268-1290.

Bernard, A.B. and J.B. Jensen (2004), "Exporting and productivity in the USA," Oxford Review of Economic Policy 20(3): 343-357.

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

Van Biesebroeck, J. (2005), "Exporting raises productivity in sub-Saharan African manufacturing firms," *Journal of international economics* 67(2): 373-391.

Blyde, J. and G. Iberti (2010), Trade cost, resource reallocation and productivity in developing countries," MPRA Paper 21317, University Library of Munich, Germany

Buchinsky, M. (1994), “Changes in the U.S. wage structure 1963-1987: application of quantile regression,” Econometrica, 62(2): 405-458.

Chen, T.J. and D.P. Tang (1987), “Comparing technical efficiency between Import-substitution-oriented and export oriented foreign firms in a developing. economy,” Journal of Development Economics 26(2): 277-89.

Clerides, S. K., S. Lach, and J.R. Tybout (1998), “Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco,” The Quarterly Journal of Economics 113(3): 903-947.

Das, S., M.J. Roberts, and J.R. Tybout (2007), “Market entry costs, producer heterogeneity, and export dynamics,” Econometrica 75(3): 837-873.

Demidova, S., and A. Rodríguez-Clare (2009), “Trade policy under firm-level heterogeneity in a small economy,” Journal of International Economics, 78(1): 100-112.

Fernandes, A.M. and A.E. Isgut (2005), “Learning-by-doing, learning-by exporting, and productivity: evidence from Colombia,” World Bank Working Paper 3544.

Girma, S., H. Görg and E. Strobl (2004b), “Exports, international investment, and plant performance: evidence from a non-parametric test,” Economics Letters 83: 317-324.

Greenaway, D., J. Gullstrand and R. Kneller (2005), “Exporting may not always boost firm productivity”, Review of World Economics, 141(4):561-582.

Greenaway, D. and R.Kneller (2008), “Exporting, productivity and agglomeration,” *European economic review* 52(5): 919-939.

Haddad, M. and A. Harrison (1993), “Are there positive spillovers from direct foreign investment? Evidence from panel data for Morocco,” Journal of Development Economics 42: 51-74.

Hao, L. and D.Q. Naiman (2007), “Quantile regression,” Quantitative Applications in the Social Sciences, no.149 Thousand Oaks CA: Sage.

Handoussa, J., M. Nishimizu and J. Page (1986), “Productivity change in Egyptian public sector industries after the 'opening',” Journal of Development Economics 20(l): 53-74.

Helpman, E., M. J. Melitz and S.R. Yeaple (2004), “Export versus FDI with heterogeneous firms,” The American Economic Review 94(1): 300-316.

Helpman, E. (2006), “Trade, FDI and the organization of firms,” Journal of Economics Literature 44: 589-630.

IMF (2009), Balance of Payments and International Investment Position Manual, sixth edition, IMF, Washington DC.

Jean S. (2002), “International trade and firms' heterogeneity under monopolistic competition,” *Open Economies Review* 13(3): 291-311.

Koenker, R. (2005), Quantile Regression, New York: Cambridge University Press.

Koenker, R. and K.F. Hallock (2001), “Quantile regression: an introduction,” Journal of Economic Perspectives 15(4): 143-156.

Kolmogorov, A.N. (1933), “Sulla determinizazione empirica delle leggidi probabilita”, *Giornale dell ’Istituto Italiano Attuari* 4: 1-11.

Leonidou, L.C. (1995), “Empirical research on export barriers -review assessment, and synthesis,” Journal of International Marketing 3(1): 29~43.

Leonidou, L (2000), “Barriers to export management: An organizational and internationalization analysis,” Journal of International Management 6: 121–14.

Mayer, T. and G. Ottaviano (2007), “The happy few: the internationalization of European firms: new facts based on firm level evidence,” Brueghel Blueprint Series Vol III.

Mayer, T. and G. Ottaviano (2008), “The happy few: the internationalization of European firms,” Intereconomics 43(3): 135-148.

Medin, H. (2003), “Firms’ export decisions – fixed trade costs and the size of the export Market,” Journal of International Economics 61: 225-241.

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

Pavcnik, N. (2002), “Trade liberalization, exit, and productivity improvements: evidence from Chilean plants,” Review of Economic Studies 69(1): 245-276.

Roberts, M.J. and J.R. Tybout (1997), “The decision to export in Colombia: an empirical model of entry with sunk costs,” American Economic Review, 87: 545-564.

Schmitt, N. and Z. Yu (2001), “Economies of scale and the volume of intra-industry trade,” *Economics letters* 74(1): 127-132.

Sinha, R. (1993), “Foreign participation and technical efficiency in Indian industry,” Applied Economics 25: 583-588.

Sjöholm, F. (2003), “Which Indonesian firms export? The importance of foreign networks,” Regional Science, 82(3): 333–50.

Smirnov, N.V., (1939), “Estimate of deviation between empirical distribution functions in two independent samples,” *Bulletin Moscow University* 2(2): 3-16.

Tomiura, E. (2007), “Foreign outsourcing, exporting, and FDI: a productivity comparison at the firm level,” *Journal of International Economics* 72(1): 113-127.

Tybout, J.R. and M.D. Westbrook (1995), “Trade liberalization and dimensions of efficiency change in Mexican manufacturing industries,” Journal of International Economics 31: 53-78.

Tybout, J. (2000), “Manufacturing firms in developing countries: How well do they do, and why?” *Journal of Economic Literature* 38(1): 11-44.

Wagner, J. and A.B. Bernard (1997), “Exports and success in German manufacturing,” *Review of World Economics*: *Weltwirtschaftliches Archiv* 133(1): 34-157.

Wagner, J. (2006), “Export intensity and plant characteristics: what can we learn from quantile regression?” *Review of World Economics: Weltwirtschaftliches Archiv* 142(1): 195-203.

Wagner, J. and A. Vogel (2010), “Higher productivity in importing German manufacturing firms: self-selection, learning from importing, or both?” *Review of World Economics: Wweltwirtschaftliches Archiv* 145(4): 641-665.

Yeaple, S.R. (2005), “A simple model of firm heterogeneity, international trade, and wages,” *Journal of International Economics* 65(1): 1-20.

Yeaple, S.R. (2009), “Firm heterogeneity and the structure of U.S. multinational activity,” *Journal of International Economics* 78(2): 206-215.

2.8 Appendix

A2.1 Number of observations across countries and sectors

	ARG	BOL	CHL	COL	ECU	SAL	GTM	HND	MEX	NIC	PAN	PAR	PER	URY	VEN	Total	%
Food	167	123	160	154	105	131	90	83	158	83	69	93	120	119	72	1,727	15.8
Garments	119	121	72	172	27	114	38	15	162	20	19	56	120	74	37	1,166	10.67
Textiles	117	0	49	147	44	25	45	24	155	8	3	7	35	44	22	725	6.63
Machinery & Equipment	127	0	33	0	6	2	3	5	236	8	8	10	0	1	12	451	4.13
Chemicals	67	59	74	160	97	29	15	22	169	24	12	108	83	122	15	1,056	9.66
Electronics[18]	1	0	0	0	1	1	0	0	77	0	0	4	0	0	5	89	0.81
Non-Metallic mineral	3	20	4	1	10	29	9	23	165	23	10	37	0	5	9	348	3.18
Other Manufacturing	145	86	305	15	104	136	128	91	39	199	122	125	3	31	111	1,640	15
Retail	123	123	123	121	138	54	67	66	119	42	119	127	123	125	91	1561	14.28
Information technology	106	2	119	120	0	7	8	1	118	4	0	3	0	4	2	494	4.52
Other Services	64	33	43	28	106	132	73	52	52	48	54	13	128	51	87	964	8.82
Construction	24	46	35	82	20	33	46	54	30	19	117	30	20	45	37	638	5.84
Wholesale	0	0	0	0	0	0	0	0	0	0	71	0	0	0	0	71	0.65
Total	1,063	613	1,017	1,000	658	693	522	436	1480	478	604	613	632	621	500	10,930	100
%	9.73	5.61	9.3	9.15	6.02	6.34	4.78	3.99	13.54	4.37	5.53	5.61	5.78	5.68	4.57	100	

[18] Figures and results for the electronic and the wholesale sectors are not provided as there are not enough observations available.

A2.2.1 Summary statistics of the log of productivity and normalized productivity by sector

Sectors	mean	variance	min	median	max
log of productivity					
Food	9.83	1.612	4.36	9.76	17.32
Garments	9.26	1.544	4.24	9.33	17.94
Textiles	9.76	1.310	3.83	9.80	16.33
Machinery & Equipment	10.24	1.197	4.24	10.26	13.75
Chemicals	10.23	1.484	2.72	10.24	14.65
Electronics	10.13	1.304	3.60	10.04	12.12
Non-Metallic mineral	9.32	1.664	4.34	9.28	13.26
Other Manufacturing	9.72	1.947	5.21	9.73	18.96
Retail	10.38	1.838	4.24	10.47	19.46
Information technology	10.19	1.169	5.70	10.11	17.63
Other Services	10.33	2.329	1.40	10.32	17.52
Construction	9.92	2.071	1.08	9.96	15.11
Wholesale	10.70	0.952	8.18	10.82	12.60
average	9.94	1.829	1.38	9.93	19.46
Normalized productivity					
Food	0.38	0.021	0.00	0.369	1.00
Garments	0.37	0.022	0.00	0.368	1.00
Textiles	0.47	0.027	0.00	0.470	1.00
Machinery & Equipment	0.48	0.025	0.00	0.477	1.00
Chemicals	0.44	0.020	0.00	0.435	1.00
Electronics	0.57	0.039	0.00	0.531	1.00
Non-Metallic mineral	0.40	0.034	0.00	0.386	1.00
Other Manufacturing	0.38	0.025	0.00	0.381	1.00
Retail	0.40	0.020	0.00	0.415	1.00
Information technology	0.48	0.021	0.00	0.464	1.00
Other Services	0.41	0.028	0.00	0.410	1.00
Construction	0.38	0.024	0.00	0.381	1.00
Wholesale	0.57	0.049	0.00	0.597	1.00
average	0.41	0.025	0.00	0.406	1.00

A2.2.2 Summary statistics of productivity by ISIC sector classification

ISIC-normalized productivity	mean	variance	min	median	max
Food	0.38	0.020	0.00	0.369	1.00
Garments	0.37	0.022	0.00	0.371	1.00
Textiles	0.46	0.027	0.00	0.458	1.00
Machinery & Equipment	0.47	0.026	0.00	0.472	1.00
Chemicals	0.44	0.021	0.00	0.443	1.00
Electronics	0.58	0.036	0.00	0.555	1.00
Non-Metallic mineral	0.37	0.036	0.00	0.355	1.00
Other Manufacturing	0.36	0.023	0.00	0.364	1.00
average	0.40	0.025	0.00	0.398	1.00

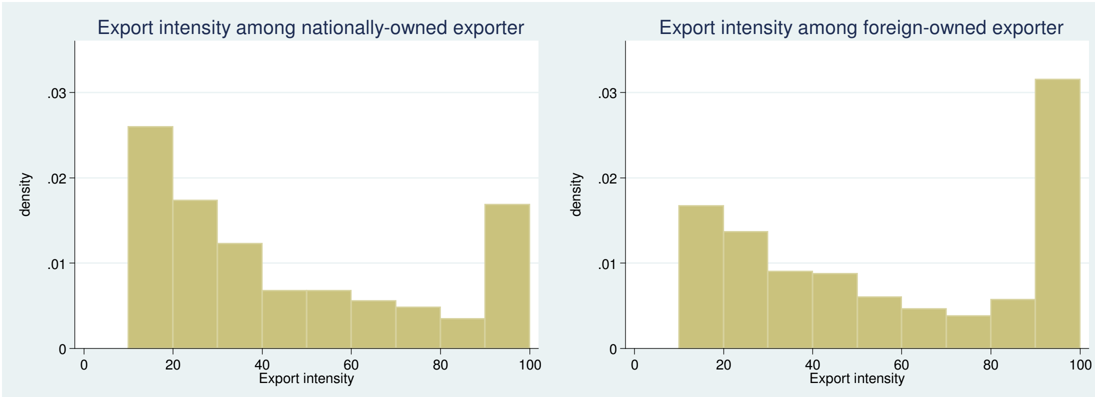
A2.2.3 Normalized productivity by countries in the food sector

Development	Country	country code	mean	variance	min	medium	max
Middle-High	Mexico	MEX	0.37	0.017	0.02	0.35	0.74
	Chile	CHL	0.45	0.021	0.12	0.42	0.97
	Argentina	ARG	0.44	0.018	0.12	0.43	0.85
	Venezuela	VEN	0.38	0.021	0.04	0.37	0.83
	Uruguay	URY	0.41	0.022	0.12	0.39	0.78
Middle-Middle	Panama	PAN	0.44	0.018	0.17	0.41	1.00
	Colombia	COL	0.35	0.013	0.08	0.35	0.89
	Ecuador	ECU	0.42	0.014	0.15	0.42	0.73
	Peru	PER	0.42	0.027	0.00	0.41	0.80
	El Salvador	SAL	0.37	0.016	0.10	0.36	0.86
Middle-Low	Guatemala	GTM	0.33	0.017	0.05	0.34	0.66
	Paraguay	PAR	0.31	0.021	0.03	0.31	0.82
	Bolivia	BOL	0.30	0.019	0.00	0.29	0.81
	Honduras	HND	0.33	0.018	0.05	0.32	0.72
	Nicaragua	NIC	0.31	0.019	0.03	0.28	0.68

A2.3.1 Percentage of firm types within each sector (updated with full data)

Sectors	National Domestic	National Exporter	Foreign Domestic	Foreign Exporter	# firms
Manufactures					
Food	76.63	13.01	6.11	4.24	1,652
Garments	77.11	17.57	2.66	2.66	1,127
Textiles	76.71	16.43	3.57	3.29	700
Machinery & Equipment	70.84	17.54	3.87	7.74	439
Chemicals	73.48	12.63	8.49	5.40	1,037
Electronics	67.86	13.10	4.76	14.29	84
Non-Metallic mineral	82.60	10.03	4.42	2.95	339
Other Manufacturing	77.41	13.66	4.92	4.01	1,523
Services					
Retail	87.93	2.86	8.79	0.41	1,467
Information technology	76.73	10.41	8.78	4.08	490
Other Services	74.54	6.96	15.53	2.97	876
Construction	78.76	8.19	10.70	2.34	598
Wholesale	78.87	7.04	9.86	4.23	71

A2.3.2 Export intensity of National Exporters and Foreign Exporters



A2.3.3. Export intensity by country

Country	Export intensity (%) 10~20	21~60	61~100	Total # of exporters	Total # of firms	% of firms exporting
Mexico	38	58	37	133	1,480	8.99
Chile	58	45	32	135	1,015	13.30
Argentina	141	115	43	299	1,058	*28.26*
Venezuela	9	6	0	15	500	3.00
Uruguay	22	35	44	101	617	16.37
Panama	14	19	47	80	603	13.27
Colombia	40	45	18	103	1,000	10.30
Ecuador	34	22	23	79	656	12.04
Peru	37	36	48	121	632	19.15
El Salvador	49	60	54	163	693	*23.52*
Guatemala	40	39	28	107	522	*20.50*
Paraguay	16	25	32	73	611	11.95
Bolivia	25	21	28	74	612	12.09
Honduras	15	14	23	52	436	11.93
Nicaragua	16	11	16	43	478	9.00
All countries	554	551	473	1578	10,913	14.46

A2.4 Kolmogorov-Smirnov (KS) test

A2.4.1 Country pair wise KS test, p-values

	ARG	BOL	CHL	COL	ECU	SAL	GTM	HND	MEX	NIC	PAN	PAR	PER	URY	VEN
ARG	-	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BOL		-	0.00	0.00	0.00	0.00	0.07	0.18	0.00	0.00	0.00	0.11	0.00	0.00	0.00
CHL			-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
COL				-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
ECU					-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.61	0.00
SAL						-	0.07	0.06	0.00	0.00	0.00	0.20	0.00	0.00	0.00
GTM							-	0.62	0.00	0.00	0.00	1.00	0.00	0.00	0.00
HND								-	0.00	0.00	0.00	0.86	0.00	0.00	0.00
MEX									-	0.00	0.37	0.00	0.00	0.00	0.01
NIC										-	0.00	0.00	0.00	0.00	0.00
PAN											-	0.00	0.00	0.02	0.02
PAR												-	0.00	0.00	0.00
PER													-	0.55	0.00
URY														-	0.00
VEN															-

A2.4.2 Sector pair wise KS test, p-values

	F	G	T	M	Ch	E	N	OM.	R	IT	OS	Co	W
F – Food	-	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.20	0.00
G – Garments		-	0.00	0.00	0.00	0.00	0.03	0.04	0.00	0.00	0.00	0.39	0.00
T – Textile			-	0.00	0.38	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00
M – Machinery				-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ch – Chemical					-	0.00	0.00	0.00	0.00	0.45	0.00	0.00	0.00
E – Electronics						-	0.00	0.00	0.00	0.00	0.00	0.00	0.68
N – Non-metallic							-	0.09	0.01	0.00	0.06	0.39	0.00
OM – Other Manuf.								-	0.00	0.00	0.00	0.83	0.00
R – Retail									-	0.00	0.00	0.00	0.00
IT										-	0.00	0.00	0.00
OS – Other Services											-	0.00	0.00
Co – Construction												-	0.00
W – Wholesale													-

A2.4.3 Firm type pair wise KS test, p-values

	National Domestic	National Exporter	Foreign Domestic	Foreign Exporter
National Domestic	-	0.00	0.00	0.00
National Exporter		-	0.00	0.00
Foreign Domestic			-	0.03
Foreign Exporter				-

A2.4.4 Size pair wise KS test, p-values

	Small	Medium	Large
Small	-	0.00	0.00
Medium		-	0.00
Large			-

A2.4.5 KS significant test, p-values

Norm. Productivity (ISIC)	ND v NE	FD v FE	ND v FD	NE v FE	NE v FD	ND v FE
Food	0.00	0.03	0.00	0.13	0.12	0.00
Garments	0.00	0.34	0.18	0.06	0.90	0.08
Textile	0.00	0.99	0.01	0.71	0.36	0.09
Machinery	0.00	0.12	0.00	0.00	0.02	0.00
Chemical	0.00	0.27	0.02	0.62	0.30	0.01
Electronics	0.48	0.74	0.13	0.06	0.09	0.01
Non-metallic material	0.46	0.82	0.00	0.00	0.00	0.01
Other manufacturing	0.00	0.01	0.18	0.02	0.01	0.00

A5. Regression related

A2.5.1 Definition of variables used in regressions

Normalized productivity	Log of sales per worker, normalized by sector
ISIC-normalized productivity*	Log of sales per worker, normalized by ISIC-sector
Normalized value added*	Log of value added (sales – total cost of raw material and intermediate inputs) per worker, normalized by ISIC-sector.
National Domestic	Dummy variable for nationally owned firms (less than 10% foreign ownership) with more than 90% domestic sales
National Exporter	Dummy variable for nationally owned firms (or less than 10% foreign ownership) and export at least 10% of their outputs
Foreign Domestic	Dummy variable for foreign-owned firms (with over 10% foreign ownership) and make over 90% of sales domestically
Foreign Exporter	Dummy variable for foreign-owned firms (with over 10% foreign ownership) and export at least 10% of their outputs.
Small	Dummy variable for firms employed less than 20 people
Medium	Dummy variable for firms employed between 20 to 99 people
Large	Dummy variable for firms employed more than 100 people
Ln(GDP/cap)	Log of GDP per capita, PPP
Conglomerate	Dummy for subsidiary firms (part of larger firm)
Capital city	Dummy for firms located in the capital of their country
Ln(Cap int)*	Log of capital intensity, calculated as fixed cost expenditure (annual expenditure on machinery, vehicles, equipment, land and building, and compensation on non-production workers) per worker, normalized by ISIC-sector.
* Variable only available only for firms in the manufacturing sectors	

A2.5.2 Correlation table

	Norm prod	ND	NE	FD	FE	Small	Medium	Large	log(GDPcap)	Congl
ND	-0.19*	1.00								
NE	0.09*	-0.59*	1.00							
FD	0.14*	-0.45*	-0.09*	1.00						
FE	0.13*	-0.32*	-0.07*	-0.05*	1.00					
Small	-0.17*	0.21*	-0.19*	-0.10*	-0.13*	1.00				
Medium	0.05*	-0.01	0.05*	0.02	-0.04*	-0.67*	1.00			
Large	0.16*	-0.26*	0.18*	0.10*	0.21*	-0.43*	-0.38*	1.00		
log(GDPcap)	0.31*	-0.05*	0.00	-0.05*	0.01	-0.01	-0.02	0.04*	1.00	
Conglomerate	0.14*	-0.16*	-0.01	0.16*	0.11*	-0.15*	-0.01	0.20*	0.09*	1.00
Capital city	0.04*	-0.07*	0.00	0.06*	0.00	0.01	0.00	-0.02	-0.02	0.02

ND = National Domestic; NE = National Exporter; FD = Foreign Domestic; FE = Foreign Exporter

Chapter 3 Impact of trade on viability and exporter selection with heterogeneous fixed cost in the Melitz model

3.1 Introduction

In the past decade, economists started to incorporate firm heterogeneity into the models of international trade and further analyze the impact of trade on these firms. The seminal work of Melitz (2003) is a particularly tractable model and has stimulated a great deal of research into analyzing the implication of firm heterogeneity for a wide range of issues in international trade. However, the productivity distribution depicted in the Melitz model contradicts the empirical findings. This paper reconciles the discrepancy between the empirical distribution found and the implied distribution from the theoretical work by introducing heterogeneous fixed cost into the Melitz model.

A growing body of literature followed the Meliz modeling structure in incorporating firm level heterogeneous marginal cost, i.e. productivity (Aw and Lee, 2008; Helpman et al. 2004; Bernard et al., 2003). These models show trade induces resource reallocation from the least productive firms to most productive ones. It revealed the unequal impact of trade on heterogeneous firms, while leaving the aggregated outcome comparable to a model with representative firms.

Despite its ability to capture many stylized facts that differentiate exporters from non-exporters (Greenaway and Kneller, 2007; Bernard et al., 2006; Bernard and Jensen, 1999; 2004), there are consequences in imposing productivity as the sole heterogeneous dimension among these firms. In particular, the model predicts a strong *pecking order*. The ordering suggests only firms with above a certain productivity threshold can carry out economic production, with another productivity threshold drawing between the exporters and non-exporters (also further between exporters and FDI firms, see Helpman et al. 2004). Based on productivity differences, the distinction between each type of firms is clear. However, the productivity distribution depicted based on the Melitz-type model is inconsistent with the empirical findings (Mayer and Ottaviano, 2007). Their empirical evidence suggests that there is no survival cut-off productivity threshold and there exist no productivity threshold to tell exporter apart from domestic firms.

This paper reconciles the discrepancy between the empirical distribution found and the implied distribution from the theoretical work by introducing heterogeneous fixed cost into the Melitz model. The model show how incorporating heterogeneous fixed cost and heterogeneous exporting fixed cost can successfully depicts the productivity distributions of active firms in the market that resembles empirical findings. In the equilibrium, the profitable firms that remain in operation have marginal productivity within a range. Especially, there are firms with very low marginal productivity still making economic production. This is because now the selection is not solely base on productivity but efficiency. Thus, firms that are less productive make economic production as long as their fixed cost is sufficiently low, while highly productive firms exit if their fixed cost is excessively high. Modeling with heterogeneous exporting fixed cost further made possible to reproduce the overlapping pattern of the productivity distributions for the exporters and non-exporters. The model while retained the idea analytical features of the Melitz model (resource reallocation to more efficient firms) further point out that different level of exporting fixed cost can affect the weighted productivity in the trade equilibrium. Irrespective of the mean to further liberalize trade, we find that trade is welfare improving.

Identifying and modeling with heterogeneous exporting fixed cost is not new in the literature (Robert and Tybout, 1997; Schmitt and Yu, 2001; Jørgensen and Schröder, 2006). However, it has not been the center of focus compare to other heterogeneous aspects, such as productivity. Nonetheless, there are reasons to take heterogeneous exporting fixed cost into consideration. First, treating market entry cost to be homogeneous is not realistic. Fixed cost related to exporting included cost on market research, establishing foreign contact and distribution networks, training human resource sent to foreign office and adapting to foreign preference and regulations, ect. Business managerial studies reveal the impact of export barriers on management decisions (Leonidou, 1995; 2000). They found that different types of barrier are associated with different cost. Since different firms take different types of barriers more difficult to tackle than another, the fixed cost incurred for market entry varies substantially among the exporting firms. Econometric evidence based on Italian manufacturing panel data find substantial differences among firms' abilities to collect and operationalize information about foreign markets and consumer tastes, which is the main entry barrier (Bugamelli and Infante, 2003). Das et al. (2007) also report that the sunk entry cost vary considerably across Colombian manufacturers.

Second, organizational literatures point out the change in industry structure with the activities in the value chain broken down and produced by different firms (Fine, 1998; Sturgeon, 2000; 2002). More recent case studies by Linden et al. (2009; 2011) broke down the global value chain network of Apple's products and reveal how production of a single product is fragmented. Along the supply chain, each firm provides different inputs, namely from product innovation, component production and assembly to sales and distribution (Hess and Coe, 2006; Dedrick et al., 2010). We see that despite being in the same narrowly defined industry, firms in different parts of the value chain requires different fixed cost invest for the relevant activities (Jørgensen and Schröder, 2008). Thus, the fragmentation of the production activities in the same industry provides additional motivation to model with heterogeneous fixed cost.

In the consecutive sections, we follow the Melitz structure in presenting our model and point out the additional feature in modeling with heterogeneous fixed cost. The third section emphasize the additional uncertainty arise from the additional heterogeneous firm dimension we incorporated in the closed economy. The fourth extended the model with the possibility to trade, in which the firms face additional uncertainty, the heterogeneous exporting fixed cost. The fifth section compares the trade equilibrium with autarky equilibrium. We analyze the effect of trade liberalization on exporter and viable firm selection, and the effect of trade liberalization on the development of the overall weighted productivity and welfare. The last section concludes.

3.2 Basic model setup

3.2.1 Consumption

Consumer preferences are given in equation 1 by a constant elasticity of substitution (CES) utility function U with a continuum of goods available for consumption indexed by $\omega \in \Omega$, where $q(\omega)$ denotes the quantity consumed of variety ω and Q denotes the aggregate quantities demanded. The parameter ρ represents consumers' love of variety effect. It ranges between zero and one ensuring that the product varieties are imperfect substitutes and not complements for each other. This in turn determined the lower bound of the elasticity of substitution, which is always greater than one $\varepsilon \equiv 1/(1-\rho) > 1$. Consumers maximize utility subject to the budget constraint, with $p(\omega)$ and $q(\omega)$ as the price charged and quantity consumed for each variety ω and R as the total expenditure (equal to aggregate firms' revenue).

(1) $$U = Q = \left[\int_{\omega \in \Omega} q(\omega)^{\rho} \, d\omega \right]^{1/\rho} ; \quad 0 < \rho < 1$$ (Utility function)

(2) $$R = \int_{\omega \in \Omega} p(\omega) q(\omega) \, d\omega$$ (Budget restriction)

The price index P is defined as in equation 3. *Technical Note 1* derives the individual demand functions as in equation 4; see also Dixit and Stiglitz (1977). The associated individual consumption constraint is also the revenue for the firm producing variety ω (equation 5). Notice that equations 4 and 5 hold for all $\omega \in \Omega$, as is henceforth implicit.

(3) $$P = \left[\int_{\omega \in \Omega} p(\omega)^{1-\varepsilon} \, d\omega \right]^{1/1-\varepsilon} , \varepsilon \equiv 1/(1-\rho)$$ (Price index)

(4) $$q(\omega) = R \, p(\omega)^{-\varepsilon} P^{\varepsilon - 1} = Q\big(p(\omega)/P\big)^{-\varepsilon}$$ (Demand function)

(5) $$r(\omega) \equiv p(\omega) q(\omega) = R\big(p(\omega)/P\big)^{1-\varepsilon}$$ (Firm revenue)

3.2.2 Production

There are many firms active in the market. Production involves a fixed and variable cost each period. Due to the fixed cost requirement for production, each firm is capable of and chooses to produce a single and unique product ω using labor as the only input in the equilibrium. Firms are heterogeneous in terms of marginal productivity. The marginal productivity of the firm producing good ω is denoted as $\varphi(\omega)$, which is always greater than zero. Production enjoys increasing returns to scale with firm-specific marginal cost $1/\varphi(\omega)$.

To produce $q(\omega)$ goods, firms have to employ $l(\omega)$ workers. With labor as the only input, firms also have to incur a fixed cost $f(\omega)$ as the basis for production activity:

(6) $q(\omega) = \varphi(\omega)[l(\omega) - f(\omega)]$ (Production function)

Rearranging the above equation gives us the labor demand function: $l(\omega) = f(\omega) + q(\omega)/\varphi(\omega)$, where $1/\varphi(\omega)$ is the marginal labor input requirement. The fixed labor input $f(\omega)$ ensures that as production expands, less labor is needed to produce an additional unit of output, which means that there are internal economies of scale (increasing return to production)[19]. The average labor requirement decreases as production increases. Input and output follow a linear yet stochastic relationship since the marginal productivity and fixed cost are determined stochastically.

It must be noted that not only variable cost varies with firm productivity $\varphi(\omega)$, which is different between firms, the per period fixed costs $f(\omega)$ differ between firms that produce differentiated goods ω as well, which are both determined stochastically. This stochastic setting of per period fixed cost is where the model presented in this paper deviates from the original Melitz (2003) model.

Each worker is paid the same wage rate *w*, so firm profits are as given in equation 7. Each firm maximizes profits by choosing its optimal price, subject to its demand function from equation 4, taking the economy-wide price level (P) and expenditure level (R) as given. The solution gives the optimal price rule (equation 8), where the last equality results from taking the wage rate as the numéraire (see *Technical Note 2* for detail). In view of the constant price elasticity of demand, each firm charges a constant mark-up $1/\rho$ over marginal costs. It should be noted that although the stochastic determination of per period fixed cost influenced the profit, it does not influence the optimal pricing rule derived from maximizing the firm's profit.

(7) $\pi(\omega) = r(\omega) - wl(\omega)$ (Firm profits)

(8) $p(\omega) = w/\rho\varphi(\omega) = 1/\rho\varphi(\omega)$ (Optimal price rule)

[19] Implicitly, the Dixit-Stiglitz framework assumes no economies of scope. Therefore, there is no reason for firms to produce multiple varieties. Each firm produces a distinct variety and each variety is only produced by one firm. In this case, firms do not lose profit to competition involved in producing the same type of varieties. Hence, the number of firms is also the number of varieties in equilibrium.

With the optimal pricing rule, firms' output and revenue are proportional to their productivity ($\varphi(\omega)$, $f(\omega)$ and $\varepsilon = 1/(1-\rho)$), and other macro-parameters, including: the total expenditure *R*, and the price index *P* (see equation 4′, 5′, 7′).

(4′) $$q(\omega) = Q\left(\rho\varphi(\omega)P\right)^{\varepsilon}$$ (Output function)

(5′) $$r(\omega) = R\left(P\rho\varphi(\omega)\right)^{\varepsilon-1}$$ (Firm revenue)

Given firm productivity $\varphi(\omega)$, the revenue earned is increasing in the aggregate expenditure and price level of the economy, which is also an inverse measure of the competition intensity in the market, and increasing in the inverse of mark-up prices, ρ . Note that two parameters: the firm's marginal productivity $\varphi(\omega)$ and the elasticity of substitution ε alone determine the price, output quantity and revenue ratio of any two firms that produce ω and ω' respectively (equation 9). In other words, the relative revenue of two firms solely depends on their relative productivity.

(9) $$\frac{p(\omega)}{p(\omega')} = \left(\frac{\varphi(\omega)}{\varphi(\omega')}\right)^{-1}; \qquad \frac{q(\omega)}{q(\omega')} = \left(\frac{\varphi(\omega)}{\varphi(\omega')}\right)^{\varepsilon}; \qquad \frac{r(\omega)}{r(\omega')} = \left(\frac{\varphi(\omega)}{\varphi(\omega')}\right)^{\varepsilon-1}$$

Intuitively, a more productive firm charges a lower price, sells more goods, and earns higher revenue. The latter two effects are magnified by the elasticity of substitution. At this stage, the elasticity of substitution ε plays a central role, as it determines the ratio of sales and revenues.

To determine a firm's viability in the market, the absolute value of the operating profits as given in equation 7′ is crucial. Entrepreneurs must earn non-negative profit to prove their ability to produce economically, and those who cannot earn a positive profit will immediately exit the market after knowing their productivity and equivalent fixed cost.

(7′) $$\pi(\omega) = (R/\varepsilon)\left(\rho P\varphi(\omega)\right)^{\varepsilon-1} - f(\omega)$$ (Firm profits[20])

---

20 $\pi(\omega) = r(\omega) - w[f(\omega) + q(\omega)/\varphi(\omega)] = r(\omega) - 1[f(\omega) + \rho p(\omega)q(\omega)] = (1-\rho)r(\omega) - f(\omega) = [r(\omega)/\varepsilon] - f(\omega)$

Unlike Melitz (2003), the endogenously determined cut-off productivity is not unique in our setting. But similar to Melitz' model, the model presented here provides a clear divide between profitable entrepreneurs and those who earn negative profit. Equation 7′ suggests that firms do not necessarily need to have high productivity as long as their stochastic fixed cost drawn is sufficiently low, such that it enables a firm to earn non-negative profit. Thus, it is no longer the case that the firms remaining in the market must be those with the highest productivity. Instead, they are those firms that earn a non-negative profit, which we call viable firms. So, even for firms with very high productivity, their viability depends on their fixed cost level. In short, the stochastic determined fixed cost plays a crucial role.

3.3 Closed economy viability selection

For given productivity draw, there exists a cut-off fixed cost level such that any firm with fixed cost below this level can make non-negative profit, allowing economic production. At the exact point where profit equals zero $\pi(\varphi, f) = 0$, we derived the following relation between productivity and fixed cost by rearranging equation 7'.

(10) $$f^*(\varphi) = (R/\varepsilon)\left(\rho P\varphi\right)^{\varepsilon-1} \Leftrightarrow \pi(\varphi, f^*(\varphi)) = 0$$

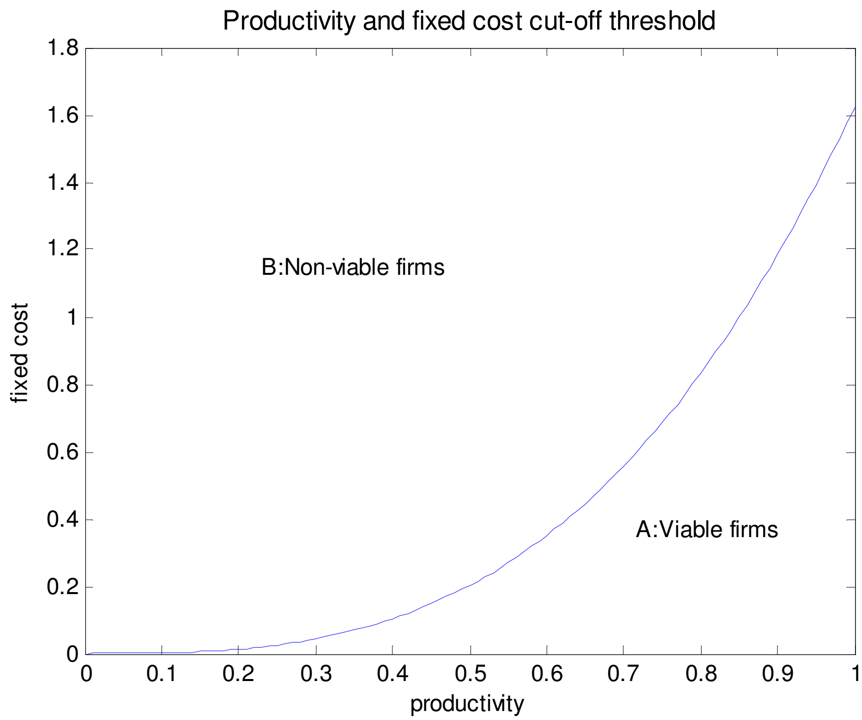
We define this specific productivity and fixed cost combination as the cut-off profit level. This cut-off profit combination can be expressed in terms of productivity and pinned down a specific fixed cost level. A curve can be drawn on the fixed cost and marginal productivity axes for equation 10. The curve shows all combinations of the productivity and fixed cost $(\varphi, f^*(\varphi))$ for which the profit is zero $\pi(\varphi, f^*(\varphi)) = 0$, and is later referred to as the *cut-off profit curve*. As shown in Figure 3.1, a clear division between the viable and non-viable firms is distinguished. For non-negative profit $\pi(\varphi, f(\varphi)) \geq 0$, productivity and fixed cost have the following relationship:

(10′) $$\pi(\varphi) \geq 0 \quad iff \quad f(\varphi) \leq f^*(\varphi)$$

Simply stated, for a given productivity draw $\varphi(\omega)$, firm ω must have a fixed cost level $f(\omega)$ below the cut-off fixed cost level $f^*(\varphi)$ to be viable and earn non-negative profit. These are the firms with productivity $\varphi = \varphi'$ and fixed cost below the cut-off fixed cost level $f^*(\varphi')$. Entrant firms producing ω'' with specific productivity and fixed cost combination (φ'', f'')

that generates a negative profit $\pi(\varphi'', f'') < 0$ will be forced to exit the market immediately since economic production is not possible.

Figure 3.1 The cut-off profit curve



We collect any pairs of productivity and fixed cost combinations that have non-negative profit in a set named: **A**. All productivity and fixed cost combinations in this set allows firms to make non-negative profit: $\forall(\varphi, f) \in \mathrm{A} \Leftrightarrow \pi(\varphi, f) \geq 0$. Along with the cut-off profit condition as given in equations 10 and 10′, set **A** can be formally defined by equation 11 as follows:

(11) $\mathrm{A} \equiv \left\{ (\varphi, f) \,\middle|\, f \leq f^*(\varphi) \right\}$

The exiting firms are in set **B**. *Figure 3,1* shows a division between viable and non-viable firms by the cut-off profit curve, in which the two subsets of productivity and fixed cost combinations are distinguished. The two sets are exhaustive dividing firms by their productivity and fixed cost combination.

(11′) $\mathrm{B} \equiv \left\{ (\varphi, f) \,\middle|\, f > f^*(\varphi) \right\}$

Notice that the cut-off profit curve is function of the elasticity of substitution (ε), aggregate revenue (R) and price (P). The cut-off profit curve can have different degrees of curvature, resulting in different realizations of productivity and fixed cost combinations where profit is zero: $\pi(\varphi, f^*(\varphi)) = 0$.

3.3.1 Uncertainty

Upon paying the entry fixed cost f_{en}, firms draw their fixed cost parameter $f(\omega)$ and their marginal productivity parameter $\varphi(\omega)$ from a common distribution with probability density function: $\kappa(\varphi, f)$. Let Φ be the support of φ and similarly F for f, which by economic sense must be non-negative with support $[0, \infty)$. Then the equivalent marginal probability density functions for productivity and fixed cost are given by: $\kappa_\varphi(\varphi) = \int_F \kappa(\varphi, f)\, df$ and $\kappa_f(f) = \int_\Phi \kappa(\varphi, f)\, d\varphi$.

If the productivity parameters and fixed cost are independently distributed, then the joint probability density function is the product of the two independent probability density functions: $\kappa(\varphi, f) = \kappa_\varphi(\varphi)\kappa_f(f)$. However, it would be naïve to assume the productivity and fixed cost to be independent of each other since a higher marginal productivity usually comes at a higher fixed cost. We can also think of higher marginal productivity as a production that produces higher quality. In this sense, firms with higher per period fixed cost produce product with higher quality. Therefore, we expect firms with higher marginal productivity to be associated with higher fixed costs. Nevertheless, without the information regarding the actual distribution of the two, we assume a general distribution form $\kappa(\varphi, f)$ to denote the *ex ante* joint probability density function of φ and *f*, and $\mu(\varphi, f)$ as the *ex post* joint probability density function. Hence, $\mu(\varphi, f)$ is the conditional distribution of $\kappa(\varphi, f)$ on set **A**:

(12) $$\mu(\varphi, f) = \begin{cases} \kappa(\varphi, f)/p_{in}, & (\varphi, f) \in A \\ 0 \quad , & otherwise \end{cases}, \text{ for } A \equiv \left\{ (\varphi, f) \,\middle|\, f \le f^*(\varphi) \right\} \quad \text{and}$$

$$p_{in}(f^*) \equiv \iint_A \kappa(\varphi, f)\, d\varphi\, df = \int_0^\infty \int_0^{f^*(\varphi)} \kappa(\varphi, f)\, df\, d\varphi = \int_0^\infty \int_{\varphi(f^*)}^\infty \kappa(\varphi, f)\, d\varphi\, df < 1$$

Since we assume that the subsequent firm exit is uncorrelated with firms' productivity but exogenously given, the exit process will not affect the equilibrium productivity distribution

$\mu(\varphi, f)$. Instead, the *ex post* distribution will be determined by the initial joint probability $\kappa(\varphi, f)$, conditioned on successful entry, as in equation 12, where p_{in} is the *ex ante* probability of successful entry for all potential entrants. By definition, p_{in} aggregates the probability over set A. Thus, in the second equality following p_{in}, we first aggregate the probability over the range $[0, f^*(\varphi)]$ for given φ and later the probability over all possible f, hence $[0, \infty)$, covering all viable firms. In short, p_{in} is a function of f^*.

3.3.2 Aggregation

We consider only the stationary equilibrium where aggregate variables remain constant over time. An entering firm with marginal productivity φ and fixed cost f will immediately exit if its profit level is negative, while a firm with non-negative profit will enter the market and continue its operation until hit by an adverse shock that forces it to exit with probability δ in each period. In the absence of time discounting, a firm's value function is:

$$(13) \qquad v(\varphi, f) = \max\left\{0, \sum_{t=0}^{\infty}(1-\delta)^t \pi(\varphi, f)\right\} = \max\left\{0, \frac{\pi(\varphi, f)}{\delta}\right\} \quad \text{(Value function)}$$

Here firms' optimal expected profits remain constant unless hit by an adverse shock. With δM firms simultaneously entering and exiting each period at the equilibrium, the *ex post* joint probability density function of productivity and fixed cost $\mu(\varphi, f)$ will not be affected. The equilibrium is characterized by a continuous mass of M viable firms (with equivalently M varieties produced).

We calculate the *ex post* weighted average productivity as the average productivity level for all viable firms. Recall that the cut-off profit level is: $\pi(\varphi, f^*(\varphi)) = 0$, the weighted average marginal productivity can be therefore written as the expected productivity $\varphi^{\varepsilon-1}$ conditional on a non-negative profit level, and similarly for the weighted average fixed cost (equation 14). Notice that the *ex post* weighted average productivity depends on the underlying joint distribution of the productivity and fixed cost, but independent of the number of active firms in the equilibrium market.

(14) $\tilde{\varphi} \equiv \left[E(\varphi^{\varepsilon-1} \middle| \pi \geq 0) \right]^{1/(\varepsilon-1)} = \tilde{\varphi}(\varphi, f) \quad \tilde{f} \equiv \left[E(f \middle| \pi \geq 0) \right] = \tilde{f}(\varphi, f)$

Here the weighted average productivity and profit can be understood as the productivity and profit level of the 'representative' agent in this heterogeneous world. Correspondingly, the associated weighted average price, quantity and revenue can be written as simple functions of the weighted productivity, while the weighted profit is also associated with the weighted fixed cost (equation 15).

(15) $\tilde{p} \equiv 1/\rho\tilde{\varphi}; \quad \tilde{q} \equiv Q\left(\tilde{p}/P\right)^{-\varepsilon}; \quad \tilde{r} \equiv \tilde{p}\tilde{q}; \quad \bar{\pi} = (\tilde{r}/\varepsilon) - \tilde{f} = (R/\varepsilon)\left(\rho P\tilde{\varphi}\right)^{\varepsilon-1} - \tilde{f}$

(16) $P = M^{1/(1-\varepsilon)} p(\tilde{\varphi}); \quad Q = M^{1/\rho} q(\tilde{\varphi}); \quad R = PQ = M\, r(\tilde{\varphi})$

The weighted productivity $\tilde{\varphi}$ and fixed cost $\tilde{f}$ summarize the information in the joint distribution of the productivity and fixed cost relevant for all aggregate variables. We rewrite the associated aggregate price, quantity, revenue, and profit levels as functions of the weighted marginal productivity $\tilde{\varphi}$ as in equation 16 (see detail in Technical note 4). The negative power attached to the number of active firms for the price index and the positive power attached to the number of active firms for the quantity index suggest positive externalities from the mass of active firms M at the aggregated level. This is a direct result of the love-of-variety effect embedded in the utility function (Brakman, Garretsen, and van Marrewijk, 2009, chapter 3).

The aggregate profit however is not solely related to the weighted average productivity, but is also related to the fixed cost. Depending on the fixed cost level, firms with the same marginal productivity can have different profit. We estimate the aggregate profit level with the information of the joint distribution of productivity and fixed cost. Firms' average expected profit is denoted as $\bar{\pi}(f^*)$, which at the same time captures the unspecified distribution.

(17) $\Pi = \int_0^\infty \int_0^{f^*(\varphi)} M\left(\frac{r(\varphi)}{\varepsilon} - f\right)\kappa(\varphi, f)\, df\, d\varphi = M\bar{\pi}(f^*)$

Substituting the aggregate variables P and R derived in equation 16 into the cut-off profit level (equation 10), we rewrite the cut-off fixed cost as a function of productivity:

$$(18) \qquad f^*(\varphi) = \left(\frac{r(\tilde{\varphi})}{\varepsilon}\right)\left(\frac{\varphi}{\tilde{\varphi}}\right)^{\varepsilon-1}$$

The f^* locus is important for determining the average viable firms' productivity. The location of f^* is independent of the mass of firms M in the stationary equilibrium, but depends on the productivity level and the underlying ex ante joint distribution of productivity and fixed cost. Denoting the part influenced by the distributions by $a = a(\tilde{\varphi})$, we further simplifies the cut-off fixed cost as:

$$(18') \qquad f^*(\varphi) = a\varphi^{\varepsilon-1}, \text{ where } a = \left(\frac{r(\tilde{\varphi})}{\varepsilon}\right)\left(\frac{1}{\tilde{\varphi}}\right)^{\varepsilon-1} = a(\tilde{\varphi}) \qquad \text{(Cut-off fixed cost)}$$

Notice that from equation 18 and 18′, we implicitly define $a = (R/\varepsilon)(\rho P)^{\varepsilon-1}$. It follows that a is a production parameter since: $q(\varphi) = Q(\rho\varphi P)^{\varepsilon} = (\varepsilon-1)a\varphi^{\varepsilon}$.

3.3.3 Entry and Exit

To construct a minimalist dynamic model, we impose an exogenous probability of firm exit (through some shock or bad luck) equal to δ, which is common to all firms. In each period, there are δM number of firms exiting the market. This is counter-balanced by an endogenous determination of firm entry. The active firms in the market are earning zero or positive profits as given in equation 7′. This makes the market attractive for potential entrants to enter.

There are two caveats. First, potential entrants are identical as they do not yet know their productivity level before entry. Second, to enter the market, firm have to incur a one-time fixed entry cost f_{en}, which is thereafter sunk. Once the entry cost is paid, the remaining uncertainty regarding firms' productivity and fixed cost level is revealed. The productivity and fixed cost once drawn do not change. The firms that earn non-negative profits after the uncertainty is resolved will therefore earn the same optimal profit level over time until hit by a shock that forces them to exit. In the equilibrium, there are δM number of firms simultaneously entering and exiting the market.

3.3.4 Equilibrium in the closed economy

Two conditions are crucial for determining the equilibrium of this economy, (i) the viability condition and (ii) the free entry condition. Since each particular fixed cost suggests different cut-off productivity, the combinations of different productivities and fixed costs result in different profit levels. Therefore, unlike the setting in Melitz model (2003) with a unique fixed cost level and a unique cut-off productivity level, the equilibrium derived in this paper is solved with a continuum level of cut-off fixed cost with respect to different productivity level. The equilibrium condition is underpinned by two equilibrium conditions: the viability condition and the free entry condition.

The viability condition (VC) provides the relationship between the representative firm's profit level $\bar{\pi}$ and the cut-off fixed cost f^*. It states the representative firm's expected profit level, which had already been implicitly defined in equation 17.

$$(19) \qquad \bar{\pi}(f^*) \equiv \int_0^{\infty} \int_0^{f^*(\varphi)} \left(\frac{r(\varphi)}{\varepsilon} - f \right) \kappa(\varphi, f)\, df\, d\varphi \qquad \text{(VC)}$$

The free entry condition (FE) provides the relationship between expected profit $\bar{\pi}$ and the attractiveness of entry. Note that a successful entrant will earn an average expected profit $\bar{\pi}$, with probability of termination equal to δ in each period. The associated expected net present value, $\bar{v}$ say, over all time periods *t* is equal to:

$$\bar{v} = \sum_{t=0}^{\infty} \frac{\tilde{\pi}}{(1-\delta)^t} = \frac{\bar{\pi}}{\delta}.$$

The equilibrium requires that the net value of entry equals zero: $\bar{v} p_{in} - f_{en} = (\bar{\pi}/\delta) p_{in} - f_{en} = 0$. This is because that no firm will want to enter if the net value of entry is negative and that it cannot be positive either as there is an unlimited pool of potential profit seeking entrants. Thus, the net value of entry must be zero in equilibrium, which again links the average profit level with the cut-off fixed cost. We call this the free entry condition (FE):

$$(19) \qquad \bar{\pi}(f^*) = \frac{\delta f_{en}}{\iint_A \kappa(\varphi, f)\, d\varphi\, df} = \frac{\delta f_{en}}{p_{in}(f^*)} \qquad \text{(FE)}$$

In equilibrium, the aggregate variables remain constant over time. To ensure a constant output supplied in the market, it requires a (sufficiently large) mass of potential new entrants M_e with probability of successful entry p_{in} entering in each period to replace the exiting incumbents that are hit by bad luck δM. Since the market is competitive and the incumbents earn non-negative profits, the market remains attractive to enter. With sufficient number of potential entrants, $p_{in}M_e = \delta M$ always hold.

Labor market requires that the labor supply in the economy is fully matched to the labor demand in the economy. The aggregate labor supply (L) available for production (and demand) reflects the size of the economy and is the major and only factor constraining the size of the entire economic activity. The aggregated labor supply is divided into two parts: $L = L_{in} + L_{en}$, including the labor employed by the incumbent firms (L_{in}) and the labor employed by the potential entrant firms (L_{en}). Labor employed by the incumbent firms are paid with the differences between aggregate revenue and profit: $L_{in} = R - \Pi$. This is also the market clearing condition for these production workers working for the incumbent firms.

The market clearing condition for labor employed by the new entrant firms requires $L_{en} = M_e f_{en}$ to hold. Using the aggregate stability condition $p_{in}M_e = \delta M$, and the free entry condition, the number of potential entrants L_{en} can be written as: $L_{en} = f_{en}M_e = f_{en}\delta M / p_{in}$ $= M\bar{\pi} = \Pi$. Hence, the aggregate revenue R must equal the total payment to labor L, which is exogenously fixed by the size of the economy: $R = L_{in} + \Pi = L_{in} + L_{en} = L$. The number of firms in any period can be determined from the average profit level by rearranging equation 15 and 16.

(20) $M = R / \tilde{r} = L / \varepsilon(\bar{\pi} + \tilde{f})$ (Number of firms)

With the number of firms determined, we complete the characterization of the stable equilibrium in the closed economy.

3.4 Open economy

We now expand the model with trade opportunity and analyze the impact of trade in a world that is composed of two symmetric countries. As in Krugman (1980) and Melitz (2003), if there are no additional costs to trade, individual countries can replicate the outcome of the integrated world economy, which is equivalent to an increase in the size of the market such that consumers enjoy welfare gains through an increase in product variety without firm-level effect.

In line with the empirical evidence (see Roberts and Tybout, 1997; Bugamelli and Infante, 2003) and the heterogeneity literature, we know that expending sale to foreign markets is not without extra costs. Besides per unit cost such as tariff and transportation cost, which we assume to be common for all firms, a huge fixed cost investment is inevitable for firms to become exporters. This included the sunk start-up cost which is needed to learn the bureaucratic procedure, to establish the distribution channel, and to adapt the product (taste/packaging…ect.) for the foreign market. Moreover, there is the per period fixed cost that needs to be covered, including the cost in maintaining a presence in foreign market and mounting to foreign custom procedure and product standard. The important notion we want to argue here is that per period exporting fixed cost bared by each firms also differs. Besides the differences in ability to cost down in the above mentioned activities, some firms may benefit from having spontaneous connection to foreign buyers in a trade exposition while others have managers experienced with exporting activities. In other words, the advantageous position on firms' exporting fixed cost is unevenly distributed across firms. Under such circumstances, this exporting fixed cost is not necessary related to firm's productivity, but may be positively linked to the domestic fixed cost invested for production.

The standard iceberg transportation cost is applied, whereby a fraction of $\tau > 1$ units of good must be shipped in order for 1 unit to arrive. In addition, there is the fixed export cost per period f_x. Since there is no additional time discounting other than the exit shock with probability δ, this amortized fixed cost per period is equivalent to an up-front entry cost of f_x / δ. We assume that the exporting fixed cost is related to the domestic fixed cost by the following relation $f_x = \eta f$. For $0<\eta<1$ means that the exporting fixed cost firms bear per period to serve the foreign market is less than the fixed cost necessary for the production in the domestic market. With the additional fixed and variable cost associated with export sales, only the most productive firms will find it profitable making additional sales in the foreign market.

We focus on two identical countries (henceforth have the same wage, which is normalized to one). This is because allowing multiple identical countries to trade within the model will not add much unless geographical differences between countries are being considered. That is, unless we assume the transportation cost between each pairs of countries in trade to be heterogeneous, the value added for introducing additional identical trading countries is limited. While the inclusion of the possibility to trade with multiple countries will only strengthen our assumption in modeling with heterogeneous fixed exporting cost. This is because as firms exporting to more countries, the fixed cost is likely to increase disproportionally as they enter more distant or less familiar markets (due differences in terms of language, culture and business negotiation rules). To focus on the heterogeneous domestic and exporting fixed cost dimension and elaborate the importance of modeling with heterogeneous fixed cost, we model the trade economy with only one other symmetric country.

3.4.1 Consumption and production

The export decision is made only after uncertainty is resolved (so φ and f are known). For firm with marginal productivity φ, profit maximization implies that equilibrium pricing rule for the domestic market is a constant mark-up over the marginal cost: $p_d(\varphi) = w/\rho\varphi = 1/\rho\varphi$ (wage normalized to one). The export price with additional variable cost to trade is directly added to the mark-up for the foreign buyer: $p_x(\varphi) = \tau/\rho\varphi = \tau\, p_d(\varphi)$.

Given firms' optimal pricing rule, equilibrium revenue obtained from the domestic market is $r_d(\varphi) = R(P\rho\varphi)^{\varepsilon-1}$, where R and P are aggregate revenue and price index of the home market. Revenue from the exporting sales is proportional to that in the domestic market due to pricing differences between the two markets $r_x(\varphi) = R\left(P\rho\varphi/\tau\right)^{\varepsilon-1} = \tau^{1-\varepsilon}\, r_d(\varphi)$. In view of the balance of payments condition, the aggregate revenue R is equivalent to the aggregate income. The combined revenue for firms with marginal productivity φ is:

(21) $$r_{tr}(\varphi) = \begin{cases} r_d(\varphi) & \text{if does not export} \\ r_d(\varphi) + r_x(\varphi) & \text{if export} \end{cases}$$

Due to the fixed production cost, no firms will export without also producing for the domestic market. Therefore, all viable firms are those that at least produce for the domestic market, while the firms capable of making non-negative profits from exporting will make sales abroad. Accounting for the overhead production cost for the domestic production and export activity, we separate each firm's profit into components earned from the domestic sales π_d and foreign sales π_x; where the share of production fixed cost is deducted from the domestic revenue and the exporting fixed cost is deducted from the foreign revenue. For firm with productivity φ and fixed costs f, the profit from domestic and foreign market is:

(22) $$\pi_d(\varphi, f) = \frac{r_d(\varphi)}{\varepsilon} - f$$ (Domestic profit)

(23) $$\pi_x(\varphi, f) = \frac{r_x(\varphi)}{\varepsilon} - f_x = \frac{\tau^{1-\varepsilon} r_d(\varphi)}{\varepsilon} - \eta f$$ (Export profit)

A firm exports to the other market if profit made from export is non-negative ($\pi_x(\varphi, f) \geq 0$). The total firm profit in the open trade economy is therefore:

(24) $$\pi_{tr}(\varphi, f) = \pi_d(\varphi, f) + \max\{0, \pi_x(\varphi, f)\}$$ (Total profit)

The value function for each firm is: $v(\varphi, f) = \max\{0, \pi_{tr}(\varphi, f)/\delta\}$.

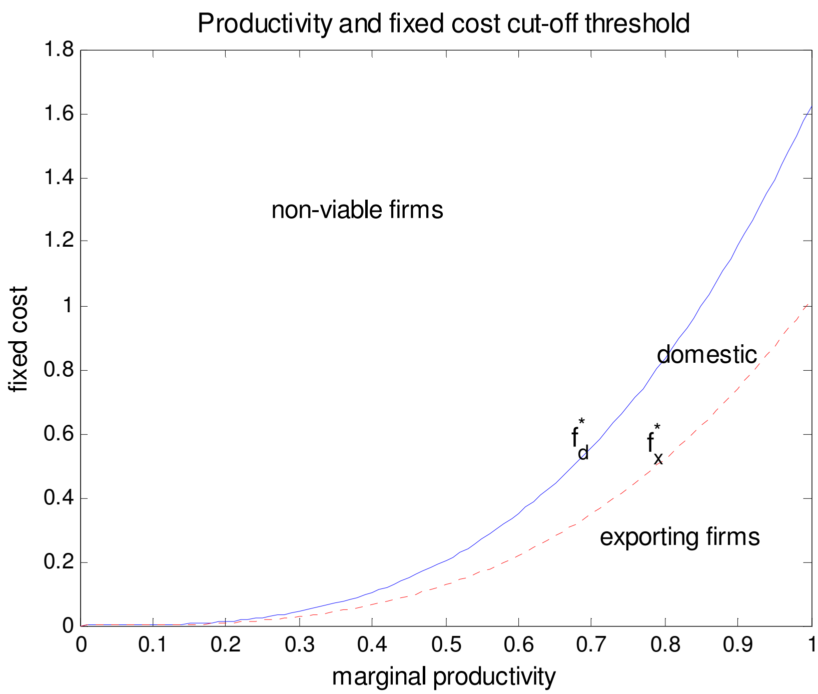
3.4.2 Entry and exit decision and exporter selection

Similar to the closed economy case, the cut-off productivity level for identifying successful entry firms and exporters is not unique. Instead, we have two *zero-profit curves* that pin down a series firms with productivity and fixed cost combinations that earn zero profit in either the domestic market or the exporting market. The first, <u>zero-profit *viability* curve</u> identifies those firms that make zero profit from production and are those at the margin of exiting the market. Second, different from the closed economy case, <u>zero-profit *exporting* curve</u> further distinguishes the firms that are indifferent between being exporter and non-exporter as they make zero profit from the exporting activity. The zero-profit viability and exporting curves (ie. the two cut-offs) are both functions of the marginal productivity (equation 25).

(25) $$\begin{aligned} f_d^*(\varphi) &= (R/\varepsilon)\left(P\rho\varphi\right)^{\varepsilon-1} && \Leftrightarrow \quad \pi_d(\varphi, f_d^*(\varphi)) = 0 \\ f_x^*(\varphi) &= (\tau^{1-\varepsilon}/\eta) f_d^*(\varphi) < f_d^*(\varphi) && \Leftrightarrow \quad \pi_x(\varphi, f_x^*(\varphi)) = 0 \end{aligned}$$

The relationship identified from equation 25 divides the (φ, f) -plane into three areas (Figure 3.2). Firms with different productivity and fixed cost combination fall in different areas divided by the two zero-profit curves. Those firms with (φ, f) to the left of the viability curve, $f_d^*(\varphi)$, will not be able to make economic profit in the domestic market (non-viable) and therefore exit; while firms with productivity and fixed cost combination to the right of the exporting curve, $f_x^*(\varphi)$, make non-negative profit in both the domestic and foreign market. Firms that lie in the area between the two curves are those with productivity and fixed cost combinations that allow them to make economic production only in the domestic market. The partitioning of firms by exporting status occurs if and only if $f_d^*(\varphi) > f_x^*(\varphi)$. Base on abundant empirical evidence and as in Melitz (2003), only a fraction of viable firms make additional investment to engage in positive export activity, therefore the following parameter restriction: $\eta\tau^{\varepsilon-1} > 1$ must hold[21].

Figure 3.2 *The zero-profit viability and exporting curve*



As before, the *ex-ante* probability of successful entry into the market is given by:

---

[21] Since $\tau^{1-\varepsilon} r_d(\varphi)/\varepsilon\eta = f_x^* < f_d^* = r_d(\varphi)/\varepsilon$, therefore $\tau^{1-\varepsilon}/\eta < 1$.

(26) $$p_{in} \equiv \int_0^{\infty} \int_0^{f_d^*(\varphi)} \kappa(\varphi, f)\, df\, d\varphi \le 1$$

Likewise, the *ex-ante* probability of successful entry as an exporting firm is given by equation 27, with the cut-off threshold set to the cut-off fixed exporting cost determine in equation 25.

(27) $$\overline{p}_x \equiv \int_0^{\infty} \int_0^{f_x^*(\varphi)} \kappa(\varphi, f)\, df\, d\varphi < p_{in}$$

The share of exporting firms among the viable firms is thus: $\overline{p}_x / p_{in} \equiv p_x$.

3.4.3 Uncertainty and aggregation

Provided with the *ex-ante* joint distribution of firm productivity and fixed cost, the *ex-post* distribution of incumbent firms is a subset of firms that survived in the competition.

(28) $$\mu(\varphi, f) = \begin{cases} \kappa(\varphi, f) / p_{in}, & f \le f_d^*(\varphi) \\ 0 & , \; otherwise \end{cases},$$

Similarly, the *ex-post* probability density function for successful exporters is:

(29) $$\chi(\varphi, f) = \begin{cases} \kappa(\varphi, f) / \overline{p}_x, & f \le f_x^*(\varphi) \\ 0 & , \; otherwise \end{cases}.$$

In equilibrium, the number of firms in a country is M and the number of exporting firms is $M_x = p_x M$. The total number of active firms selling goods to consumers in either market is thus $M_{tr} = M + M_x = (1 + p_x) M$, which is equivalent to the number of varieties available. Now let $\tilde{\varphi}_d$ denote the weighted average marginal productivity of all viable firms and $\tilde{\varphi}_x$ as the weighted average marginal productivity of exporting firms. As long as the partitioning condition holds, it insures only the most efficient firms make non-negative profit in the exporting market. Therefore, the weighted average marginal productivity of exporters must be greater than the weighted average marginal productivity of all viable firms.

When the most efficient firms export their product and enter the foreign market, these exporting goods no longer preserved their most competitive position, for these products being negatively adjusted by the transit cost. Therefore it is important to account for the discounted competitiveness (productivity) of the importing products when they enter the distant market. We define the weighted productivity of all importing goods as the productivity discounted by the marginal exporting cost (equation 30). This term will be used later for ease of comparison. Note that the difference between $\tilde{\varphi}_x$ and $\tilde{\varphi}_i$ can also be understood as the damaged product loss during transportation or as the difference between CIF (cost, insurance and freight) and FOB (free on board) price in business terminology.

(30)
$$\tilde{\varphi}_d = \left[E(\varphi^{\varepsilon-1}) \middle| \pi_d \geq 0\right]^{1/(\varepsilon-1)} = \left[\int_0^{\infty}\int_0^{f_d^*} \varphi^{\varepsilon-1}\mu(\varphi, f)\, df\, d\varphi\right]^{1/(\varepsilon-1)};$$
$$\tilde{\varphi}_x = \left[E(\varphi^{\varepsilon-1}) \middle| \pi_x \geq 0\right]^{1/(\varepsilon-1)} = \left[\int_0^{\infty}\int_0^{f_x^*} \varphi^{\varepsilon-1}\chi(\varphi, f)\, df\, d\varphi\right]^{1/(\varepsilon-1)};$$
$$\tilde{\varphi}_i = \left[E((\varphi/\tau)^{\varepsilon-1}) \middle| \pi_x \geq 0\right]^{1/(\varepsilon-1)} = \left[\int_0^{\infty}\int_0^{f_x^*} (\varphi/\tau)^{\varepsilon-1}\chi(\varphi, f)\, df\, d\varphi\right]^{1/(\varepsilon-1)}$$

Taken the effect of allowing foreign varieties in the domestic market, $\tilde{\varphi}_{tr}$ denotes the weighted average marginal productivity of all varieties competing in a market.

(31)
$$\tilde{\varphi}_{tr} = \left\{\frac{1}{M_{tr}}\left[M\,\tilde{\varphi}_d^{\,\varepsilon-1} + M_x(\tilde{\varphi}_x/\tau)^{\varepsilon-1}\right]\right\}^{1/(\varepsilon-1)}$$

We rewrite the aggregate price index P and expenditure level R as functions of the weighted productivity $\tilde{\varphi}_{tr}$ and equilibrium numbers of varieties competing in a market:

(32)
$$P_{tr} = M_{tr}^{1/(1-\varepsilon)}\, p(\tilde{\varphi}_{tr}), \qquad R = M_{tr} r_d(\tilde{\varphi}_{tr})$$

These aggregate expressions of price and revenue are use to rewrite the domestic cut-off fixed cost $f_d^*(\varphi)$ from equation 25.

(33)
$$f_d^*(\varphi) = \left(\frac{r_d(\tilde{\varphi}_{tr})}{\varepsilon}\right)\left(\frac{\varphi}{\tilde{\varphi}_{tr}}\right)^{\varepsilon-1}$$

The relationship implies that the cut-off domestic fixed cost curve is independent of the mass of firms M in a stationary equilibrium, but a function of productivity and the production parameter a_{tr} in the trade equilibrium. Note that the production parameter is constant since the weighted productivity $\tilde{\varphi}_{tr}$ is a definite value in the equilibrium.

(34) $$a_{tr}(\tilde{\varphi}_{tr}) \equiv \left(\frac{r_d(\tilde{\varphi}_{tr})}{\varepsilon}\right)\left(\frac{1}{\tilde{\varphi}_{tr}}\right)^{\varepsilon-1}$$

As a result, the fixed cost cut-off for viability selection and exporter selection in equation 25 can be further simplified as:

(25') $$f_d^*(\varphi) = a_{tr}\varphi^{\varepsilon-1} \text{ and } f_x^*(\varphi) = (\tau^{1-\varepsilon}/\eta)f_d^*(\varphi) = (\tau^{1-\varepsilon}/\eta)a_{tr}\varphi^{\varepsilon-1}.$$

Lastly, we make use of the production parameter a_{tr} to express the average profit firms earn from the domestic market $\bar{\pi}_d(a_{tr})$ and from the foreign market $\bar{\pi}_x(a_{tr})$:

(35) $$\begin{aligned} \bar{\pi}_d(a_{tr}) &\equiv \int_0^{\infty}\int_0^{f_d^*(a_{tr})}\left(\frac{r_d(\varphi)}{\varepsilon} - f\right)\mu(\varphi, f)\,df\,d\varphi \\ \bar{\pi}_x(a_{tr}) &\equiv \int_0^{\infty}\int_0^{f_x^*(a_{tr})}\left(\frac{r_x(\varphi)}{\varepsilon} - f_x\right)\chi(\varphi, f)\,df\,d\varphi \end{aligned}$$

3.4.4 Equilibrium in the open economy

As in the closed economy case, the viability condition and free entry condition pins down the equilibrium. The viability condition (VC) identifies the relationship between the average profit and the production parameter in trade in the equilibrium:

(36) $$\bar{\pi}_{tr}(a_{tr}) = \bar{\pi}_d + p_x\bar{\pi}_x \qquad \text{(VC)}$$

The average profit for any domestic firm in the open economy is the sum of the average profit from the domestic market and the additional profit from the exporting activity. Since both components in the average trade profit are function of the production parameter a_{tr}, the average profit determined from the viability condition is also a function of a_{tr}.

The free entry condition (FE) is comparable to that of the closed economy but change endogenously in accordance with the probability of entry, where the probability of entry is implicitly defined as function of the production parameter in trade: $p_{in}(f_d^*(a_{tr}))$. Let $\bar{\upsilon}$ be the present value of average profit flows of viable firms: $\overline{\upsilon} = \overline{\pi}_{tr} / \delta$. With unrestricted entry and an unbounded mass of potential entrants, the equilibrium requires that the cost of entry equals the expected net value of entry, $f_{en} = \overline{\upsilon} p_{in}$, which is the average viable firms' present value multiply by the probability of successful entry. Thus we have the *Free Entry* (FE) condition, in which the average profit is a function of the production parameter a_{tr}, while the exit rate and sunk cost of entry are given exogenously:

$$\overline{\pi}_{tr}(a_{tr}) = \frac{\delta f_{en}}{p_{in}(f_d^*(a_{tr}))}, \tag{37} \quad \text{(FE)}$$

As in the closed economy case, the numbers of firms entering and the number of firms exiting the market by exogenous shock must equal in the steady-state equilibrium $p_{in} M_e = \delta M$. Likewise, the aggregated fixed entry cost must equal to the aggregated excessive profits incumbents earn $f_{en} M_e = L_{en} = \Pi$, while the aggregated revenue remains exogenously fixed by the size of the economy $R = L$. Therefore, by dividing total revenue by the revenue per firm gives the equilibrium numbers of viable firms:

$$M = \frac{M_{tr}}{1 + p_x} = \frac{R}{r_d(\tilde{\varphi}_{tr})} \frac{1}{1 + p_x} = \frac{R}{\overline{r}_{tr}} \tag{38}$$

Note that $r_d(\tilde{\varphi}_{tr})$ denotes the revenue earn per varieties, which is different from the average revenue earn per firm $\overline{r}_{tr}$ (see Technical note 5 for disambiguate).

3.4.5 Welfare

Welfare defined as the inverse of the aggregated price index $W = P^{-1}$. Since the price index are the weighted averages of the prices charged by individual firms with different productivity (with the weights determined by the *ex post* productivity distribution), the welfare is thus positively related to the weighted productivity in trade and numbers of varieties in the equilibrium.

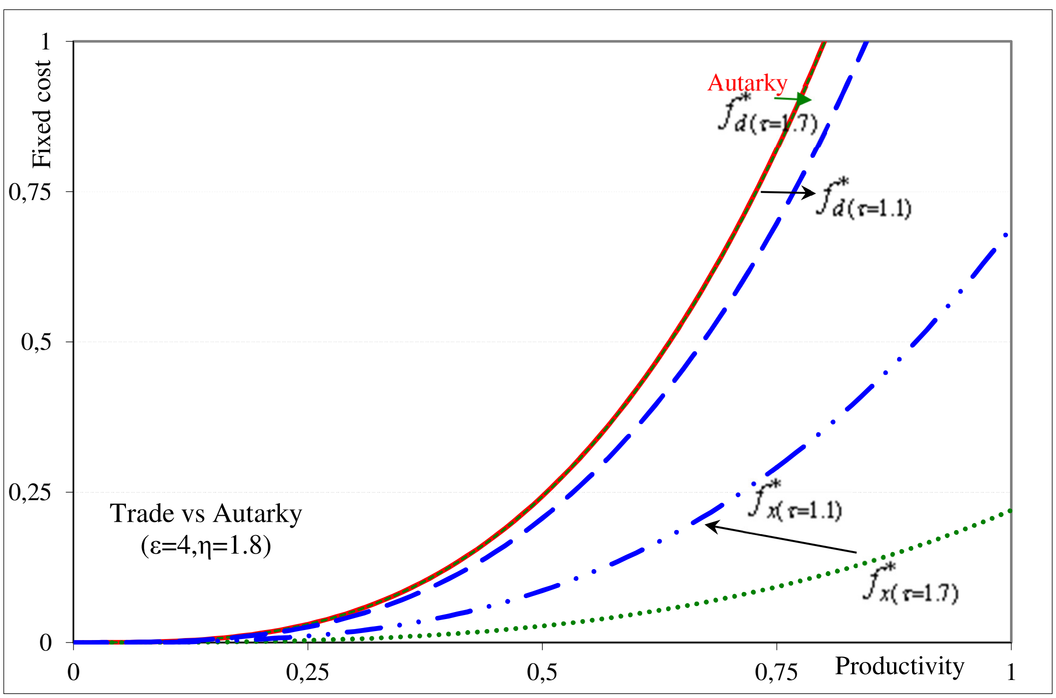
$$W = P^{-1} = M_{tr}^{1/(\varepsilon-1)} p(\tilde{\varphi}_{tr})^{-1} = M_{tr}^{1/(\varepsilon-1)} \rho \tilde{\varphi}_{tr} = (a_{tr} \varepsilon / R)^{1/(1-\varepsilon)} \rho \tag{39}$$

3.5 Impact of trade

The current section analyzes the effect of trade by comparing the autarky equilibrium with trade equilibriums. Due to the complexity of the model, simulated numerical solution is provided to analyze the result. We report the changes in the equilibrium production parameter a_{tr} with decreasing trade barriers and compare the result with the equilibrium a in autarky.

Recall that the production parameter is positively associated with the zero-profit viability and exporting curve (equation 25'). Smaller production parameter ($a_{tr} < a$)in trade equilibrium imply a stronger viability selection since at any level of marginal productivity, the maximum fixed cost viable firms can have in trade is lower than in autarky. Thus, we see that as transportation cost decreases, the zero-profit viability curve rotate to the right, while the zero-profit exporting curve rotate to the left —suggesting stronger viability selection and increasing numbers of firms becoming exporters (Figure 3.3).

Figure 3.3 The lower the transportation cost, the stronger the viability selection.



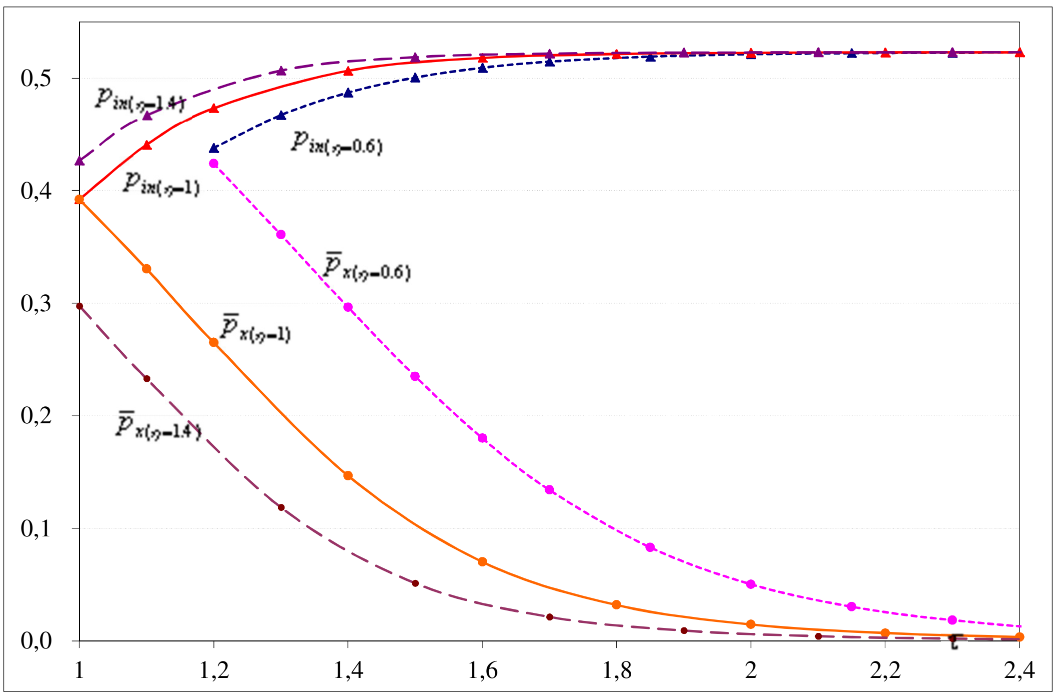
The consecutive sections show how gradual trade liberalization from decreasing transportation cost at various exporting fixed cost level affect the selection of viable firms and exporting firms. Note that the result presented shows the long-run steady state equilibrium after all mechanism described above is at work. Furthermore, the figures presented in this section are based on simulation results with exogenous parameters $\varepsilon = 4$, $\delta = 0.1$ and $f_{en} = 2$. We construct the underlying distribution by taking the product of two independent beta distributions for fixed cost and productivity respectively: $\kappa(\varphi, f) = \kappa(f; \alpha_f, \beta_f)\kappa(\varphi; \alpha_\varphi, \beta_\varphi)$ with shape parameters: $\alpha_f = 3, \beta_f = 5$, $\alpha_\varphi = 4$ and $\beta_\varphi = 3$.

3.5.1 Trade liberalization, decreasing transportation cost

When a country is open to trade, the most productive firms start to export. With the zero profit exporter selection being very stringent at the beginning of opening to trade, only a small percentage of viable firms are able to make non-negative profits from export ($\overline{p}_x$, Figure 3.4). Successful exporters are either those with high productivity or those with low fixed cost, or the firms with both qualities. At high level of transportation cost, foreign products' competitiveness is heavily discounted when entering domestic market. As a result, the weighted productivity of these imported goods is much smaller than the weighted productivity of the domestic goods at high level of transportation cost (Figure 3.5). Despite the weighted productivity of imports being significantly lower than the weighted productivity of the domestic varieties, the negative influence on the overall productivity is small when the transportation cost is high (Figure 3.5; eq.31). In addition, despite that weighted productivity of imports start with a disadvantage, the least productive product among these imports is still more productive than the average productivity of domestic products before exported.

As trade further liberalizes through decreasing transportation cost, the viability selection become more stringent as the competition intensity increased. With more completive products entering and sharing the market, the zero profit viability selection becomes stronger, resulting in a downward pressure in the equilibrium production parameter a_{tr}. As a result, the probability of entry for the domestic firms decreases, while the probability of becoming an exporter increases (movement along p_{in} and $\overline{p}_x$, Figure 3.4). In addition, we see that the weighted productivity of the domestic products increases ($\tilde{\varphi}_d$, Figure 3.5). The weighted productivity of the exporting products in contrast decreases since the zero profit exporting selection is weakened as trade liberalize ($\tilde{\varphi}_x$, Figure 3.5). Regardless of the decreasing trend, the weighted productivity of the same goods that become imports in the foreign country is increasing as the transportation cost decreases. The weighted productivity difference between export goods and the perceived weighted productivity of imported goods (ie. ($\tilde{\varphi}_x$ and $\tilde{\varphi}_i$) is caused by the transportation cost. Thus, as transportation cost decreases, the discrepancy between the two exhausted. The overall effect on the weighted productivity in trade $\tilde{\varphi}_{tr}$ will first decreases before increasing again. This holds for all levels of exporting fixed cost (Figure 3.5). Despite the non-monotonic development in the weighted productivity, the equilibrium production parameter a_{tr} decreases monotonically as the transportation cost decreases (Figure 3.6).

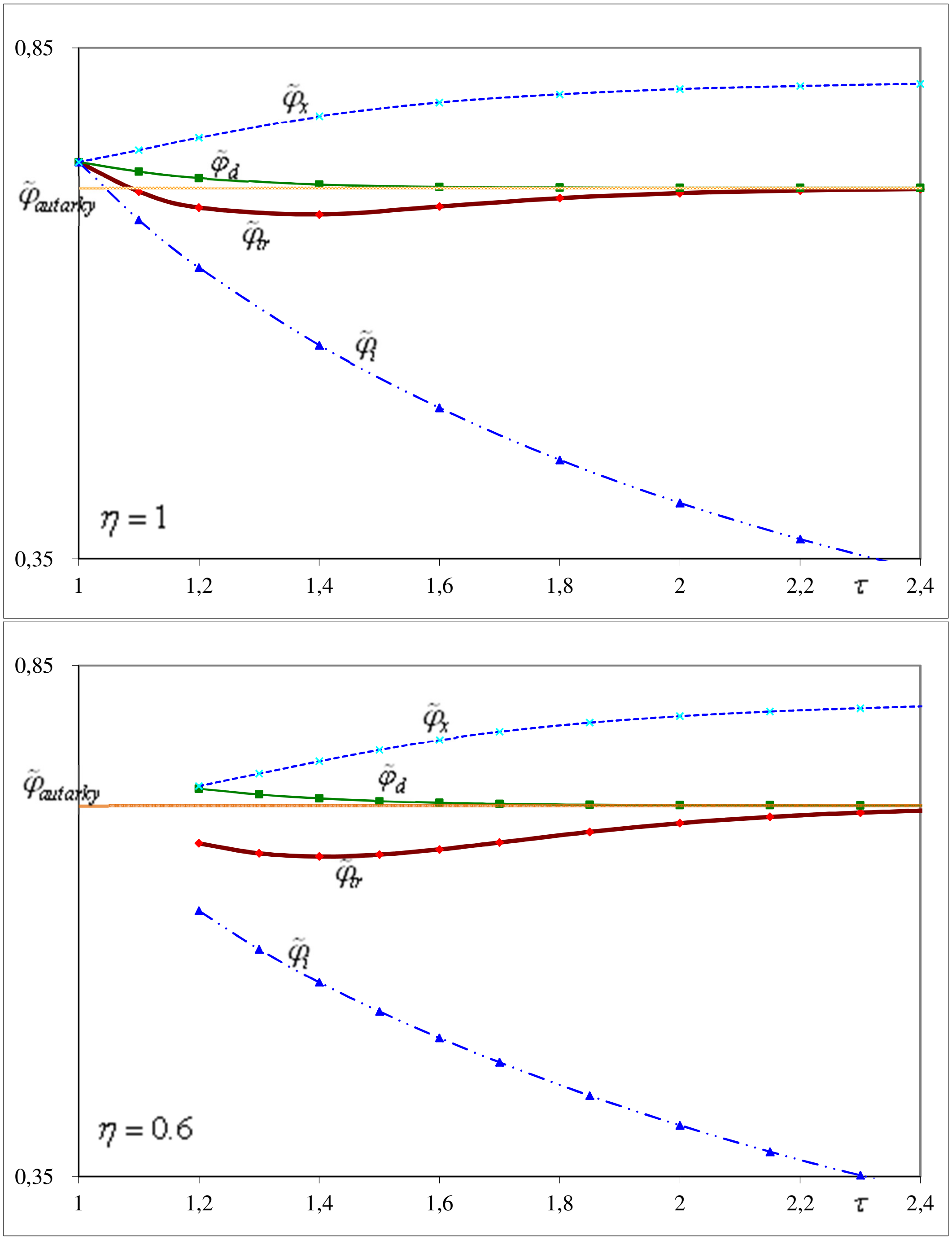
Figure 3.4 Development of probability of entry and the ex ante probability of successful entry of exporting firms for different exporting fixed cost

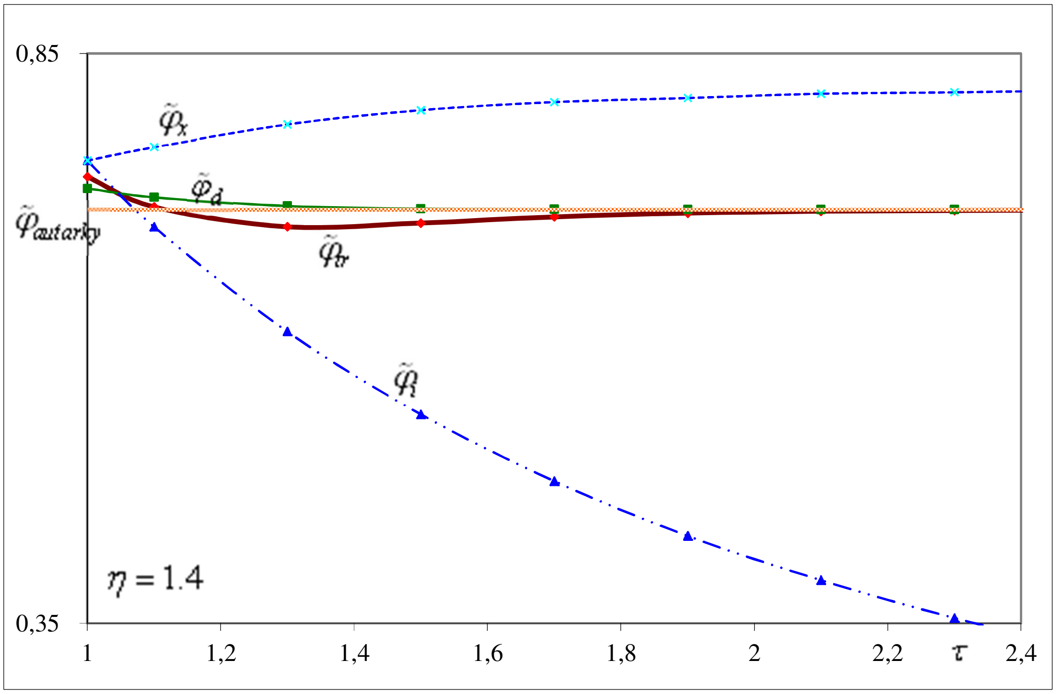


3.5.2 Trade liberalization— the effect of different exporting fixed cost

Having an overview of the general development of the weighted productivities as the transportation cost decreases, we move on to compare the effect of exporting fixed cost on the development of the weighted productivity. As the transportation cost decreases, the decreasing intensity of the exporter selection is relaxed more quickly when the exporting fixed cost is less than proportion to the domestic fixed cost ($\eta < 1$). In other words, at any given level of transportation cost, the probability of becoming an exporter is higher when the exporting fixed cost is low: $\bar{p}_{x(\eta<1)} > \bar{p}_{x(\eta=1)} > \bar{p}_{x(\eta>1)}$ (Figure 3.4). In short, all else equal, the smaller the exporting fixed cost, the stronger the exporter selection (smaller a_{tr}), ie. $f^*_{x(\eta<1)} > f^*_{x(\eta=1)} > f^*_{x(\eta>1)}$ (Figure 3.6).

Figure 3.5 Development of weighted productivity and probability of entry





Since the percentage of exporters is the weight used in determining the weighted productivity in trade, the negative pressure of weighted productivity of imports on the overall weighted productivity in trade is greater when the exporting fixed cost is low (ie. the percentage of firms exporting is high). In addition, since the weighted productivity of imports is always smaller than the weighted productivity of domestic varieties when the exporting fixed cost is not greater than one ($\eta \leq 1$), the reverse increasing trend of the weighted productivity in trade is limited as the transportation decreases (Figure 3.5a). In contrast, the weighted productivity of the imports will at some point become greater than the weighted productivity of the domestic varieties, resulting in an overshoot in the reverse development of the weighted productivity above the weighted productivity in autarky (Figure 3.5c).

To summarize, the overall weighted productivity in trade is greatest when the fixed exporting fixed cost is a bit more than proportion to the domestic fixed cost (in our simulation scenario, when η is approximately 1.6) and when the transportation cost is zero ($\tau = 1$). The equilibrium development of the production parameter a_{tr} decreases monotonically as the transportation cost decreases irrespective of the exporting fixed cost. The smallest a_{tr} is achieved when there is no transportation cost (free trade) and when the exporting fixed cost equal to the domestic fixed cost (Figure 3.6).

3.5.3 Welfare implication

The results above suggest that free trade environment as compared to autarky always imposes a stronger viability selection, while allowing more firms to become exporters. During the transition of trade liberalization, we observe an optimizing intra-sector resource reallocation as market share is reallocated from the less competitive firms to the more competitive ones. The numbers of domestic firms in the trade equilibrium is less than in autarky. The decreasing number of less efficient product varieties produced by domestic firms is replaced by an increasing number of foreign varieties.

The overall welfare is evaluated based on a combined measure of the price and varieties consumers enjoy. With the price as the inverse of the weighted productivity, consumers experience an increase in welfare as trade liberalizes (Figure 3.6). At the same time, consumers observe increasing product varieties as the country move from autarky to an open economy. Despite the less favorable price change, the gain from the increase in product variety out-weights the loss during the transition. As a result, the welfare in trade is always greater than in autarky and trade liberalization is always welfare improving.

3.5.4 Correlated fixed cost and productivity

Before closing up the discussion, we briefly explores how the assumption regarding the ex ante joint distribution between the productivity and fixed cost affect the *ex post* distribution of marginal productivity (Figure 3.7). Below we show the simulation result based on independent distributed productivity and fixed cost (left column) and one with the productivity and fixed cost jointly, positively correlated distributed (right column). The simulation result of the *ex post* marginal productivity distribution shows the lack of lower bound cut-off for viability selection. From the upper to lower panel, we see that decreasing transportation cost is closely related to smaller differences between the *ex post* marginal productivity distribution of the viable firms and *ex post* marginal productivity distribution of the exporters, as expected. When the transportation cost is small, most viable firms are also exporting. This holds for both types of distribution, independent or jointly correlated marginal productivity and fixed cost.

Figure 3.6 Equilibrium trajectory of a_{tr} at various level of exporting fixed cost.

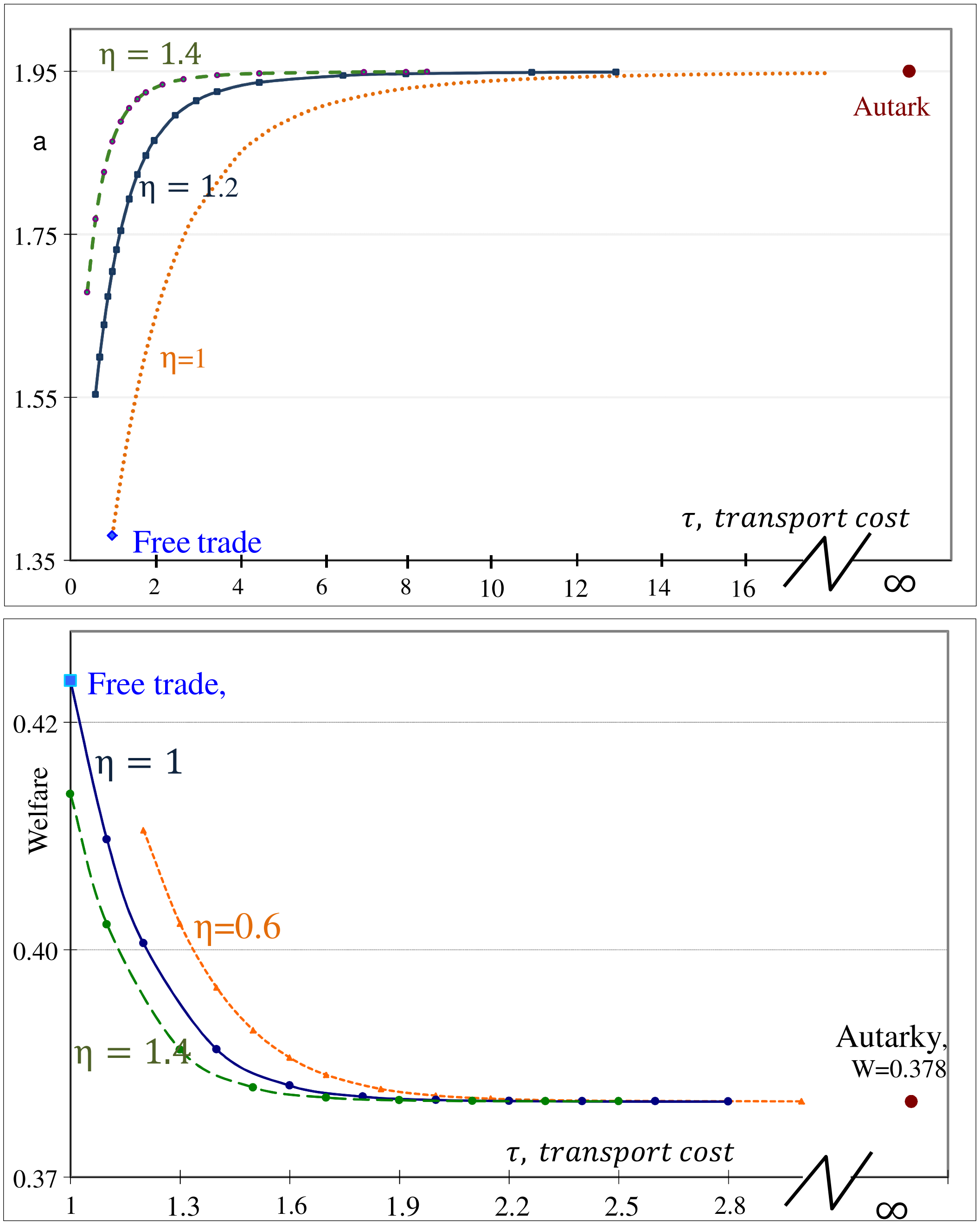
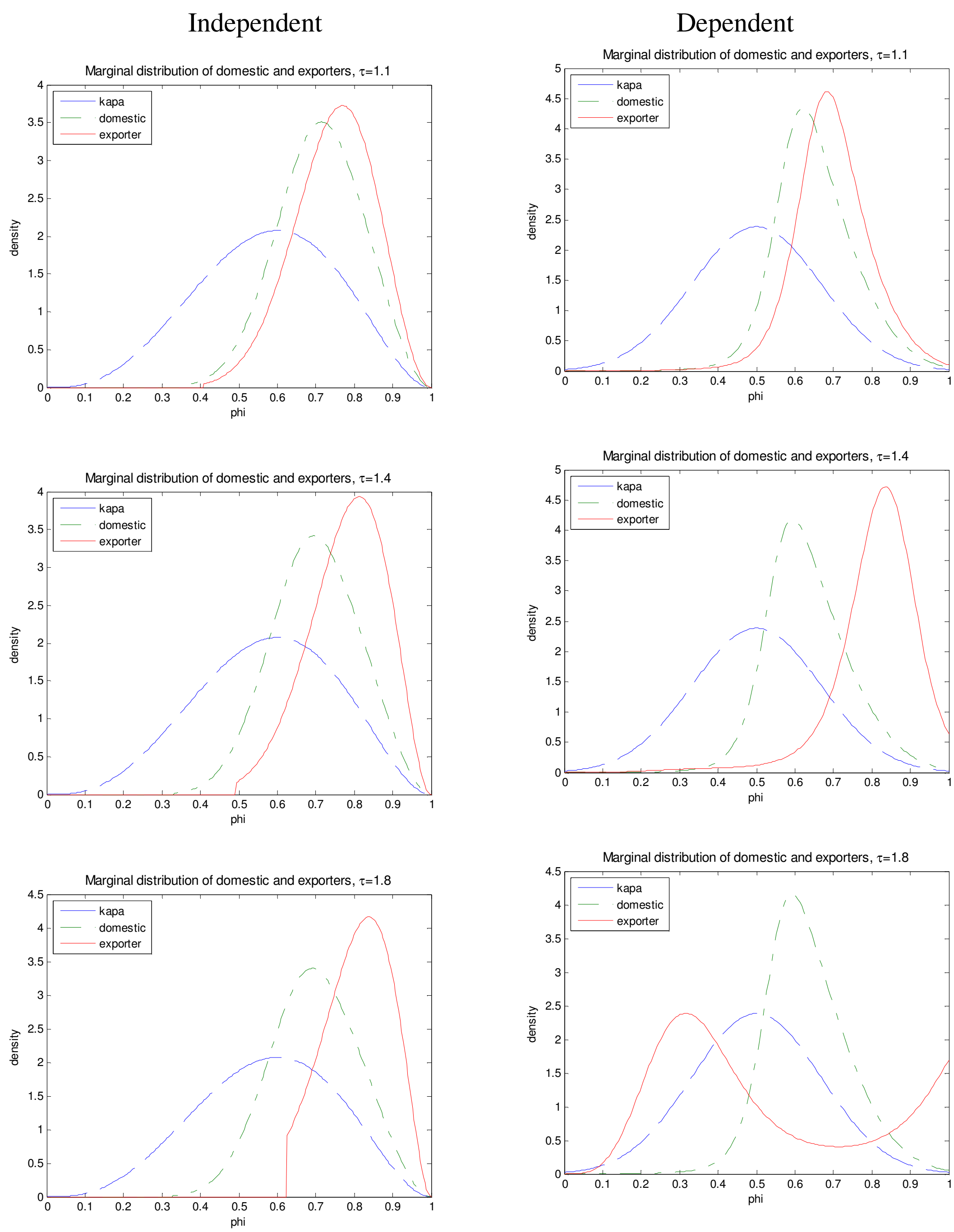


Figure 3.7 Ex ante and ex post marginal productivity distributions



More interestingly, we see that for relatively high transportation cost where the ex ante marginal productivity and fixed cost is jointly correlated (top panel, right), the ex post marginal productivity has a bimodal (two distinct peaks/two local maxima) distribution. The distribution suggests that exporter selection goes against firms of average productivity and preference towards relatively low and high productivity firms. The distinct difference in equilibrium prediction of the ex post distribution implies that future research should take the assumption regarding the underlying distribution of the marginal productivity and fixed cost more seriously.

3.6 Conclusion

The model in this paper reconciles the discrepancy in productivity distribution depicted between the theoretical Melitz type model and the empirical findings. By introducing heterogeneous fixed cost into the Melitz (2003) model, the cut-off productivity is no longer unique but depends on the level of fixed cost. The resulting equilibrium from this adapted model retains the idea quality of the original model. It shows that trade liberalization will always induce more stringent selection and provide welfare improvement. In addition, the model more accurately reflects the empirical productivity distribution of viable firms and the overlapping productivity distribution of the domestic and exporting firms. The resource (market share) is not reallocated from less productive firms to more *productive* firms but from less to more *efficient* ones. Here, the most efficient firms are those who make cheap product (by their high marginal productivity) with a fixed cost investment lower than their peers.

Introducing heterogeneous fixed cost not only made the model more closely fitting to the empirical distribution, but it also unravel how the level of exporting fixed cost relative to the domestic fixed cost may affect the equilibrium weighted productivity. We see that only when the fixed exporting cost is *no less* than proportionate to the domestic fixed cost will the ultimate free trade scenario ensures that the weighted productivity in trade be greater than the weighted productivity in autarky. Despite resource reallocating towards imports discounted the benefit that could otherwise be given to the consumers (for the utility is melted during the transportation), we observe welfare gain.

3.7. References

Bernard, A.B. and B.J. Jensen (1999), "Exceptional exporter performance: cause, effect, or both?", *Journal of International Economics*, **47**(1): 1–25.

Bernard, A., J. Eaton, J.B. Jensen and S. Kortum (2003), "Plants and productivity in international trade", *American Economic Review*, **93**: 1268-1290.

Bernard, A. and J.B. Jensen (2004), "Why some firms export", *Review of Economics and Statistics*, **86**: 561–9.

Bernard, A. and J. Jensen and P. Schott (2006), "Trade costs, firms and productivity**",** *Journal of Monetary Economics*, 53(5): 917–937.

Bugamelli, M. and L. Infante (2003), “Sunk costs to exports”, Bank of Italy Discussion papers, No. 469.

Das, S., M.J. Roberts and J.R. Tybout (2007), “Market entry costs, producer heterogeneity and export dynamics” *Econometrica*, **75**(3): 837-873.

Dedrick, J., K.L. Kraemer and G. Linden (2010), “Who profits from innovation in global value chains? A study of the iPod and notebook PCs”, *Industrial and Corporate Change* 19(1), 81-116.

Greenaway, D. and R. Kneller (2007), “Firm heterogeneity, exporting and foreign direct investment”, *The Economic Journal*, **117** :134–161

Helpman, E., M.J. Melitz and S.R. Yeaple (2004) ″Export versus FDI with Heterogeneous Firms”, *The American Economic Review*, **94**(1): 300-316.

Hess, M. and N.M. Coe (2006), “Making connections: global production networks, standards, and embeddedness in the mobile telecommunications industries”, *Environment and Planning*, 38(7): 1205-1227.

Krugman, Paul R. (1980), “Scale Economies, Product Differentiation, and the Pattern of Trade”, *American Economic Review*, **70**, 950-959.

Jørgensen, J.G. and P.J.H Schröder (2006), “Tariffs and firm-level heterogeneous fixed export costs”, *Contributions to Economic Analysis and Policy*, 5(1): 1–15.

Jørgensen, J.G. and P.J.H. Schröder (2008), “Fixed export cost heterogeneity, trade and welfare”, *European Economic Review*, **52**: 1256-1274.

Leonidou, L.C. (1995), “Empirical research on export barriers: Review, assessment, and synthesis”, *Journal of International Marketing*, **3**(1): 29-43.

Leonidou, L.C. (2000), “Barriers to export management: An organizational and internationalization analysis”, *Journal of International Management*, **6**(2): 1-28.

Leonidou, L.C. (2004), “An analysis of the barriers hindering small business export development”, *Journal of Small Business Management*, **24**(3): 279-302.

Linden, G., K.L. Kraemer and J. Dedrick (2009), “Who captures value in a global innovation network? The case of Apple’s iPod”, *Communications of the ACM*, **52**(3): 140-144.

Linden, G., J. Dedrick and K. L. Kraemer (2011) “Innovation and job creation in a global economy: The case of Apple's iPod”, *Journal of International Commerce and Economics*, 3(1): 223-239.

Melitz, M. J. (2003), “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”, *Econometrica,* Econometric Society, **71**(6): 1695-1725.

Mayer, T and G. Ottaviano (2007), “The happy few: new facts on the internationalisation of European firms,” Bruegel Blueprint 3.

Roberts, M.J. and J.R. Tybout (1997), “The decision to export in Colombia: an empirical model of entry with sunk costs”, *American Economic Review*, **87**: 545-564.

Schmitt, N. and Z. Yu (2001), “Economics of scale and the volume of intra-industry trade”, *Economics Letters*, **74**: 127–132.

Yeaple, S.R. (2009), “Firm heterogeneity and the structure of U.S. multinational activity”, *Journal of International Economics,* **78**(2): 206-215.

3.8 Appendix: technical notes

***Technical Note* 1 Derivation of demand function**

To maximize equation 1 subject to the budget constraint given in equation 2, we define the Lagrangian Γ, using the multiplier λ:

$$\Gamma = \left[\int_{\omega\in\Omega} q(\omega)^{\rho}\, d\omega \right]^{(1/\rho)} + \lambda \left[R - \int_{\omega\in\Omega} p(\omega) q(\omega)\, d\omega \right]$$

Differentiating Γ with respect to $q(\omega)$ and equating to 0 gives the first order conditions. Taking the ratio of the first order conditions for two varieties ω and ω', cancels the multiplier λ. Define $\varepsilon \equiv 1/(1-\rho)$ as discussed in the main text. Then:

$$\frac{q(\omega)^{\rho-1}}{q(\omega')^{\rho-1}} = \frac{p(\omega)}{p(\omega')}, \text{ or } q(\omega) = p(\omega)^{-\varepsilon} p(\omega')^{\varepsilon} q(\omega')$$

It is evident that the elasticity of substitution is constant $\varepsilon = \dfrac{-d \ln q(\omega)/q(\omega')}{d \ln p(\omega)/p(\omega')}$, hence a CES demand function. Substituting the utility maximizing quantity back to the budget constrain (R) and aggregated quantity (Q) function, while noting that $-\varepsilon\rho = 1-\varepsilon$ and $1/\rho = -\varepsilon/(1-\varepsilon)$, the definition of the price index *P* is then derived by dividing the aggregated expenditure by the aggregated quantity consumed:

$$R = \int_{\omega\in\Omega} p(\omega) q(\omega) d\omega = \int_{\omega\in\Omega} p(\omega)^{1-\varepsilon} p(\omega')^{\varepsilon} q(\omega') d\omega = p(\omega')^{\varepsilon} q(\omega') \int_{\omega\in\Omega} p(\omega)^{1-\varepsilon} d\omega$$

$$Q = \left[\int_{\omega\in\Omega} q(\omega)^{\rho}\, d\omega \right]^{1/\rho} = p(\omega')^{\varepsilon} q(\omega') \left[\int_{\omega\in\Omega} p(\omega)^{-\varepsilon\rho} d\omega \right]^{1/\rho}$$

$$P = \frac{R}{Q} = \left[\int_{\Omega} p(\omega)^{1-\varepsilon}\, d\omega \right]^{1/(1-\varepsilon)}$$

The optimal consumption and expenditure decision for individual varieties can thus be expressed by the aggregated term as in equation 4 and 5.

The aggregate price index *P* is a perfect reflection of consumer welfare: if income *R* increases by more than the price index *P*, welfare rises, and *vice versa* if this is not the case (see Diewert (1976) on the exact indices).

***Technical Note 2* Deriving the optimal pricing rule**

Substituting the demand $q(\omega)$ for good ω from equation 4 and the production function 6 in the profit function (7) gives:

$$\pi(\omega) = p(\omega)Q\left(p(\omega)/P\right)^{-\varepsilon} - w\left[f(\omega) + Q\left(p(\omega)/P\right)^{-\varepsilon}/\varphi(\omega)\right]$$
$$= p(\omega)^{1-\varepsilon}QP^{\varepsilon} - wf(\omega) - wp(\omega)^{-\varepsilon}QP^{\varepsilon}/\varphi(\omega)$$

Differentiating with respect to price $p(\omega)$ and equating to zero gives:

$$\partial\pi(\omega)/\partial p(\omega) = (1-\varepsilon)\,p(\omega)^{-\varepsilon}QP^{\varepsilon} + \varepsilon wp(\omega)^{-\varepsilon-1}QP^{\varepsilon}/\varphi(\omega) = 0$$

Cancelling the term $p^{-\varepsilon}QP^{\varepsilon}$ and solving the above equation gives the optimal pricing rule: $p(\omega) = [\varepsilon/(\varepsilon-1)]w/\varphi(\omega) = w/\rho\varphi(\omega)$.

Technical Note 3 Deriving the ex post weighted average productivity

The joint probability density function of the productivity and fixed cost is $\kappa(\varphi, f)$. To calculate the weighted average over the domain in which profit is non-negative, we multiply $\varphi^{\varepsilon-1}$ by its *ex ante* distribution and integrate over all viable firms in set **A**.

$$\tilde{\varphi}(\varphi, f) \equiv \left[E(\varphi^{\varepsilon-1}\middle|\pi \geq 0)\right]^{1/(\varepsilon-1)} = \left[\frac{1}{\int\int_A \kappa(\varphi, f)\,d\varphi df}\int\int_A \varphi^{\varepsilon-1}\kappa(\varphi, f)\,d\varphi df\right]^{1/\varepsilon-1}$$
$$= \left[\frac{1}{p_{in}}\int\int_A \varphi^{\varepsilon-1}\kappa(\varphi, f)\,d\varphi df\right]^{1/\varepsilon-1} = \left[\int_0^\infty\int_0^\infty \varphi^{\varepsilon-1}\mu(\varphi, f)\,d\varphi df\right]^{1/(\varepsilon-1)} = \left[\int_0^\infty \varphi^{\varepsilon-1}\mu_\varphi(\varphi)\,d\varphi\right]^{1/(\varepsilon-1)}$$

Technical Note 4 Deriving the aggregate price

With a continuum of active firms M in the market and a marginal distribution $\mu_\varphi(\varphi)$ of marginal productivity level, the price index *P* is then given by:

$$P = \left[\int_\Omega p(\omega)^{1-\varepsilon}\,d\omega\right]^{\frac{1}{(1-\varepsilon)}} = \left[\int_0^\infty p(\varphi)^{1-\varepsilon}M\mu_\varphi(\varphi)\,d\varphi\right]^{\frac{1}{(1-\varepsilon)}} = M^{\frac{1}{(1-\varepsilon)}}\left[\int_0^\infty (\rho\varphi)^{\varepsilon-1}\mu_\varphi(\varphi)\,d\varphi\right]^{\frac{1}{(1-\varepsilon)}}$$
$$= \frac{M^{1/(1-\varepsilon)}}{\rho\tilde{\varphi}} = M^{1/(1-\varepsilon)}p(\tilde{\varphi}); \quad where \quad \tilde{\varphi} \equiv \left[\int_0^\infty \varphi^{\varepsilon-1}\mu_\varphi(\varphi)\,d\varphi\right]^{\frac{1}{(\varepsilon-1)}}$$

Define $\tilde{\varphi}$ as the weighted average of the marginal productivity levels. All firms with the same marginal productivity sell the same quantity, such that:

$$Q=\left[\int_{\Omega} q(\omega)^{\rho}\, d\omega\right]^{1/\rho}=\left[\int_{0}^{\infty} q(\varphi)^{\rho} M \mu_{\varphi}(\varphi)\, d\varphi\right]^{1/\rho}=M^{\frac{1}{\rho}}\left[\int_{0}^{\infty} q(\tilde{\varphi})^{\rho}\left(\frac{\varphi}{\tilde{\varphi}}\right)^{\rho\varepsilon} \mu_{\varphi}(\varphi)\, d\varphi\right]^{1/\rho}$$
$$=M^{\frac{1}{\rho}} q(\tilde{\varphi})\left[\int_{0}^{\infty} \frac{\varphi^{\varepsilon-1}}{\tilde{\varphi}} \mu_{\varphi}(\varphi)\, d\varphi\right]^{\frac{\varepsilon}{\varepsilon-1}}=M^{\frac{1}{\rho}} \frac{q(\tilde{\varphi})}{\tilde{\varphi}^{\varepsilon}}\left[\int_{0}^{\infty} \varphi^{\varepsilon-1} \mu_{\varphi}(\varphi)\, d\varphi\right]^{\frac{\varepsilon}{\varepsilon-1}}=M^{\frac{1}{\rho}} \frac{q(\tilde{\varphi})}{\tilde{\varphi}^{\varepsilon}} \tilde{\varphi}^{\varepsilon}=M^{\frac{1}{\rho}} q(\tilde{\varphi})$$

Similarly, firms with the same marginal productivity have the same revenue:

$$R=\int_{\Omega} p(\omega) q(\omega)\, d\omega=\int_{\Omega} r(\omega)\, d\omega=\int_{0}^{\infty} r(\varphi) M \mu_{\varphi}(\varphi)\, d\varphi=M \int_{0}^{\infty} r(\tilde{\varphi})\left(\frac{\varphi}{\tilde{\varphi}}\right)^{\varepsilon-1} \mu_{\varphi}(\varphi)\, d\varphi$$
$$=M \frac{r(\tilde{\varphi})}{\tilde{\varphi}^{\varepsilon-1}} \int_{0}^{\infty} \varphi^{\varepsilon-1} \mu_{\varphi}(\varphi)\, d\varphi=M\, r(\tilde{\varphi})$$

Technical Note 5 Difference between the average revenue per firm and average revenue per product

Weighted average revenue from domestic sales:

$$r_d(\tilde{\varphi}_{tr})=R(P\rho\tilde{\varphi}_{tr})^{\varepsilon-1}$$
$$=R(P\rho)^{\varepsilon-1}\left\{\left(\frac{1}{1+p_x}\tilde{\varphi}_d^{\ \varepsilon-1}+\frac{p_x}{1+p_x}(\tilde{\varphi}_x/\tau)^{\varepsilon-1}\right)^{1/(\varepsilon-1)}\right\}^{\varepsilon-1}$$
$$=R(P\rho)^{\varepsilon-1}\left(\frac{1}{1+p_x}\tilde{\varphi}_d^{\ \varepsilon-1}+\frac{p_x}{1+p_x}(\tilde{\varphi}_x/\tau)^{\varepsilon-1}\right)$$

Weighted average revenue earn from domestic and foreign market:

$$\overline{r}_{tr}=r_d(\tilde{\varphi}_d)+p_x r_x(\tilde{\varphi}_x)$$
$$=R(P\rho\tilde{\varphi}_d)^{\varepsilon-1}+p_x R(P\rho\tilde{\varphi}_x/\tau)^{\varepsilon-1}$$
$$=R(P\rho)^{\varepsilon-1}\left[\tilde{\varphi}_d^{\ \varepsilon-1}+p_x(\tilde{\varphi}_x/\tau)^{\varepsilon-1}\right]$$

Therefore, to account for the numbers of firms

$$M=M_{tr}\frac{1}{1+p_x}=\frac{R}{r_d(\tilde{\varphi}_{tr})}\frac{1}{1+p_x}$$

Chapter 4 Investment, supermarket and fixed cost heterogeneity, evidence from chili farmers in Indonesia

4.1 Introduction

The increase in market penetration of supermarkets continues to transform the food-retail sector in developing countries (see Reardon et al, 2009; Reardon et al, 2007; Neven et al., 2006; Reardon et al., 2005; 2003; Weatherspoon and Reardon, 2003). In much of Asia, more than a decade of rising incomes and urbanization has resulted in large shifts in food consumption patterns and shopping behavior. Consumers are eating many more calories on average, while shifting from staple foods consumption towards higher-value fruits, vegetables and processed foods in consumption. Shoppers are willing and able to pay price premiums for convenience, food quality, and food safety (Mergenthaler et al, 2009a,b).[22]

These ongoing trends also affect agricultural producers through foreign investments and the higher standards imposed by supermarkets (Schipmann and Qaim, 2011a,b).[23] A growing literature aims to understand how the transformation of modern food retail chains influences smallholder market decisions. In its broadest sense, research on evolving food chains addresses how market liberalization, trade and foreign investment policies impact food producers, traders, wholesale markets, processors and retailers. This literature analyzes how rapid income growth and urbanization transform traditional chains, exploring poverty reduction opportunities, distributional implications and efficiency effects of smallholder participation (Reardon, Barrett, Berdegué and Swinnen, 2009). Key research issues are whether small producers are able to participate in more modern food chains and when they do participate, what are the growth, equity and development consequences?

To date the evidence suggests that supermarkets provide challenges, threats and opportunities to small farmers (Reardon et al, 2009). On the one hand, better prices and a steady relationship with the buyer potentially improves the welfare for those who participate in the

[22] See also Minot and Roy, 2006; Traill, 2006; Regmi and Dyck, 2001; Jaffee and Morton, 1994; Minot, 1986.
[23] See also Reardon, Berdegué et al., 2007; Berdegué et al., 2007; Reardon, Henson et al., 2007; Berdegué et al., 2005; Henson and Reardon, 2005; Reardon and Swinnen, 2004; Neven and Reardon, 2004; Reardon et al, 2003; Reardon et al., 2001; Reardon and Barrett, 2000.

supermarket retail channel.[24] On the other hand, the quantity and quality requirements set by supermarkets are a challenge for small farmers. A growing body of literature has been looking into the determinants of the modern channel participation and the welfare implications of the potential exclusion of small-scale producers from entering the modern retail channel (Rao and Qaim, 2010; Miyata et al., 2009; Moustier et al., 2009; Neven et al., 2009).[25] Exclusion from potentially more profitable food markets implies small farmers may be stuck in lower income poverty traps, and unable to gain the new technologies and management practices that modern food markets require.

This paper contributes to this literature in three ways. First, we develop a simple model that determines the sorting of famers, based on their household characteristics, into high- or low-cost investment for production, rather than channel selection. The investment level correlates positively with the probability of producing high-quality output meeting the quality standards of the supermarket, thus increasing the likelihood of accessing the modern retail channel. The model thus provides a theoretical foundation for the self-selection of investment, which influences a farmer's modern channel participation, and the economic impact of that decision. The model provides a testable hypothesis regarding differences in profitability.

Second, we incorporate an aspect overlooked in the current literature, namely that not all modern channel producers are aware of the final marketing channel of their product. This unawareness may arise because farmers do not sell directly to retailers but to intermediate traders, who represent and collect the fresh products for either the traditional market wholesaler or the dedicated supermarket wholesalers (Gulati et al., 2007). Asymmetric information regarding the value and quality of the product and its marketing channel places the intermediate trader in a better bargaining position over farmers with less market information. Some farmers are therefore unable to receive the maximum price the traders are potentially willing to offer. We model this aspect as a factor that affects traders' search cost. Empirically, we show how awareness in the marketing channel is an important aspect influencing farmers' profitability.

---

[24] Sexton (2012) argues that the agricultural market is characterized by imperfect competition in view of product heterogeneity, consolidation, and the dominate role of a few large processing, trading, and retailing firms.
[25] See also Hernández et al., 2007; Reardon, et al., 2007; Reardon, et al., 2007; Natawidjaja et al., 2007.

Third, based on our model we apply the treatment effect model (Heckman self-selection) to analyze the heterogeneous welfare implications for three types of farmers: aware, unaware, and traditional. The treatment effect is implemented in a two-stage estimation procedure, with the first stage dealing with the self-selection issue (the probability of making a high-cost investment decision) and the second stage dealing with the welfare implications.

The empirical analysis uses household data from a survey of 597 chili farmers in Java, Indonesia.[26] Chili producers are chosen because chilies are a high-value agricultural product in Indonesia, providing farm households a source of income in addition to traditional or staple crops. The chili is an important vegetable consumed daily by most Indonesian families and supermarkets are increasingly including chilies as part of their product offerings. Like many other fresh fruit and vegetables (FFV) production, the production of chili is labor intensive, providing additional employment opportunities (Weinberger and Lunpkin, 2005). Previous studies suggest that FFV sold in the supermarket are usually marketed as high quality produce and that traders selling directly to the supermarkets usually receive a premium for their produce. However, production of high-quality FFV products that meet the stringent specifications of the supermarket channel usually comes at a higher input cost, resulting from the use of high-quality seeds, hired labor, capital investments required to ensure product quality standards (in terms of safety and appearance), and the packaging, storage and delivery requirements (Schipmann and Qaim, 2011b; Neven et al., 2009). These costs are challenging for the credit-constrained farmers. In addition, production uncertainty implies that even though investments are made and additional costs incurred, not all output will be of sufficient quality to meet the modern retail standards. Thus, investing in high-quality output is not necessarily rewarded with higher prices.

We find that aware and unaware farmers have similar basic household characteristics and production attributes. Our production sorting and welfare optimization model suggests that farmers who self-select into making high-cost investments for production earn higher profits. The treatment effect model confirms that endogenous sorting takes place based on farmers' productivity and other household characteristics. More interestingly, among modern channel producers, aware farmers benefit much more from participating in the modern channel than

[26] See Chowdhury el al. (2005) for an overview of structural change in the agricultural sector in Indonesia. See Awfaw et al (2010), Awfaw (2007), Minten et al (2009), and Miyata et al (2009) for treatment effect models.

unaware farmers, after taking into account their investment level. This suggests that farmers benefit unequally even among supermarket suppliers. Our research thus shows that (i) farmers have to overcome a production and investment barrier in order to produce high-quality products suitable for the modern food retail channel and (ii) farmers need to be aware of the marketing channel of their product to obtain the full benefit from their investments as suppliers of the modern food retail channel.

4.2 Production sorting model

Let F be the total number of farmers, each producing one unit of output. Output is heterogeneous in terms of quality and can be sold in two separate channels: the modern channel and the traditional channel. The product quality is evaluated partially according to the appearance of a product, while other unobservable characteristics are evaluated through close monitoring during the production stage. An exogenous quality threshold set by the supermarket leads to sorting of outputs into two levels: high-quality and low-quality. Consumers willing and able to pay a premium for guaranteed quality and food safety shop in the supermarket.

There are different prices offered to the famers in the two channels. Traders who procure for the modern channel buy only high-quality products. Although quality is partially observable, traders have to search for the farmers that produce high-quality output and incur search cost, S_i. These search cost are farmer-specific and depend on the ease with which the supermarket traders can find the farmers with high-quality products. The easier it is for the traders to find or contact the farmers with the output they require the lower the search cost for the trader to reach a specific farmer. The maximum price (or reservation price) the modern channel traders are willing to pay for the high-quality output is the difference between the price supermarkets pay to the trader P_S and the search costs supermarket traders incur to find a specific farmer: $P_{M,i} = P_S - S_i$, where the subscript M indicates Modern, S indicates Supermarket, and i is a farmer index. Traders who procure for the traditional channel do not care about quality and buy whatever output is available in the market. The search cost for the traditional traders is therefore negligible. The traditional channel traders offer a maximum price P_T.

Farmers are heterogeneous in terms of production efficiency γ. The efficiency is related to household characteristics, such as age, education, experience, and endowments. More educated farmers, for example, are expected to have better understanding of the specific requirements and details specified in the contract or agreement with the traders compared to the less educated farmers. Since the maximum price the supermarket pays to the traders is fixed, the price these traders are willing to offer to each farmer is a function of the search cost that differs between farmers: $P_M(S_i)$. Let $g(\gamma, S)$ denote the joint probability distribution of farmers' production efficiency and the search cost incurred to reach them. Farmers choose between making a high- or low-cost production decision, with investment costs C_H and C_L respectively ($C_H > C_L$). Production involves uncertainty. Farmers with high-cost investment produce high-quality output with probability β_H and low-quality output with probability 1-β_H. Similarly, farmers with low-cost investment produce high-quality output with probability β_L and low-quality output with probability 1-β_L. We assume that high-cost investment leads to higher expected quality: $\beta_H > \beta_L$.

Let $\delta \in (0,1]$ denote the probability that high-quality output is sold to the supermarket trader. The rejection rate $1-\delta$ is exogenous and independent of individual farmers' search cost.[27] Based on the exogenous market condition parameters $(\delta, \beta_H, \beta_L, P_S, P_T, C_H, C_L)$, farmers have three options: investing in high-cost production, low-cost production, or exiting the market (not entering into production). The expected profit π for farmer i with production efficiency parameter γ_i for whom the trader incurs search costs S_i is :

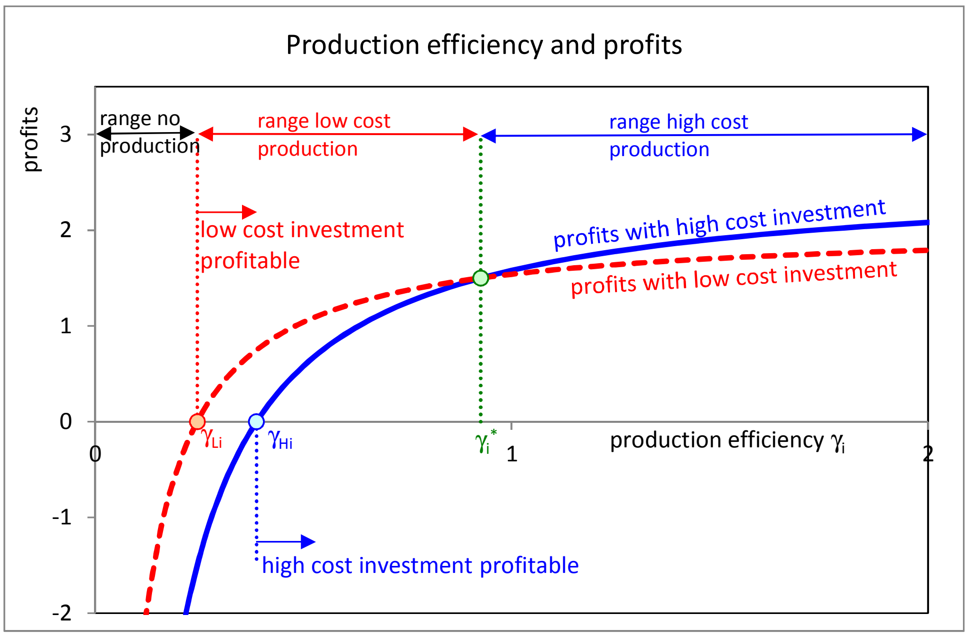
$$(1) \qquad E(\pi_i) = \begin{cases} \delta\beta_H(P_S - S_i) + (1-\delta\beta_H)P_T - C_H/\gamma_i & \textit{high cost investment} \\ \delta\beta_L(P_S - S_i) + (1-\delta\beta_L)P_T - C_L/\gamma_i \quad \textit{for} & \textit{low cost investment} \\ 0 & \textit{no production} \end{cases}$$

Figure 4.1 illustrates the production choice facing a farmer as a function of his production efficiency, given a particular set of other parameters. If production efficiency is below some critical value γ_{Li} even low cost production is not profitable. Similarly, high cost production

27 The parameter captures the exogenous product rejection shock for farmers.

with efficiency below some critical value γ_{Hi} is not profitable.28 Note that these critical values are producer-specific as they depend on the search costs S_i. Figure 4.1 is drawn under the assumption that $\gamma_{Li} < \gamma_{Hi}$, in which case we can identify three ranges of production for the producer in question. If production efficiency is too low, namely $\gamma_i \in (0, \gamma_{Li})$, the farmer does not produce. If production efficiency is intermediate, namely $\gamma_i \in [\gamma_{Li}, \gamma_i^*)$, the farmer produces with low cost investment. If production efficiency is high, namely $\gamma_i \in [\gamma_i^*, \infty)$, the farmer produces with high cost investment.

Figure 4.1 Expected profits from investing in high and low cost production



Parameters: $P_S = 4$; $S_i = 1$; $P_T = 1.5$; $\delta = 0.9$; $\beta_H = 0.8$; $\beta_L = 0.4$; $C_H = 1$; $C_L = 0.5$.

The critical value γ_i^* which determines the borderline between low cost and high cost investment is determined by the switching condition, leading to equal profits for low and high cost investment.

(2) $$\gamma_i^* = \frac{(C_H - C_L)}{\delta(\beta_H - \beta_L)(P_S - S_i - P_T)}$$ (switching condition)

Clearly, for this switching condition to be operative two conditions must be fulfilled. First, the solution should be positive and finite. This implies that the term $(P_S - S_i - P_T)$ must be

---

28 Note that $\gamma_{Li} = C_L / [P_T + \delta\beta_L(P_S - S_i - P_T)]$ and $\gamma_{Hi} = C_H / [P_T + \delta\beta_H(P_S - S_i - P_T)]$.

positive, or the search costs must be less than the supermarket price premium: $S_i < (P_S - P_T)$. In other words: farmer with too high search costs will never switch to high cost investments. Second, for a switch to occur it must be worthwhile to produce at low cost for intermediately efficient farmers, as illustrated in Figure 4.1. However, if $\gamma_{Hi} < \gamma_{Li}$ high cost investment profits always dominate low cost investment and a switch does not occur. As can be easily verified, this condition is equivalent to assuming $\gamma_{Hi} > \gamma_i^*$, that is switching occurs to the left of the point where high cost investment becomes profitable. Note that the conditions imposed so far do not exclude this possibility from occuring, as $\gamma_{Hi} < \gamma_{Li}$ is equivalent to assuming $(C_H - C_L)P_T < \delta(\beta_H C_L - \beta_L C_H)(P_S - S_i - P_T)$. Other things equal, it is therefore more likely to hold if (i) the probability of high quality production with high cost investment β_H rises, (ii) the probability of high quality production with low cost investment β_L falls, and (iii) the search costs S_i fall. We can combine the various possibilities by defining the *switching* efficiency level $\tilde{\gamma}_i \equiv max\{\gamma_i^*, \gamma_{Hi}\}$. High cost investment production then occurs if efficiency exceeds the switching level: $\gamma_i \geq \tilde{\gamma}_i$. Similarly, we can define the *minimum* efficiency for positive production level $\underline{\gamma_i} \equiv min\{\gamma_{Li}, \gamma_{Hi}\}$. Low cost investment production then occurs if efficiency is in between the minimum efficiency and the switching level: $\underline{\gamma_i} \leq \gamma_i < \tilde{\gamma}_i$. As noted above, depending on the specific parameters this range may be empty for certain levels of low searching costs.

4.2.1 Equilibrium

The market is in equilibrium when high quality output sold in the supermarket matches the demand for high quality output that pays a high price: $Q^M = D_H$, while the rest of the output is sold though the traditional channel regardless of quality: $Q^T = Q - Q^M$. To determine the share of farmers producing with high and low investment costs and the share exiting the market we need to take both the search costs and the production efficiency into consideration. This is illustrated in Figure 4.2 in case there is no range of low cost investment for sufficiently low search costs. The figure shows γ_H and γ_L as defined in footnote 7and γ^* as defined in equation (2). All three curves intersect at point A. To the left of this point there is no low cost investment area, such that: $\underline{\gamma} = \gamma_H = \tilde{\gamma}$. To the right of point A we have: $\underline{\gamma} = \gamma_L < \tilde{\gamma}$. The range to the left of point A does not occur, of course, if point A is to the left of the vertical axis. A sufficient condition for that to happen is: $\beta_L C_H > \beta_H C_L$, that is the appropriately weighted probability of high quality output with low cost investment is sufficiently high.

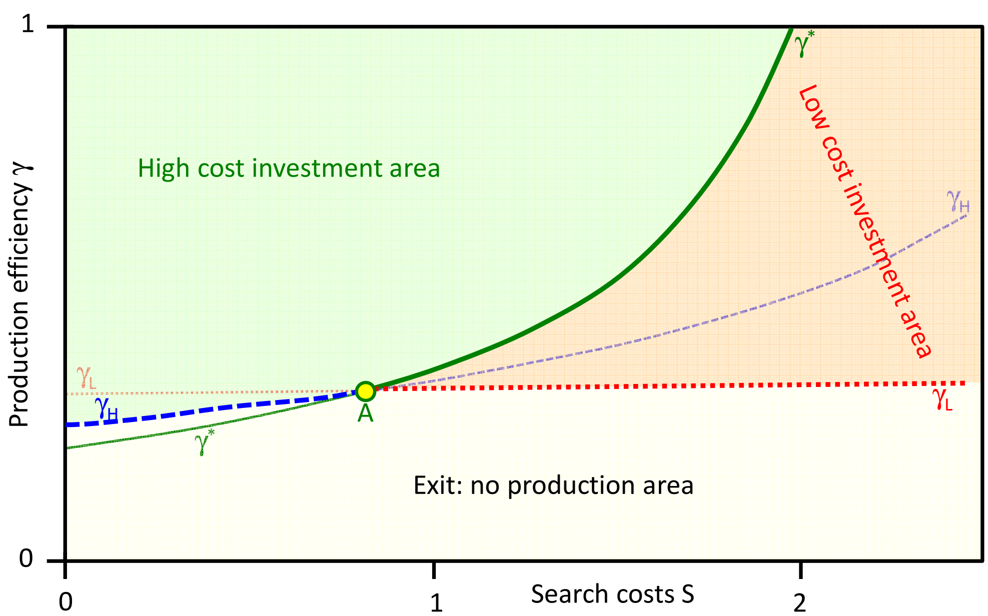
The curves in Figure 4.2 separate the plain into three regions and depict the sorting of farmers into high cost production, low cost production, and exit depending on their production efficiency and the search cost. Taking the joint density $g(\gamma, S)$ over production efficiency and search costs into consideration, we can now determine the share μ_H of farmers which produce with high costs investment, the share μ_L which produces with low cost investment, and the share $\mu_E = 1 - \mu_H - \mu_L$ which exits the market:

(3) $\mu_H = \int_0^\infty \int_{\tilde{\gamma}}^\infty g(\gamma, S) d\gamma dS$ high cost investment

(4) $\mu_L = \int_0^\infty \int_{\underline{\gamma}}^{\tilde{\gamma}} g(\gamma, S) d\gamma dS$ low cost investment

(5) $\mu_E = \int_0^\infty \int_0^{\underline{\gamma}} g(\gamma, S) d\gamma dS$ exit

Figure 4.2 Sorting of farmers into high and low investment costs and exit



Parameters: $P_S = 4$; $P_T = 1.5$; $\delta = 0.99$; $\beta_H = 0.99$; $\beta_L = 0.04$; $C_H = 1$; $C_L = 0.5$.

Since there is a total of F farmers available, and the probability of producing high quality output is equal to β_H for farmers with high cost investment and equal to β_L for farmers with low cost investment we can now determine the total volume of high quality and low quality production.

(6) $Q_H = (\mu_H \beta_H + \mu_L \beta_L) F$ high quality output

(7) $Q_L = \big(\mu_H (1 - \beta_H) + \mu_L (1 - \beta_L)\big) F$ low quality output

The total output produced is the sum of high and low quality outputs: $Q = Q_H + Q_L$ and the total number of farmers that exit production is $F - Q$. Recall that only a fraction δ of high quality output will be sold through the modern supermarket channel: $Q^M = \delta Q_H$. The remaining output is sold through the traditional channel regardless of quality. As long as some high quality product is rejected by the supermarket traders ($\delta \neq 1$) the total high quality produced is greater than the quantity sold through the supermarket channel. Therefore, although only Q^M is priced and purchased as high quality product, the true amount of high quality food consumed is larger. This is an indirect welfare effect for the public often neglected. There will be a welfare loss as there will be less than the optimal number of farmers joining high quality output production when a proportion of high quality output is not rewarded with a high price. In this case there will be fewer farmers engaging in high cost investment with high quality production. At the same time, the rejected high quality product will be sold at a lower price to the consumers who shop at the traditional market either because they cannot afford standardized quality food or because they place a lower value on food quality.

4.2.2 Comparative static analysis

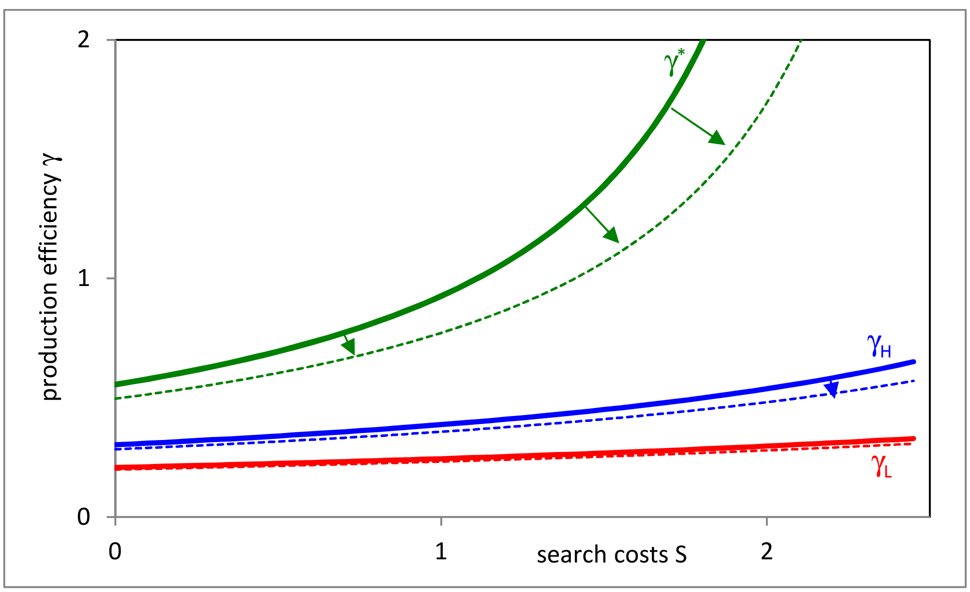
During the stage of economic transition in which the growing demand for high quality food exceeds the supply in the short-run, the excess demand drives up the procurement price in the modern channel. Recall that the maximum price the supermarket trader is willing to pay is positively related with the price P_S paid by supermarkets and negatively related to traders' search cost S_i, which differs per farmer. As economies develop and the price P_S supermarkets offer to traders rises, both the switching production efficiency and the minimum production efficiency fall.[29] This is illustrated in Figure 4.3. Since the high investment cost area increases and the exit area decreases, we can be sure that the share of high cost investment farmers rises ($\partial \mu_H / \partial P_S > 0$) and the share of exiting farmers falls ($\partial \mu_E / \partial P_S < 0$). Whether

[29] Note, for example, that $\partial \gamma_{Li} / \partial P_S = -\delta \beta_L \gamma_{Li}^2 / C_L$. Similarly for the other curves.

or not the share of low cost investment farmers rises or falls depends on the joint density $g(\gamma, S)$.

Up to this point, we reflect little on supermarket traders' search cost and assume it is exogenously determined by the remoteness of the farmer. As mentioned, the easier it is for the traders to find or contact the farmers with the output they want, the lower the search cost for the trader to reach a specific farmer. Search costs are negatively correlated with the ease of communication traders experience with farmers and with the quality signals provided by the farmers. Although the farmers located further away from the market in principle have higher search costs, ownership of transportation or communication facilities, such as trucks and mobile phone can mediate the impact of physical distance. Other farm-related assets that signal producer's ability to preserve freshness or delay the product from perishing also decrease traders' search cost and are expected to increase profit. Note in particular that for an individual farmer it is the net premium paid by the supermarket channel over the traditional channel which determines its investment decision. The impact of a fall in search costs is therefore the same as the impact of a rise in the price paid by the supermarket.

Figure 4.3 Impact of supermarket price increase; from solid to dashed lines



Parameters: $P_S = 4$ *and* 4.3 (*solid and dashed*); $P_T = 1.5$; $\delta = 0.9$; $\beta_H = 0.8$; $\beta_L = 0.4$; $C_H = 1$; $C_L = 0.5$.

4.2.3 Hypothesis

To sum up the model and connect the conceptual model with the empirical work, following hypotheses are made. First, profit maximization ensures farmers sort into different investment levels based on their productivity. Thus, the most productive farmers self-select into high effort investment, and less productive farmers invest in low cost production, while the least productive farmers will not enter production.

- Hypothesis 1: More productive farmers self-select into high cost production and earn higher profits than farmers who invest in low cost production.

Second, although farmers make production decisions based on market price for high and low quality output, the search costs traders incur to find each farmer depends on how easy it is to find them. Thus, we expect that search cost will affect the maximum price the trader is willing to offer to each farmer.

- Hypothesis 2: The expected profits for farmers with lower search costs for traders are higher as these farmers are more likely to engage in high cost investment.

Third, awareness of the marketing channel explains the heterogeneity in profitability among farmers whose product enters the modern retail channel. On the one hand, we expect that aware farmers stand at a relatively better bargaining position as compared to their unaware neighbors, given the knowledge of their products' marketing channel. On the other hand, awareness can be thought of to lower search cost. As aware farmers are more likely to initiate contact or signal their ability to produce high quality output to the modern channel traders themselves, they lower traders search cost to find them.

- Hypothesis 3: Aware farmers earn higher profit than unaware farmers.

4.3 Empirical Methodology

Farm household income is determined by the income earned from agricultural production activities and by various socioeconomic factors. Since higher production capacity and productivity can lead to more investment and thus result in better output quality, we assume that the investment level directly affects the quality of the output produced, raising the probability that the farmer will receive a premium price for the product. High or low cost investment is thus a binary decision farm households make in order to maximize their net return from production (chili profit) and net household income.

Since the investment decision is endogenously determined, such that producers self-select into making high cost investment for production, the coefficient estimated based on OLS may be biased due to omitted variables in the specification. To eliminate this bias we use the treatment effect model, also known as the Heckman self-selection model, to get an unbiased and consistent estimate of the effect of making high cost production investment on profit and income. The method is more common in the impact assessment literature, particularly when the treatment is not randomly applied (Greene, 2008; Maddala, 1986, 1983; Heckman, 1979, 1978). The strategy is to apply a two-stage estimation procedure. While the first stage evaluates the probability of making high cost investment (modern channel participation) with a probit model, the second stage uses the prediction from the first stage to evaluate the impact of the treatment on the welfare outcome. Maximum likelihood estimation is used, with all parameters estimated simultaneously.

The model estimates the effect of an endogenous binary treatment z_i on a continuous and fully observable variable y_i, conditioned on the independent variables x_i and w_i. The primary interest is in the welfare/output model:

(8) $y_i = x_i\psi + z_i\vartheta + v_i$ (2nd stage, output model)

where z_i is the endogenous dummy variable indicating whether the treatment is applied. The coefficient ϑ captures the effect of the endogenous decision of making high cost investment. The binary decision of whether the treatment is applied is modeled as the outcome of an unobserved latent variable z_i^*. It is assumed that z_i^* is a linear function of the exogenous covariates w_i and a random component u_i, that is: $z_i^* = w_i\theta + u_i$. The observed binary investment decision is:

(9) $z_i = \begin{cases} 1 & if\ z_i^* > 0 \\ 0 & otherwise. \end{cases}$ (1st stage selection model)

where u and v are bivariate normal with mean zero and covariance matrix $\begin{bmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{bmatrix}$.

Identification requires that at least one variable appearing in the selection equation is not included in the outcome equation. The identification variable often used in the literature is distance. Rao and Quim (2010), for example, use distance from a farmer's house to a tarmac road, while others used distance to a paved highway (Hernández et al., 2007) and to village head's land (Miyata et al., 2009). Distance from house to asphalt road is used as identification variable in this paper.

4.4 Data description

In this section, we first describe the survey and show how the three groups of producers (aware, unaware and traditional) are identified. The next sections provide summary statistics of the main variables and t-tests of difference in mean value between each group of farmers.

4.4.1 Survey design, awareness, and channel identification

Data for this study is collected by Sahara[30] in 2010 through a household survey of chili producers in West Java, Indonesia. There are in total 597 valid samples. Although chilies are one of the most important food crops consumed daily by Indonesians, the amount purchased through the modern channel is still limited. As a result, the share of chili supplying to the supermarket is small. The survey sampling is stratified based on the marketing channel to insure sufficient representation of producers from both the traditional and modern channel.

Producers who are aware of selling their chili to the supermarket are identified by their response from "experience with the modern channel". Those with a positive response of having sold chili to supermarkets are categorized into the aware farmers group. Identifying unaware modern channel producers is more difficult. This is done by an additional 'channel' identifier in the data, which identifies modern channel producers apart from those traditional channel producers. The channel identifier is available by asking the supermarket for a list of suppliers in which they source their chili product from at the stage of the survey design to insure the modern channel producers are sufficiently represented in the sampling. With the channel identifier and famers response in their experience to the modern channel, unaware modern channel producers can be singled out as they denied selling chili product to the modern channel (have no experience), while they are at the same time identified as supplier by the supermarket. This gives us three separate groups of chili producers: aware (A), unaware (U) and traditional (T) farmers. In our data there are 485 (81%) traditional farmers, 66 (11%) unaware farmers, and 46 (7.7%) aware farmers (Table 4.1).

[30] See Sahara (2012) for more information on survey design, data collection and identification of the modern channel supplier (see also Sahara et al., 2011).

Table 4.1. Numbers of observations in each group			
	Channels		
Farmer type	Traditional	Modern	sum
Traditional (T)	485	0	485
Unaware (U)	0	66	66
Aware (A)	0	46	46
sum	485	112	597

4.4.2 Descriptive analysis— dependent variables: Profit and Net income

Farmers' net income is derived by aggregating income from all of the income-earning activities for the household. The income from agricultural production related activities was calculated by subtracting the revenue earned from the costs associated with production. Aware farmers on average have significantly more net income than unaware farmers by 13 million Rupiah (Rp), approximately USD 1,350 a year. On average, modern channel producers' (aware and unaware combined) net income is significantly greater than traditional channel producers, but the difference between unaware farmers and traditional producers alone is not significantly different (Table 4.2, row 1).

*Table 4.2 Source of net income**

	mean			test statistics (p-value)			
Unit: million Rp	Traditional	Unaware	Aware	AvU	AvT	UvT	Channel
Net income	23.71	26.61	41.49	**0.10**	**0.00**	0.43	**0.01**
Net income from various income activities							
Chili	6.12	8.79	18.74	**0.03**	**0.00**	0.15	**0.00**
other Agri.	7.06	6.45	6.38	0.98	0.83	0.78	0.74
Livestock	0.72	0.47	0.65	0.63	0.88	0.45	0.52
Aquaculture	0.08	0.16	0.38	0.16	**0.00**	0.32	**0.01**
Agri. Trade	1.34	0.00	3.84	0.15	**0.06**	**0.09**	0.75
other trading	2.16	2.82	4.88	0.31	**0.09**	0.60	0.15
Rice / milling	0.28	0.00	0.00	na	0.60	0.53	0.41
Food processing	0.14	0.32	0.20	0.71	0.73	0.29	0.32
Other business	1.71	2.92	1.43	0.49	0.83	0.27	0.49
Agri. Wage labor	0.81	0.59	1.00	0.28	0.63	0.55	0.89
non-agi. Employment	2.48	3.32	2.47	0.66	0.97	0.39	0.50
pension	0.21	0.02	0.91	0.10	**0.07**	0.51	0.47
Remittances	0.38	0.62	0.42	0.57	0.86	0.28	0.37
Assistance program	0.01	0.10	0.05	0.37	**0.08**	**0.00**	**0.00**
Other.	0.20	0.06	0.15	0.44	0.87	0.54	0.57

* Excluding 17 obs. with negative net income, of which 15 traditional farmers and 2 unaware farmers group.

Disaggregating the net income by source, we see that net profit from chili production is on average 18.7 million for the aware farmers, followed by 8.8 million for the unaware and 6.1 million for the traditional farmers (Table 4.2, row 2). Net income from chili production thus accounts for a high percentage of income for most farmers, namely 26, 33, and 45 per cent for traditional, unaware, and aware farmers, respectively. Traditional channel producers have a greater proportion of income derived from other agricultural activities. Net income from other agricultural activities alone contributed to 30, 24 and 15 percent of the total net income for the traditional, unaware and aware farmers, respectively. Other significantly different income sources between the groups of farmers are: agricultural trade, other trading activities, and net income from assistance program.[31]

4.4.3 Descriptive analysis—independent variables

Five broad classes of explanatory factors are used in the models explaining farmers' welfare: household characteristics, land size & ownership attributes, non-land assets ownerships, chili production attribute, and habit of record keeping. Summary statistics of the variables used in the analysis are provided in Table 4.3, with cross group comparison of the mean value reported.

Household characteristics

As already pointed out by Sahara (2012), and similar to other studies, the data show that the modern channel participants are younger and receive more education as compared to the traditional channel producers (Rao and Qaim, 2010; Hernández et al., 2007). Among modern channel producers, household characteristics of unaware famers are not significantly different from aware farmers (Table 4.3, column 4: AvU). The significantly higher education received among supermarket producers suggests that these households have higher human capital than traditional channel producers. Boselie et al. (2003) have shown that farmers with higher human capital are more likely to improve their negotiation skills in dealing with the buyer. Young producers are more willing to adapt to new technologies and production methods. The percentage of family members in different age categories (young / working / retired) gives an indication of the dependency ratio and the percentage of family member that can contribute to the labor force. The difference in age categories is, however, not significant.

[31] We use total net income and net income from chili production (chili profit) in logarithm form as dependent variables for our analysis. This reduces the number of observation due to negative net income values by 17 (2 from unaware and 15 from traditional) and by 120 due to negative profit from chili production (101 from traditional, 14 from unawary, and 5 from aware).

Table 4.3 Summary statistics of variables used, across groups comparison

	mean			Test statistics (p-value)			
	Traditional	Unaware	Aware	AvU	AvT	UvT	Channel
Household characteristics variables							
Household size (number)	4.58	4.23	4.46	0.40	0.61	0.09	0.12
Age (years)	46.29	44.03	42.78	0.50	**0.04**	0.12	**0.02**
Age spouse	40.13	38.88	36.09	0.19	0.01	0.37	**0.03**
Education (years)	6.43	8.02	7.96	0.92	**0.00**	**0.00**	**0.00**
Education spouse	6.63	7.66	8.05	0.49	**0.00**	**0.00**	**0.00**
Age between15-65 (%)	65.95	64.05	66.43	0.52	0.87	0.47	0.66
Age above 65 (%)	2.42	3.66	3.91	0.90	0.31	0.34	0.19
distance to road (km)	0.30	0.11	0.12	0.70	**0.07**	**0.02**	**0.00**
--- to sub-district market (km)	6.05	3.95	7.80	**0.00**	**0.03**	**0.00**	0.34
Total land size owned and cultivating land ownership and types							
Land owned							
Year 2005	4,079.00	4,876.41	3,195.34	0.40	0.35	0.40	0.89
Year 2010	4,544.28	5,145.24	3,432.92	0.40	0.29	0.56	0.90
Land cultivated							
Total (m^2)	6,433.41	8,089.61	8,097.52	1.00	0.18	0.11	**0.05**
% owned	51.65	44.72	36.22	0.23	**0.01**	0.20	**0.01**
% rented	24.35	32.39	50.75	**0.01**	**0.00**	**0.10**	**0.00**
% others	24.00	22.89	13.03	**0.10**	**0.05**	0.81	0.16
% irrigated	46.28	45.32	65.04	**0.01**	**0.01**	0.87	0.13
% Rain fed	11.51	7.14	5.90	0.75	0.21	0.24	**0.09**
% Dry land	42.21	47.54	29.06	**0.01**	**0.05**	0.36	0.62
Non-land assets variables (numbers of following asset)							
Assets in 2010							
mobile phone	1.19	1.47	1.61	0.57	**0.02**	**0.06**	**0.00**
motorbike	0.64	0.67	0.76	0.45	0.28	0.75	0.37
water pump	0.28	0.35	0.24	0.26	0.58	0.32	0.69
mist blower	1.12	1.30	1.41	0.52	**0.02**	**0.10**	**0.01**
power tiller	0.01	0.00	0.07	0.11	0.04	0.39	0.44
storage house	0.19	0.41	0.39	0.88	**0.00**	**0.00**	**0.00**
Assets in 2005							
mobile phone	0.53	0.45	0.87	**0.03**	**0.02**	0.52	0.31
motorbike	0.42	0.52	0.61	0.49	**0.06**	0.25	**0.05**
water pump	0.25	0.29	0.26	0.77	0.92	0.67	0.70
mist blower	0.96	0.79	1.17	**0.02**	0.17	0.18	0.87
power tiller	0.01	0.00	0.07	0.11	**0.03**	0.33	0.40
storage house	0.15	0.20	0.35	0.11	**0.00**	0.35	**0.01**
Chili production							
Area (m2)	3,137.45	4,465.80	5,304.97	0.55	**0.01**	**0.02**	**0.00**
Cost (million Rp)	11.28	20.81	22.56	0.79	**0.00**	**0.00**	**0.00**
Output(kg)	4,408.72	4,927.61	6,366.65	0.42	**0.08**	0.53	0.13
Types of chili planted	1.15	1.41	1.41	0.97	**0.00**	**0.00**	**0.00**
Seasons	1.24	1.62	1.67	0.68	**0.00**	**0.00**	**0.00**
Productivity (kg/ m^2)*	2.40	1.09	1.13	0.73	**0.07**	**0.02**	**0.00**
Cost per m^2 (Rp/ m^2) *	6408.72	4862.84	3583.01	**0.03**	0.14	0.34	**0.10**
Grading+	2.98	3.22	3.53	**0.02**	**0.00**	**0.06**	**0.00**
Price (Rp/kg)	6,236.60	7,637.88	8,976.27	**0.07**	**0.00**	**0.01**	**0.00**
Cost per kg (Rp/kg)*	3,052.92	4,992.44	3,900.03	**0.02**	**0.00**	**0.00**	**0.00**
Record keeping (Bivariate dummy, yes=1, no=0)							
Pesticide	0.12	0.47	0.46	0.89	**0.00**	**0.00**	**0.00**
date pesticide	0.06	0.15	0.15	0.99	**0.01**	**0.00**	**0.00**
price received	0.22	0.76	0.89	**0.08**	**0.00**	**0.00**	**0.00**
quantity output	0.21	0.76	0.85	0.25	**0.00**	**0.00**	**0.00**
For over 1 year	0.40	0.29	0.37	0.43	0.69	0.16	0.24

* outliers dropped. + See footnote 5 for detail of the grading scale. Significant t-test statistics are marked in bold.

Distance from the house to the nearest asphalt road is on average 200 meters shorter for modern channel producers as compared to traditional channel producers, and not statistically different between aware and unaware farmers. Distance from the house to a district center, in contrast, is significantly different between the groups. Aware farmers tend to live further away from the district market than traditional farmers, while unaware farmers live closer by.

Land size and ownership attributes

Farmers own on average 4,525 m^2 of land. The size is similar for the three groups of farmers. Ten percent have more than 1 hectare, while around 70 percent have less than 0.5 hectare, which makes it difficult to benefit from scale economies. Modern channel farmers cultivate a greater land area compared to traditional farmers (by 1,500 m^2) because they rent a larger area: aware farmers rent 3,458 m^2 on average (1.8 times greater than unaware farmers and 3.16 times greater than traditional farmers). The rented areas contribute to 51 percent of the total land area cultivated by aware farmers (compared to 32 and 24 percent for unaware and traditional farmers). The difference between each group is statistically different. Considering land quality, aware famers have a greater share of cultivated land irrigated (65%) and a much smaller share of dry land (29%) compared to both unaware and traditional farmers. Unaware farmers also have less irrigated and more dry land than the traditional farmers. Overall, 17 percent of all farmers cultivate land area over 1 hectare.

Non-land assets

Non-land assets are often pointed out in the literature as important determinants for modern channel participation. In 2005 aware farmers have on average significantly more mobile phones than unaware farmers or traditional farmers, which facilitate them to obtain market information and communicate with the traders. In 2010, a lot more farmers have mobile phones for communication, such that the difference between aware and unaware farmers disappeared, but the difference between modern and traditional farmers is significant. Transportation assets have been suggested to be critical for becoming a modern channel participant since supermarkets desire more frequent and timely delivery (Reardon et al., 2005). In our data, modern channel participants have more motor bikes in 2005 than traditional farmers. As for farming equipment, modern channel producers have more storage available. Aware farmers also have more mist blowers in 2005. Ownership of water pump and power tiller, however, do not differ across groups. As product appearance and quality consistency are important aspects emphasized by supermarkets better storage and equipment facilitating quality control can help farmers to meet supermarket standards.

Production attributes

Farmers differ in production efficiency. On average, modern channel farmers show a higher level of specialization in chili production. Aware and unaware farmers on average cultivate a larger area of chili, invest more in chili production and produce more outputs (kilos) than traditional channel farmers. The difference in output between aware and traditional farmers is nearly 2,000 kilos in 2010. As an additional indication for specialization, modern channel producers also plant more types of chili and are active in more seasons. The production costs also differ significantly between the three groups of farmers. Although modern channel farmers produce more than traditional farmers they also invest more. The total production cost is significantly higher for modern channel producers, but is insignificantly different between aware and unaware farmers.

Considering the differences in cost and land invested for chili production, the modern channel farmers seem to be the least productive (output per square meter), yet the most cost efficient ones (cost per square meters). The modern channel farmers produce on average 1.1 kilos of chili per square meter of cultivated land, while traditional producers more than double that amount. In line with the procurement quality requirements of the supermarket, chili produced by the modern producers receives better grading[32] and is sold at a higher price. The aware farmers received the highest average grading (3.53) for all three seasons, followed by the unaware farmers (3.22) and the traditional farmers with lowest grading (2.98).[33] In line with this, the average price (over three seasons) is highest for the aware farmers (8,976 Rp/kg) and lowest for traditional farmers (6,237 Rp/kg), with unaware farmers in between.

[32] The grading inquiry in the survey was of 9 grading level, starting from the highest (1. *Superfull*), followed by: (2. *Super*), (3. *Medium*), (4. *Small*), (5. mix of 1 and 2), (6. mix of 2 and 3), (7. mixed of 3 and 4), (8. Other) and (9. no grading), consecutively. The grading reported in table 6 is based on a converted 5-level grading scale, which marked *superfull* as class A, combined 2 and 5 into class B, combined 3and 6 into class C, combined 3 and 6 into class D, and lastly the original 8 and 9 combined into lowest class E. The new A to E scale is than coded into 1 to 5 for quantitative comparison, with 5 as the highest grade and 1 as the lowest.

[33] We report only the average grading received over the three seasons. Detail of the separate grading for each season is available upon request. All producers receive higher chili grading in the early season (1.April) as compared to chili output in the later seasons (2.July and 3.September). Out of three seasons, the mean difference in grading is statistically significant between aware and unaware farmers in season 1 and 2, while the mean grading differences is statistically different between farmers of different channel in all three seasons.

The costs required to produce better chili with better quality and higher price translates into higher investment per kilo of output. Unaware farmers invest most per kilo of output, followed by the aware and the traditional farmers. The difference between the three groups is statistically significant. Overall, the profit margin per kilo of chili is highest for the aware farmers, followed by the traditional and the unaware farmers.

Table 4.4 Summary statistics

Variables	Definition (units)	mean	s.d.
Dependent variables			
Net income	Net household income from various source in log (Rp.)	16.548	1.040
Chili profit	Net profit (revenue less cost) from chili production in log (Rp.)	15.228	1.413
Variables of interest			
Chili cost	Total cost invested for chili production in log (Rp.)	15.758	1.141
High Cost	Total cost invested for chili production above the median of all sampled farmers (dummy)	0.500	0.500
High cost per m^2	Total cost invested for chili production per square meter above the median of all sampled farmers (dummy)	0.499	0.500
Traditional	Whether farmers supply chili through the traditional channel (dummy)	0.812	0.391
Unaware	Farmers who supply to the supermarket but thought they are traditional channel producers (dummy=1 if unaware)	0.111	0.314
Aware	Farmers who know they are supplying to the supermarket (=1 if aware)	0.078	0.267
Productivity (γ)	Chili output produced per unit of land in log (KG/m^2)	0.155	0.996
Control variables			
Grading	Average chili grading received. Range between 1 to 5, with 5 as the best grading and 1 for the chili with the worst or without grading	3.047	0.985
Household size	Numbers of family members in the household (persons)	4.531	1.570
Age	Age of the household head (years)	45.769	10.979
Education	Years of education of the head of the household (years)	6.725	3.019
% working age	Proportion of family member between the age 15 to 65 (%)	0.658	0.197
% retired age	Proportion of family member above age 65 (%)	0.027	0.098
Experience	Years of experience in chili production (years)	8.920	6.750
Mobile phone	Mobile phone ownership in 2005 (5 years ago from the time of the survey (dummy)	0.745	0.436
Motor bike	Motor bike ownership in 2005 (dummy)	0.529	0.500
Water pump	Water pump ownership in 2005 (dummy)	0.265	0.442
Mist blower	Mist blower ownership in 2005 (dummy)	0.848	0.360
Power tiller	Power tiller ownership in 2005 (dummy)	0.013	0.115
Storage house	Storage house ownership in 2005 (dummy)	0.221	0.415
Chili land	Total land size devote to chili production (m^2) in 2010 (current asset)	8.280	1.144
Land owned	Land size owned (m^2) in 2005	6.055	3.663
% rented	Proportion of cultivated land rented (%)	0.273	0.378
Prod. record	Whether farmers keep production record, include: pesticide use and date of pesticide application on chili (dummy)	0.128	0.285
Market. record	Whether farmers keep marketing record, include: chili price, quantity sold (dummy)	0.327	0.464
Exogenous variable			
Distance	Distance to the nearest asphalt road	0.262	0.595

Record keeping

The percentage of record keeping among modern channel producers are doubled that of traditional producers. A significantly larger share keeps records on pesticide use, date of pesticide applied, price received, and quantities sold. The percentages vary a lot. Better educated farmers may not only have the ability to keep written records but also benefit from the insights delivered through better use of knowledge. We merge the 5 record keeping items into 2 record behavior categories, those related to production and those related to marketing.

Bivariate independent variables—high cost production dummy

We set an arbitrary cutoff investment level to distinguish between high cost and low cost producers because the true critical investment level required to produce high quality chili is unknown. In the treatment effect analysis, the critical bivariate investment decision variable is defined as the farmers who made investment above the median level among all sampled farmers. Thus by definition half of the households invest in high cost production. Similarly, we generate a relatively high cost per square meter dummy variable as a robustness check. Farmers whose investment in chili production is above the median level per square meter of land cultivated are considered high cost producers. Table 4 gives the summary statistics of all variables used.

4.5 Regression analysis

This section reports the empirical analysis of the self-selection phenomena and the effect of investment choice, channel choice and awareness on the profitability of chili production and on overall income.

4.5.1 Investment choice and effect on profit model

The output from the treatment effect model is split into two parts, with the estimation from the first stage selection equation reported in the lower panel and the second stage welfare output reported in the upper panel of Table 4.5. In the upper panel, the result from the ordinary least square (OLS) estimation is reported in column 1 and 3 aside to the treatment effect model in column 2 and 4. In the lower panel, we show independent probit estimation in column 1 and 3 next to the first stage outcome from the treatment effect model in column 2 and 4. We demonstrate two model specifications, with (column 3 and 4) and without (column

1 and 2) controlling for record keeping habit. The following model specifications are estimated:

$$\text{high cost} = \psi(\text{ productivity, controls}) + \theta(\text{distance}) + u \qquad (1^{st} \text{ stage})$$

$$\text{chili profit} = \vartheta(\text{highCost}) + \psi(\text{awareness, productivity, controls}) + v \ (2^{nd} \text{ stage})$$

The first stage bivariate decision between entering into relatively high and relatively low cost production, which relates to higher and lower probability of supplying to the supermarket channel, is significantly correlated with farmers' productivity (Table 4.5, lower panel). More productive farmers are more likely to invest in high cost for chili production, even after controlling for farmers' household characteristics, land ownership and other assets. Among the control factors, we see that more educated and more experienced farmers are more likely to invest in high cost production. In addition, non-land assets such as motor bike and water pump positively correlate with farmers' endogenous decision in making high cost investment.

The instrument variable, distance to road, is statistically significant. The negative coefficients suggest that farmers who live at more remote location are less likely to invest in high cost production. This is in line with what we expected that the lower the search cost the traders had to incur, the higher the farmers are to engage in high cost production.

In the second stage, the *high cost investment* decision dummy variable significantly affects farmers' profit from chili production in all specifications (Table 4.5, upper panel). The estimates from the treatment effect model indicates that other things being equal, farmers who invest in high cost production earn on average 140 percent higher profit than farmers who invest in relatively low cost production. The difference is net of the observed selection bias concerning farmers' investment choice in the first stage. This is much higher than the 70 percent as suggested from the OLS estimation. In addition, the estimated coefficient regarding farmers' productivity is a statistically significant factor affecting both the investment decision (1^{st} stage) and the profitability of chili production (2^{nd} stage). The results show that more productive farmers are not only more likely to invest in high cost; their profit is also higher as compared to less productive farmers. The result mirrors our earlier prediction from the theoretical model regarding the productivity sorting of farmers into different investment level, and how it affects profitability.

Table 4.5. Treatment effect, OLS and Probit regression output – chili profit

	(1)		(2)		(3)		(4)	
a. Outcome model: OLS and 2nd stage of the treatment effect model; dependent variable: chili profit.								
Model:	OLS		Treatment 2nd stage		OLS		Treatment 2nd stage	
High Cost	0.742***	(6.306)	1.483***	(4.184)	0.698***	(5.938)	1.405***	(4.143)
Unaware	0.424**	(2.445)	0.405**	(2.381)	0.229	(1.256)	0.223*	(1.700)
Aware	1.035***	(5.422)	0.996***	(5.306)	0.807***	(3.964)	0.772***	(4.554)
Productivity	0.392***	(5.761)	0.363***	(5.145)	0.396***	(5.866)	0.366***	(4.525)
Grading	0.148**	(2.517)	0.140**	(2.418)	0.143**	(2.447)	0.135**	(2.200)
Household size	-0.047	(-1.19)	-0.051	(-1.255)	-0.041	(-1.04)	-0.047	(-1.165)
Age	-0.013**	(-2.12)	-0.013**	(-2.036)	-0.012*	(-1.91)	-0.012**	(-1.979)
Education	-0.010	(-0.55)	-0.023	(-1.136)	-0.017	(-0.91)	-0.026	(-1.315)
% working age	0.359	(1.056)	0.431	(1.230)	0.445	(1.317)	0.481	(1.396)
% retired age	-0.891	(-1.39)	-0.807	(-1.225)	-0.912	(-1.44)	-0.819	(-1.367)
Mobile phone	0.013	(0.103)	-0.042	(-0.314)	-0.020	(-0.16)	-0.063	(-0.491)
Motor bike	0.236**	(2.031)	0.181	(1.484)	0.251**	(2.171)	0.192*	(1.689)
Water pump	0.086	(0.675)	0.004	(0.029)	0.060	(0.475)	-0.010	(-0.078)
Mist blower	0.278*	(1.759)	0.137	(0.786)	0.258	(1.646)	0.135	(0.827)
Power tiller	0.431	(0.912)	0.498	(1.024)	0.503	(1.074)	0.558	(1.255)
Storage house	0.255*	(1.910)	0.196	(1.405)	0.209	(1.569)	0.174	(1.425)
Experience	0.014	(1.619)	0.011	(1.252)	0.016*	(1.845)	0.012	(1.280)
Chili land	0.362***	(5.656)	0.348***	(5.486)	0.360***	(5.653)	0.350***	(4.975)
Prod. record					0.163	(0.775)	-0.004	(-0.017)
Market. record					0.358**	(2.432)	0.319**	(2.345)
Constant	11.16***	(17.99)	11.23***	(18.042)	11.04***	(17.91)	11.13***	(18.026)
adj. R^2	0.383				0.395			
b. Selection model: bivariate investment decision; dependent variable: high cost (dummy)								
Model:	Probit		Treatment 1st stage		Probit		Treatment 1st stage	
Productivity	0.158***	(2.695)	0.115	(1.527)	0.171***	(2.888)	0.132*	(1.745)
Household size	-0.008	(-0.207)	0.007	(0.150)	-0.005	(-0.12)	0.011	(0.245)
Age	-0.003	(-0.493)	-0.003	(-0.381)	-0.001	(-0.11)	0.001	(0.090)
Education	0.050**	(2.301)	0.039*	(1.676)	0.042*	(1.852)	0.027	(1.192)
% working age	-0.335	(-0.933)	-0.315	(-0.784)	-0.277	(-0.77)	-0.225	(-0.543)
% retired age	-0.373	(-0.549)	-0.325	(-0.434)	-0.445	(-0.65)	-0.411	(-0.482)
Mobile phone	0.329**	(2.437)	0.262*	(1.747)	0.312**	(2.275)	0.216	(1.438)
Truck	0.195	(1.622)	0.268**	(1.998)	0.220*	(1.797)	0.316**	(2.275)
Motor bike	0.282**	(2.096)	0.262*	(1.726)	0.260*	(1.916)	0.223	(1.485)
Water pump	0.47***	(2.859)	0.491***	(2.590)	0.46***	(2.763)	0.456**	(2.366)
Mist blower	-0.222	(-0.450)	-0.206	(-0.350)	-0.232	(-0.47)	-0.230	(-0.418)
Power tiller	0.125	(0.888)	0.189	(1.246)	0.076	(0.527)	0.138	(0.907)
Storage house	0.016*	(1.758)	0.013	(1.311)	0.019**	(2.071)	0.016	(1.502)
Experience	-0.008	(-0.207)	0.058***	(2.883)	-0.005	(-0.12)	0.132*	(1.745)
Land owned	0.07***	(3.702)	0.739***	(4.049)	0.07***	(3.433)	0.05***	(2.654)
% land rented	0.71***	(4.082)	0.115	(1.527)	0.73***	(4.088)	0.76***	(4.108)
Prod. record					0.76***	(3.031)	0.81***	(3.003)
Market. record					-0.003	(-0.02)	0.049	(0.306)
Distance to road	-0.169	(-1.612)	-0.239**	(-2.183)	-0.157	(-1.49)	-0.215**	(-2.189)
Constant	-1.5***	(-3.446)	-1.39***	(-2.817)	-1.69***	(-3.83)	-1.63***	(-3.244)
Pseudo R^2	0.121				0.137			
N	462		462		462		462	
athrho			-0.436**	(-2.172)			-0.419**	(-2.063)
lnsigma			0.132**	(2.425)			0.116**	(2.168)
Wald chi^2			247.054	p=0.00^			273.568	p=0.00^
Independent test			3.536	p=0.00^			4.255	p=0.04^

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$; ^ are p-value for Prob. > chi2

The coefficient for the aware and unaware farmers dummy variables are both positive and statistically significant, suggesting that modern channel participation gives producers additional profit as compared to the traditional farmers, even after controlling for the investment differences. More interestingly, the profit premium is significantly lower for unaware farmers as compared to the aware farmers. The aware farmers earn on average 72 percent higher profit than traditional channel farmers, while the profit premium for unaware farmers is only 22 percent higher. Since both aware and unaware farmers are both modern channel producers, the difference in premium between aware and unaware farmers has important implications. The stark difference in premium suggests that participation in the modern retail channel alone though provide excessive profit but can differ significantly depending on awareness of participation. The significantly lower profit premium from supplying to the modern channel for the unaware farmers as compared to the aware farmers can potentially be explained by the difference in bargaining power or farmers' ability in assessing traders' willingness to pay. Farmers have to be aware of the marketing channel of their product to capture a greater proportion of the profit margin from supplying to the modern channel. In all specifications, we controlled for the quality of the output (grading), land area for chili production, other household characteristics and non-land assets owned.

The sign of estimated coefficients from the OLS and probit model as compared to the treatment effect model is the same for all the variables with significant estimates. The standard deviations of the estimated coefficients from the treatment effect model are much smaller for the key variables. The coefficients for most variables are not very different, with exception of the estimated effect of the high-cost bivariate dummy, which is much higher from the treatment effect model than the outcome from the OLS estimation.

An interesting result show up when record keeping behavior is distinguished between production and marketing records. We see that while production record (pesticide application dates and amount of pesticide used) is significant in determining the first stage investment decision, keeping marketing record (price and quantity) instead is significant in affecting the second stage profit outcome. This result reinforces the link we find between productivity and investment decision. Production record as knowledge base for self-evaluation of productivity can support investment decision, while marketing record provide bases for assessing the market condition (current price) rewards the farmers with higher profit.

Lastly, to justify our model choice, the reported Wald test of all coefficients (except constant) in the output (2nd stage) model being zero is rejected in both of the treatment effect specifications. The test statistics are $\chi^2 = 152$ and $\chi^2 = 282$, with $p < 0.01$, indicating the covariates used in the model are appropriate and at least one of the covariates used in the regression is not equal to zero. In addition, if $\rho = 0$, the errors u and v are independent and there is no endogeneity problem. Wald test for independence reported at the bottom of Table 4.5 compares the joint likelihood of an independent probit model for the selection equation and an OLS regression model on the observed data against the treatment effect model likelihood. The likelihood test for $H_0 : \rho = 0$ gives $\chi^2 = 3.96$ and $\chi^2 = 5.68$ with both $p < 0.01$ rejects the null hypothesis at a statistically significant level. This suggests that using OLS regression results in bias estimation of the effect of high cost investment, and instead applying the treatment effect model is indeed appropriate since it allows the correlation between the two error terms to be non-zero ($\rho \neq 0$).

4.5.2 Channel choice and effect on net income

In this section, we explore the specification many research conducted in testing the influence of participating in the modern food retail channel on income. Knowing the importance of investment on quality which may further influence modern channel participation, we estimate the following two stage treatment effect model:

$$\text{channel choice} = \theta(\text{distance}) + \psi(\text{cost, productivity, controls}) + u \qquad (1^{st}\ \text{stage})$$

$$\text{net income} = \vartheta(\text{channel choice}) + \psi(\text{awareness, productivity, controls}) + v \qquad (2^{nd}\ \text{stage})$$

Notice that we included the awareness indicator into the specification. The results are reported in Table 4.6, which is organized in the same way as in Table 4.5, with record keeping behavior in the second specification, the probit and OLS regression reported in column 1 and 3 and the first and second stage of the treatment effect model reported in column 2 and 4.

From the outcome model, the positive and significant coefficients suggest that participating in the modern food retail channel increases farmers' net income by 75 percent as compared to non-participating farmers. This corroborates the earlier finding of a positive correlation between participation and income. In contrast, using the OLS model instead gives negative, insignificant coefficients for the modern channel variable. In addition, our awareness indicator further captures the additional welfare loss unaware farmers do not enjoy. After controlling for the first stage modern channel participation choice, awareness increases farmers' net income by an additional 37 percent. The effect of channel choice is stronger when record keeping character is controlled (column 4).

From the selection model, the positive and significant coefficient of the chili cost suggested that farmers who invest more are more likely to become modern channel producers. The coefficient for the identification variable, distant to asphalt road, is negative and statistically significant in both treatment effect specifications. In the probit model, the distant coefficient instead become insignificant when record keeping behavior is included in the model.

Among the control variables, we see that most household characteristics do not have a strong correlation with farmers' channel choice, with the exception of experience and water pump ownership. Similar to other studies, we see that more experienced farmers are less likely to participate in the modern channel. It is argued that the longer the farmers have been growing chili, the stronger the relationship they formed with the buyers they are familiar with. These farmers are then less likely to seize new market opportunity, especially when it requires making adjustments in production method for better output quality. The independent probit model explains slightly less than 14 percent of the variation, but is significantly improved to nearly 30 percent when record keeping behavior is included in the specification.

In addition, the results suggest that farmers' net income also correlates positively with farm households' non-land assets. Farm household with more experience in chili production have higher income. Household who have a higher percentage of farm land rented also have higher income. This suggests that these farmers are likely to have better knowledge and capability to mobile their production capacity when necessary. Notice that unlike the previous section, keeping marketing record is significantly affecting farmers channel choice in the first stage. The result suggests that better documentation of price and quantity of ones' product increases farmers' knowledge about the market. The record keeping behavior however is costly and shows up as significantly negative factors for net income. Overall, the Wald test and independence test again confirm the appropriateness of using the treatment effect model.

Table 4.6 Treatment effect, OLS and Probit regression output -- Income

	(1)		(2)		(3)		(4)	
a. Outcome model: total net income.								
Model:	OLS		2nd stage treatment		OLS		2nd stage treatment	
Channel (0/1)	-0.179	(-1.42)	0.693***	(4.12)	-0.216	(-1.60)	0.750***	(4.26)
Aware (0/1)	0.365**	(2.01)	0.408**	(2.39)	0.367**	(2.01)	0.367**	(2.16)
Household size	0.025	(0.99)	0.029	(1.08)	0.026	(1.04)	0.029	(1.11)
Age	0.003	(0.61)	0.005	(1.05)	0.003	(0.69)	0.004	(0.89)
Education	0.033**	(2.30)	0.017	(1.15)	0.030**	(2.14)	0.023	(1.56)
% working age	0.007	(0.03)	-0.043	(-0.16)	0.036	(0.14)	-0.033	(-0.12)
% retired age	-0.260	(-0.64)	-0.588	(-1.27)	-0.266	(-0.66)	-0.515	(-1.13)
Mobile phone	0.30***	(3.22)	0.290***	(2.97)	0.296***	(3.14)	0.318***	(3.23)
Motor bike	0.44***	(5.21)	0.413***	(4.75)	0.440***	(5.24)	0.423***	(4.88)
Water pump	0.184**	(2.09)	0.222**	(2.40)	0.178**	(2.03)	0.230**	(2.50)
Mist blower	0.398***	(3.52)	0.337***	(2.88)	0.398***	(3.51)	0.352***	(3.05)
Power tiller	0.478*	(1.70)	0.507**	(2.12)	0.477*	(1.76)	0.443*	(1.91)
Storage house	0.513***	(5.75)	0.416***	(4.30)	0.498***	(5.59)	0.462***	(4.90)
Experience	0.016***	(2.81)	0.024***	(3.80)	0.017***	(2.92)	0.022***	(3.58)
Land owned	0.025*	(1.92)	0.017	(1.33)	0.023*	(1.77)	0.016	(1.28)
% land rented	0.307***	(2.60)	0.253**	(2.21)	0.312***	(2.66)	0.303***	(2.68)
Prod. record					0.240	(1.62)	0.215	(1.34)
Market. record					-0.017	(-0.15)	-0.385***	(-3.09)
Constant	14.75***	(45.03)	14.684***	(43.810)	14.714***	(44.52)	14.736***	(44.16)
Adj. R2	0.261				0.262			
b. Selection model on bivariate channel choice decision								
Model:	Probit		1st stage treatment		Probit		1st stage treatment	
Chili cost	0.255***	(3.54)	0.323***	(4.70)	0.181**	(2.28)	0.251***	(3.59)
Household size	-0.027	(-0.55)	-0.019	(-0.42)	-0.036	(-0.67)	-0.032	(-0.63)
Age	-0.008	(-1.05)	-0.008	(-1.20)	-0.001	(-0.17)	-0.005	(-0.567)
Education	0.057**	(2.39)	0.054**	(2.48)	0.033	(1.25)	0.020	(0.75)
% working age	0.244	(0.58)	0.216	(0.52)	0.334	(0.71)	0.297	(0.65)
% retired age	1.199	(1.63)	1.396*	(1.87)	1.018	(1.30)	1.291	(1.62)
Mobile phone	0.031	(0.19)	-0.063	(-0.40)	-0.088	(-0.48)	-0.178	(-1.05)
Motor bike	0.027	(0.19)	0.004	(0.03)	0.035	(0.23)	-0.020	(-0.14)
Water pump	-0.255	(-1.62)	-0.313**	(-2.11)	-0.292*	(-1.70)	-0.333**	(-1.96)
Mist blower	0.206	(1.02)	0.199	(0.99)	0.151	(0.67)	0.227	(1.17)
Power tiller	-0.290	(-0.52)	-0.149	(-0.40)	-0.004	(-0.01)	0.142	(0.43)
Storage house	0.325**	(2.06)	0.234	(1.47)	0.116	(0.66)	-0.039	(-0.22)
Experience	-0.05***	(-4.02)	-0.047***	(-4.27)	-0.040***	(-2.87)	-0.035***	(-3.02)
Chili land	0.072	(0.94)	0.125*	(1.65)	0.107	(1.178)	0.213**	(2.33)
Prod. record					-0.050	(-0.21)	-0.028	(-0.12)
Market. record					1.389***	(8.14)	1.426***	(7.96)
Distance to road	-0.447**	(-2.14)	-0.484***	(-3.552)	-0.305	(-1.22)	-0.452**	(-2.50)
Constant	-5.40***	(-4.81)	-6.810***	(-6.348)	-5.260***	(-4.25)	-6.974***	(-6.24)
PseudoR2	0.146				0.299			
N	462		462		462		462	
athrho			-0.658***	(-6.815)			-0.734***	(-6.735)
lnsigma			-0.061	(-1.475)			-0.066	(-1.603)
Wald_chi2			263.449	p=0.000			45.366	p=0.000
Independent test			46.443	p=0.000			41.815	p=0.000

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$

4.5.3 Robustness check

As a robustness check for chili profit model, we use cost per square meter invested instead of the total investment (Table 4.7). Again the treatment effect model suggests that more productive farmers engage in relatively high cost (per unit land) production, which is rewarded with higher profit. Profit premium from participating in the modern channel is still positive and significant. Aware farmers are again enjoying higher profit margin from participating in the modern channel as compared to the unaware farmers. The OLS and probit model shows similar results for most of the variables, but with insignificant coefficient for the unaware farmers and investment level in affecting chili profit. The profit premium for the aware farmers as compared to the traditional farmers is significantly positive and is doubled in size of the premium enjoyed by the unaware farmers as compared to the traditional farmers. Overall, the Wald-test is again in support of the treatment effect model.

Table 4.7 Robustness check for chili profit

Dependent var.	(1) High cost /m2		(2) High cost /m2		(3) Chili profit		(4) Chili profit	
model	Probit		Treatment 1st stage		OLS		Treatment 2nd stage	
Productivity	0.883***	(9.38)	1.022***	(8.05)	0.449***	(5.98)	0.256**	(2.15)
High Cost					0.134	(1.06)	1.008***	(2.64)
Unaware					0.305	(1.61)	0.311**	(2.27)
Aware					0.780***	(3.67)	0.780***	(4.68)
Grading					0.144**	(2.38)	0.118*	(1.87)
Distance to road	-0.092	(-0.74)	-0.223	(-1.43)				
Production record	0.305	(1.23)	0.330	(1.29)	0.261	(1.21)	0.216	(1.03)
Marketing record	0.272*	(1.71)	0.401**	(2.30)	0.386**	(2.52)	0.240	(1.55)
Constant	-1.50***	(-3.16)	-1.92***	(-3.54)	10.2***	(16.35)	10.2***	(16.74)
Pseudo /adj. $R2$	0.258				0.348			
N	577		459		562		459	

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$; Wald chi2(21) = 280.82 (prob>chi^2=0.000), independence test = 4.543 (prob>chi^2=0.033); athrho = -0.511(-2.131)

So far, we have treated investment as dummy variable, indicating the relatively high or low cost investment farmers engaged in for chili production. To expose the full information we have from the data, we run two separate OLS regressions. The first regression is the same as our probit specification in estimating the total cost farmers invested for chili production in section 4.5.1. We then use the predicted cost from the first regression to explain farmers profit from chili production for the second OLS regression. The variables used are the same as before except for the cost variable, which is now continuous. Table 4.8 reports the coefficients for the key variables of interest. The complete output is provided in the appendix A4.1 and A4.2.

Robustness check for the effect of investment decision on profit again supports our earlier findings (Table 4.8). The coefficient from the predicted cost indicates that a 1 percent increase in cost result in 45 percent higher chili profit, while a 1 percent increase in productivity increases farmers total cost invested in chili production by 25 percent. In addition, supplying to the modern channel increases farmers profit even after controlling for the investment level. Unaware modern channel producers have on average 33 percent higher income than the traditional channel producers, while aware modern channel producers have even higher profit margin (75 percent higher) as compared to the traditional channel producers. The difference in profit margin between the aware and the unaware farmers is statistically significant

Table 4.8 Robustness check for chili profit with continuous cost[+]

Dependent variable	(1) Linear total Cost		(2) Chili profit	
Productivity	0.249***	(5.679)	0.363***	(3.948)
Predicted total cost			0.455**	(1.992)
Unaware			0.329*	(1.764)
Aware			0.745***	(3.531)
Grading			0.134**	(2.199)
Distance to road	0.031	(0.375)		
adj. *R*2	0.2339		0.3499	

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$. [+]Full specification is the same as the model in column 4 of Table 5 with exception of different cost variable used. Coefficients for the control variables are omitted in the Table

4.6 Conclusion

The dynamics in the food retail sector which led to the rising importance of high-value supermarkets in developing countries provided both challenges and opportunity for farmers. s. Marginalization of small farmers has been one of the main concerns. While a growing number of studies look into the determinants of modern channel participation, a theoretical foundation is often lacking. In this article, we provide a simple model illustrating the self-selection of farmers into different level of investment for production, and how the result of the self-selection affects farmers' welfare. We thus provide a theoretical foundation for analyzing the increasing demand for high quality food supplied by the supermarkets. We link investment differentials to the likelihood of modern channel participation as private standards implicitly require higher input investment to insure output quality and consistency.

Empirical studies supported the connection between the investment level and modern channel participation. Based on evidence from micro-farm household data from Indonesia, we arrive at three main conclusions:

First, farmers who invest in high cost production is rewarded with high profit. There is clearly self-selection based on farmers' productivity for the investment level chosen. Land ownership, and ownership of other non-land assets, such as truck and water pump, are also important factors that endogenously determines farmers' investment choice.

Second, we find a positive and significant income affect from modern channel participation. Moreover, self-selection of participation is correlated with farmers' investment level. In other words, more productive farmers are more likely to make high level of investment for production. More investment increases farmers' chance of producing high quality output, which raises the likelihood to become suppliers for the modern food retail channel. These results confirm the previous concerns that the requirements of supermarket in output quality and consistency pose a challenge to the resource-poor and credit constrained farmers in participating in the modern food retail channel.

Third, depending on the farmers' awareness of their participation in the modern food retail channel, farmers who are unaware of their participation can be deprive from obtaining the full profit margin as compared to their neighbors who are aware of their participation. Controlling for investment choice, the profit premium enjoyed by the aware farmers can be as large as four times larger than the profit premium enjoyed by the unaware farmers. Our results mirror the two gaps Torero and Gulati (2004) had pointed out: the 'real access gap' and the 'market efficiency gap', which needs to be solved by institution and infrastructure (Torero, 2011; Gulati et al., 2007). We show that even after farmers overcome the 'real access gap' of becoming an efficient and capable modern channel suppliers, those who fail to address the 'market efficiency gap' regarding farmers' ability to utilize market information will be unable to reap the full benefit of engaging in the modern food retail value chain.

Acknowledgement

This study was made possible by funding from the Australian Centre for International Agriculture Research (ACIAR). The project researchers gratefully acknowledge the support and guidance provided by ACIAR project managers Simon Hearn and David Shearer. Special thanks to PhD students who participated and gave helpful comments and suggestions in the seminar for which the idea of this paper is first presented. The paper was also presented at the 2013 annual Asia Pacific Trade Seminar (APTS) in Nanjing, China and at Utrecht University School of Economics (USE). We thank the seminar participants for the valuable feedback and comments. Lastly, we thank Professor Christopher Findlay for introducing and connecting the authors of this work, as well as for all the comments and suggestions given along the way.

4.7 References

Asfaw, A. (2007), "Supermarket purchases and the dietary patterns of households in Guatemala", IFPRI discussion paper 696. Washington, DC.

Asfawa, S., D. Mithöfer and H. Waibel (2010), Agrifood supply chain, private-sector standards, and farmers' health: evidence from Kenya", *Agricultural Economics* 40: 251-263.

Berdegué, J.A, F. Balsevich, L. Flores and T. Reardon (2005), "Central American supermarkets' private standards of quality and safety in procurement of fresh fruits and vegetables", *Food Policy*, 30(3), 254–269.

Berdegué , J. A., R. Hernández, J. Ortega, and T. Reardon (2007). Strawberry growers and modern market channels in Mexico. Micro Report module 3 of component 1. Regoverning Market Program, IIED, London.

Boselie, D., S. Henson and D. Weatherspoon (2003), "Supermarket procurement practices in developing countries: redefining the roles of the public and private sectors", *American Journal Agricultural Economics*, 85(5): 1155–1161.

Chowdhury, S., A. Gulati, and E.Gumbira-Said, (2005), "High value products, supermarkets and vertical arrangements in Indonesia", MTID Discussion Paper No.83. International Food Policy Research Institute, Washington, pp.1-83.

Christiaensen, L., L. Demery and K. Jesper (2011), "The (evolving) role of agriculture in poverty reduction – An empirical perspective", *Journal of Development Economics*, 96: 239-254.

Dethier, J.J and A. Effenerger (2012), "Agricultural and development: a brief review of the literature", *Economic System*, 36: 175-205.

FAO. (2011), "The state of food and agriculture", FAO, Rome.

Greene, W. H. (2008), Econometric analysis. Upper Saddle River, NJ: *Prentice Hall.*

Gulati, A., N. Minot, C. Delgado and S. Bora (2007), "Growth in high-value agriculture in Asia and the emergence of vertical links with farmers, " p91-108 in Swinnen, J. (Eds.), *Global Supply Chains, Standards and the Poor: How the Globalization of Food Systems and Standards Affects Rural Development and Poverty*. London: CABI Press.

Heckman, J. (1978), "Dummy endogenous variables in a simultaneous equation system", *Econometrica*, 46(4): 931-959.

Heckman, J. (1979), "Sample selection bias as a specification error", *Econometrica*, 47(1): 153-162.

Henson, S. and T. Reardon (2005), "Private agri-food standards: implications for food policy and the agri-food system", *Food Policy*, 30: 241–253.

Hernández, R., T. Reardon and J.A. Berdegué (2007), "Supermarkets, wholesalers, and tomato growers in Guatemala", *Agricultural Economics*, 36(3), 281–290.

Jaffee, S., and J. Morton (1994), *Marketing Africa's high value foods: Comparative experiences of an emergent private sector*. Dubuque, Iowa: Kendall/Hunt Publishing Co..

Maddala, G. S. (1983), Limited-dependent and qualitative variables in econometrics. Cambridge: Cambridge University Press.

Maddala, G. S. (1986), "Disequilibrium, self-selection, and switching models", p.1633-1682 in Z. Griliches, and D.I. Michael (Ed.), *Handbook of Econometrics*. North-Holland: Elsevier.

Minten, B., L. Randrianarison and J.F.M. Swinnen (2009) "Global retail chains and poor farmers: Evidence from Madagascar", *World Development*, 37 (11): 1728–1741.

Minot, N. (1986), Contract farming and its effect on small farmers in less developed countries, Working paper no. 31. Lansing: Michigan State University.

Minot, N. and D. Roy (2006), "Impact of high-value agriculture and modern marketing channels on poverty: an analytical framework", Mimeo. Washington, DC: Markets, Trade and Institutions Division paper no. 120 and MTID discussion paper no 73. Washington, DC: International Food Policy Research Institute.

Miyata, S., N. Minot and D. Hu (2009), "Impact of contract farming on income: Linking small farmers, packers, and supermarkets in China", *World Development*, 37 (11): 1781–1790.

Mergenthaler, M., K. Weinberger and M. Qaim. (2009a), "Consumer valuation of food quality and good safety attributes in Vietnam", *Review of Agricultural Economics,* 31(2): 266-283.

Mergenthaler, M., K. Weinberger and M. Qaim. (2009b), "The food system transformation in developing countries: a disaggregate demand analysis for fruits and vegetables in Vietnam", *Food Policy* 34 (5): 426-436.

Moustier, P., Phan Thi Giac Tam, Dao The Anh, Vu Trong Binh and Nguyen Thi Tan Loc (2009), "The role of farmer organizations in supplying supermarkets with quality food in Vietnam", *Food Policy*, 35 (1): 69-78.

Natawidjaja, R., T. Reardon, S. Shetty, T.I. Noor, T. Perdana, E. Rasmikayati, S. Bachri, and R. Hernandez (2007), "Horticultural Producers and Supermarket Development in Indonesia", UNPAD/MSU Report published by the World Bank/Indonesia, June.

Neven, D. and T. Reardon (2004), "The rise of Kenyan supermarkets and the evolution of their horticulture product procurement systems", *Development Policy Review*, 22(6): 669-699.

Neven, D., T. Reardon, J. Chege and H. Wang (2006). "Supermarkets and consumers in Africa: the case of Nairobi, Kenya", *Journal of International Food and Agribusiness Marketing*, 18(1/2): 103-123.

Neven, D., M. M. Odera, T. Reardon, and H. Wang (2009), "Kenyan supermarkets, emerging middle-class horticultural farmers, and employment impacts on the rural poor", *World Development*, 37(11), 1802–1811.

Natawidjaja, R., T. Reardon, S. Shetty, T.I. Noor, T. Perdana, E. Rasmikayati, S. Bachri, and R. Hernandez (2007), "Horticultural producers and supermarket development in Indonesia", UNPAD /MSU/World Bank. *World Bank report* no. 38543.

Rao, E.J.O. and M. Qaim (2010), "Supermarkets, farm household income, and poverty: insights from Kenya", *World Development*, 39(5):784-796.

Reardon, T., C.B. Barrett, J.A. Berdegué and J.F.M Swinnen (2009), "Agrifood industry transformation and small farmers in developing countries", *World Development*, 37(11): 1717-1729.

Reardon, T., and C.B. Barrett (2000), "Agro-industrialisation, globalization and international development: an overview of issues, patterns and determinants", *Agricultural Economics*, 23: 195–205.

Reardon, T., J.A. Berdegué and P. Timmer (2005), "Supermarketization of the 'emerging markets' of the Pacific Rim: development and trade implications", *Journal of Food Distribution Research* 36(1):3-12.

Reardon, T., J. A. Berdegué, R. Medina, M. Aguirre, F. Echánove, A. Martínez, R. Cook, N. Tucker and F. Balsevich (2007). "The Rise of Supermarkets in Mexico: Implications for the Domestic Horticultural Supply Chain and Development", *USDA Report*, East Lansing: Michigan State University.

Reardon, T., J.M. Codron, L. Busch, J. Bingen, and C. Harris (2001), "Global change in agrifood grades and standards: agribusiness strategic responses in developing countries", *International Food and Agribusiness Management Review*, 2: 195–205.

Reardon, T., S. Henson and J. Berdegué, (2007) "'Proactive fast-tracking' diffusion of supermarkets in developing countries: implications for market institutions and trade", *Journal of Economic Geography* 7: 399-431.

Reardon, T., and J. Swinnen (2004), "Agrifood sector liberalization and the rise of supermarkets in former state-controlled economies: Comparison with other developing countries", *Development Policy Review*, 22(5): 515–524.

Reardon, T., P. Timmer, C.B. Barrett and J. Berdegué (2003), "The rise of supermarkets in Africa, Asia and Latin America", *American Journal of Agricultural Economics*, 85(5): 1140-1146

Reardon, T., P. Timmer and J.A. Berdegué (2004), "The rapid rise of supermarkets in developing countries: induced organizational, institutional, and technological change in agrifood system", *Journal of Agricultural and Development Economics*, 1(2): 15-30

Regmi, A., and J. Dyck (2001), "Effects of urbanization on global food demand". In: Regmi, A. (Ed.), Changing Structures of Global Food Consumption and Trade. ERS WRS 01-1. Economic Research Service, United States Department of Agriculture, Washington, DC.

Sahara (2012), "Channel choice decisions and market relationships between farmers and buyers in the Indonesian chili industry", University of Adelaide dissertation series; Adelaide, Australia.

Sahara, W. Umberger, N. Minot, and R. Stringer (2011), "Determinants of market channel choice Among chili farmers in Indonesia." Paper prepared for the Australian Agricultural and Resource Economics Meetings, 7-10 February 2012, Fremantle, Western Australia.

Schipmann C. and M. Qaim (2011a) "Supply chain differentiation, contract agriculture, and farmers' marketing preferences: The case of sweet pepper in Thailand", *Food Policy*, 36: 667–677.

Schipmann C. and M. Qaim (2011b), "Modern Food Retailers and Traditional Markets in Developing Countries: Comparing Quality, Prices, and Competition Strategies in Thailand", *Applied Economic Perspectives and Policy*, 33(3): 345–362. doi:10.1093/aepp/ppr018.

Sexton R.J. (2013), " Market power, misconceptions, and modern agricultural markets", *American Journal of Agricultural Economics*, 95 (2): 209-219. doi: 10.1093/ajae/aas102.

Torero, M. and A. Gulati (2004), "Connecting small holders to market: role of infrastructure and institutions", USAID Policy brief, International Food Policy Research Institute (IFPRI).

Torero, M. (2011), "A framework for linking small farmers to markets", Paper presented at the IFAD Conference on New Directions for Smallholder Agriculture 24-25 January, 2011.

Traill, W.B. (2006), "The rapid rise of supermarkets?", *Development Policy Review*, 24(2): 163-174.

Weatherspoon, D.D. and T. Reardon, (2003) "The Rise of Supermarkets in Africa: Implications for Agrifood Systems and the Rural Poor", *Development Policy Review,* 21(3): 333-55.

Weinberger, K. and T. Lumpkin (2005), "Horticulture for Poverty Alleviation - The Unfunded Revolution", AVRDC Working Paper No. 15. Available at SSRN: http://ssrn.com/abstract=781784 or http://dx.doi.org/10.2139/ssrn.781784

4.8 Appendix

A4.1 Full result of robustness check Table 4.7

Dependent var.	(1) High cost /m2		(2) High cost /m2		(3) Chili profit		(4) Chili profit	
model	Probit		Treatment 1st stage		OLS		Treatment 2nd stage	
High Cost					0.134	(1.06)	1.008***	(2.64)
Unaware					0.305	(1.61)	0.311**	(2.27)
Aware					0.775***	(3.66)	0.782***	(4.67)
Productivity	0.883***	(9.37)	1.022***	(8.05)	0.449***	(5.98)	0.256**	(2.15)
Grading					0.144**	(2.37)	0.118*	(1.87)
Household size	0.025	(0.57)	0.046	(0.92)	-0.037	(-0.90)	-0.053	(-1.22)
Age	0.004	(0.49)	0.003	(0.36)	-0.014**	(-2.16)	-0.014**	(-2.22)
Education	0.031	(1.35)	0.022	(0.94)	-0.009	(-0.44)	-0.021	(-0.97)
% working age	0.810**	(2.06)	1.297***	(2.84)	0.417	(1.18)	0.136	(0.36)
% retired age	-0.190	(-0.25)	0.233	(0.30)	-0.966	(-1.47)	-1.100*	(-1.88)
Mobile phone	0.121	(0.82)	0.013	(0.08)	0.007	(0.06)	0.036	(0.26)
Motor bike	-0.070	(-0.52)	-0.022	(-0.15)	0.268**	(2.23)	0.260**	(2.17)
Water pump	0.270*	(1.86)	0.274*	(1.77)	0.114	(0.87)	0.048	(0.37)
Mist blower	0.403**	(2.29)	0.342*	(1.66)	0.325**	(2.00)	0.175	(1.10)
Power tiller	0.225	(0.40)	0.162	(0.26)	0.494	(1.02)	0.456	(1.05)
Storage house	-0.105	(-0.69)	0.011	(0.07)	0.178	(1.29)	0.210*	(1.66)
Experience	-0.013	(-1.26)	-0.027**	(-2.25)	0.020**	(2.25)	0.029***	(3.08)
Land owned 05	-0.046**	(-2.28)	-0.032	(-1.47)				
% land rented	0.368*	(1.96)	0.618***	(2.87)				
Chili land					0.484***	(7.75)	0.498***	(7.55)
Production record	0.305	(1.23)	0.330	(1.29)	0.261	(1.21)	0.216	(1.03)
Marketing record	0.272*	(1.71)	0.401**	(2.30)	0.386**	(2.52)	0.240	(1.55)
Distance to road	-0.092	(-0.73)	-0.223	(-1.43)				
Constant	-1.50***	(-3.16)	-1.91***	(-3.54)	10.16***	(16.34)	10.17***	(16.73)
Pseudo /adj. $R2$	0.258				0.348			
N	577		459		562		459	

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$; Wald chi2(21) = 280.82 (prob>chi^2=0.000), independence test = 4.543 (prob>chi^2=0.033); athrho = -0.511(-2.131)

A4.2 Full result of robustness check Table 4.8

Dependent variable	(1) Linear total Cost		(2) Chili income	
Predicted total cost			0.455**	(1.992)
Unaware			0.329*	(1.764)
Aware			0.745***	(3.531)
Productivity	0.249***	(5.679)	0.363***	(3.948)
Grading			0.134**	(2.199)
Distance to road	0.031	(0.375)		
Household size	0.011	(0.364)	-0.046	(-1.127)
Age	-0.005	(-0.976)	-0.013*	(-1.949)
Education	0.015	(0.921)	-0.015	(-0.740)
% working age	-0.154	(-0.560)	0.479	(1.372)
% retired age	-0.265	(-0.515)	-0.823	(-1.252)
Mobile phone	0.263**	(2.552)	-0.093	(-0.650)
Motor bike	0.217**	(2.321)	0.174	(1.381)
Water pump	0.184*	(1.801)	0.016	(0.113)
Mist blower	0.357***	(2.939)	0.097	(0.522)
Power tiller	0.231	(0.631)	0.417	(0.862)
Storage house	0.184*	(1.691)	0.083	(0.577)
Experience	0.015**	(2.140)	0.014	(1.504)
Chili land			0.437***	(6.688)
Production record	0.444**	(2.444)	0.067	(0.283)
Marketing record	0.216*	(1.861)	0.294*	(1.837)
Land owned 05	0.067***	(4.825)		
% land rented	0.634***	(4.827)		
Constant	14.366***	(43.847)	3.959	(1.215)
adj. $R2$	0.234		0.350	

Chapter 5 Empirical studies in geographical economics[34]

5.1 Introduction

Since the seminal work of Krugman (1991) led the way, many researchers have further analyzed and explained the intricate connections between international trade flows, factor mobility, agglomeration, and production, see Brakman, Garretsen, and van Marrewijk (2009) for an overview of the literature. As explained in Brakman and van Marrewijk (forthcoming) there are now three 'core' models of New Economic Geography, or Geographical Economics as we prefer to label it, namely (i) Krugman's model based on labor mobility, (ii) the solvable human capital model based on Forslid and Ottaviano (2003), and the intermediate goods model based on Krugman and Venables (1995). All these models give rise to similar dynamics and core-periphery patterns with path-dependency and multiple long-run equilibria. This chapter focuses on empirical studies in geographical economics. Our contribution is limited to providing an update of the contributions regarding four key features of geographical economics as identified by Head and Mayer (2004a, p. 2616):

- *A large market potential raises local factor prices*. A large market will increase demand for local factors of production and this raises factor rewards. Regions surrounded by or close to regions with high real income (indicating strong spatial demand linkages) will have relatively higher wages.
- *A large market potential induces factor inflows*. Footloose factors of production will be attracted to those markets where firms pay relatively high factor rewards. In the Krugman core model footloose workers move to the region with highest real wage and similarly firms prefer locations with good market access.
- *Reduction in trade costs induces agglomeration*, at least beyond a critical level of transport or trade costs. For a large range of transport costs a change in these costs may not lead to a change in the equilibrium degree of agglomeration, but if a shock moves the

---

[34] This chapter is based on Chang, H., C. van Marrewijk and M. Schramm (forthcoming), "Empirical studies in geographical economics" in Charlie Karlsson and Martin Andersson (eds.), *Handbook of Research Methods and Applications in Economic Geography*, London, UK: Edward Elgar Publishing.

economy beyond its break- or sustain point the economy goes from spreading to agglomeration, or *vice versa* respectively. This also implies that more economic integration (interpreted as a lowering of transport costs) should at some point lead to (more) agglomeration of the footloose activities and factors of production.

- *Shock sensitivity*. Changes in the economic environment can (but need not!) trigger a change in the equilibrium spatial distribution of economic activity. This hypothesis goes to the heart of the idea that geographical economics models are characterized by multiple equilibria.

As we continue it will become clear that empirical research in the field has made some headway. Most empirical work on geographical economics before 2004 has focused on advanced economies. The attention for developing countries has increased considerably since then. There is also clear empirical evidence that market potential (market access) affects income per capita and that changes in freeness of trade affect the spatial distribution of economic activity. However, empirical analysis of core-periphery dynamics using the highly stylized core models of geographical economics with its reliance solely on pecuniary externalities, is also shown to be less fruitful than expected a decade ago. The concluding section discusses these issues and some recent developments.

5.2 Market access and wages

The first line of research focuses on the relationship between spatial wage variation and proximity to (i) consumers and (ii) intermediate input markets. The geographic concentration of economic activity is based on product-market linkages between regions that result from love-of-variety, economics of scale, and transportation cost. The idea can be traced back to Harris's (1954) market-potential function, which states that the demand for goods produced in a location is the sum of purchasing power in nearby locations, weighted by transportation cost. Proximity to the market, as measured by physical distance and other variables, implies lower interaction costs. Firms located in high demand locations are thus able to pay higher nominal wages. Following the familiar setup of core geographical models with Dixit-Stiglitz monopolistic competition, Cobb-Douglas production functions, scale economies, iceberg transportation costs, and intermediates and labor as inputs (see Brakman and van Marrewijk, forthcoming), the market access MA_i and supplier access SA_i for location i is the distance-weighted sum of market capacity m_j and supplier capacity s_j, respectively:

(1) $$MA_i = \sum_j \tau_{ij}^{1-\sigma} m_j = \sum_j \tau_{ij}^{1-\sigma} Y_j P_j^{\sigma-1}$$

(2) $$SA_i = \sum_j \tau_{ij}^{\sigma-1} s_j = \sum_j \tau_{ij}^{\sigma-1} n_j p_{ij}^{\sigma-1}, \quad (\tau_{ii} = 1, \tau_{ij} > 1, i \neq j)$$

Where *i* and *j* are location indices; *Y* is total expenditure; n is the number of varieties; τ_{ij} is the iceberg transportation cost for goods sent from i to j; $\tau_{ij}^{\sigma-1}$ is referred to as the *freeness of trade* between *i* and *j*; p_j is the free on board price of an individual variety; $p_{ij} = p_i \tau_{ij}$ is the delivered price; and *P* is the aggregate price index, which can also be denoted as a function of supplier access:

(3) $$P_j = \left[\sum_i n_i p_{ij}^{1-\sigma}\right]^{1/1-\sigma} = \left[\sum_i n_i (p_i \tau_{ij})_{ij}^{1-\sigma}\right]^{1/1-\sigma} = (SA_j)^{1/1-\sigma}$$

Redding and Venables (2004) derive a structural *wage equation* to provide direct links to market access and supplier access based on mark-up pricing and the zero profit condition using three input sources: (i) a tradable intermediate input with price v (input share γ), (ii) an internationally immobile factor (labor) with price w (input share β), and (iii) the composite intermediate good with price P (share α):

(4) $$\left(w_i^{\beta} v_i^{\gamma} c_i\right)^{\sigma} = \kappa\left(SA_i\right)^{\alpha\sigma/(\sigma-1)}\left(MA_i\right)$$

Where σ is the price elasticity of demand, c_i is a measure of technology differences, and κ is a constant. Empirical estimation of this wage equation requires *two steps*: (i) using the gravity equation to estimate bilateral transportation costs and further predict the market and supplier access of each location and (ii) estimating the wage equation. In all cases, wage differences are strongly associated with market access and supply access, but limited in geographic scope due to the magnitude of the trade cost parameter. Head and Mayer (2011) further theoretically and empirically generalized the Redding and Venables' findings using panel estimation to confirm the robustness of their results.

The geographical economics wage equation can also be estimated *directly* in which case we need to control for human capital and exogenous amenities. The wage equation in this type of empirical studies is based on Helpman's (1998) non-tradable housing sector approach in which intermediate inputs do not play a role and only the effect of market access on wages is explored, i.e. as if imposing $\alpha = \gamma = 0$ and $\beta = 1$ in the cost function. If one assumes real wages are equal across locations, then the wage equation takes the following form:

$$(5) \qquad \log(w_{jt}) = \beta + (1/\sigma)\log\left(\sum_i Y_{it}^{\frac{\sigma(\mu-1)+1}{\mu}} H_{it}^{\frac{(1-\mu)(\sigma-1)}{\mu}} w_{it}^{\frac{(\sigma-1)}{\mu}} e^{-\tau(\sigma-1)d_{ij}}\right) + \eta_{jt}$$

Where μ denotes the share of income spent on manufacturing goods, H_i is the housing stock, d_{ij} is the distance between locations *i* and *j*, and η is the error term. Note that the transportation costs are exponential. To eliminate the region fixed effects the empirical estimation of the wage equation is approached by taking first differences.

The early empirical studies confirm the positive impact of market access on nominal wages but mainly among developed countries.[35] More recent empirical research explores the relationship for developing countries, such as in China[36] (Hering and Poncet, 2010), Brazil (Fally et al., 2010; Paillacar, 2007), Spain (Pires, 2006) and Indonesia (Amiti and Cameron, 2007). Boulhol and De Serres (2010) show that the impact of market access is significantly higher for developing (non-OECD) countries. Most studies are based on two step estimation and include both market access and supplier access. Studying the benefits of agglomeration arising from demand and supplier linkages is particularly interesting in developing countries because industrialization has long been coupled with agglomeration activities and increasing income inequalities in a country. Table 5.1 provides an overview of the results for a selection of studies. The most important parameter for market access is $1/\sigma$, with σ as the elasticity of substitution between traded goods (which should be greater than unity). The coefficients estimated vary substantially between studies. In the studies using the direct-estimation approach estimates for σ range between 4.9 and 7.6, while in studies using the two-step approach estimates vary between 2.9 and 4.9. Studies on cross-country per-capita income

---

[35] Research includes De Bruyn (2003) for Belgium, Mion (2004) for Italy, Brakman et al. (2004) for German districts, Roos (2005) and Kiso (2005) for Japan, and Hanson (2005) for US counties using the direct approach. Knaap (2006) for US counties, Head and Mayer (2006) and Breinlich (2006) for the EU, and Lopez-Rodriguez and Faina (2006) use a two step approach (with the latter using a geographic information system instead of the gravity equation).

[36] Due to available data on trade flows between international and domestic partners at the provincial level, China has been a popular research ground in testing the wage equation, including Ma (2006), Lin (2005), Cui (2006), De Sousa and Poncet (2011), Hering and Poncet (2009, 2010), and Kamal, Lovely, and Ouyang (2011).

differential give lower elasticity estimates as compared to estimates from cross-regions/states wage differential (with average mark-ups of around 1.18 and 1.67).

Table 5.1 Overview of literature on wage inequality and market access since 2004

a. Two stage estimation	Sample	distance parameter gravity equation	Market access $1/\sigma$	σ
Redding and Venables (2004)	World countries,1996	1.74	0.26	3.91
Head and Mayer (2011)	World countries, 1965~2003	* 1.34	0.55	1.82
Knaap (2006)	US states, 1999	0.98	0.23	4.35
Breinlich (2006)	EU regions, 1975~97	* 0.78	0.13	7.75
Head and Mayer (2006)	EU regions, 1985~2000	0.52	0.12	8.51
Hering and Poncet (2009)	Chinese provinces, 1995~2002	1.24	0.16	6.25
Hering and Poncet (2010)	Chinese cities, 1995 [+]	1.53	0.16	6.33
Boulhol and De Serres (2010)	World countries, 1970-2004	0.81~0.99	0.15[♥] 0.248[♠]	6.58[♥] 4.03[♠]
De Sousa and Poncet (2011)	Chinese provinces, 1995~2007	1.21	0.11	9.43
Kamal, Lovely, Ouyang (2011)	China households[+], 1995, 2002	n.a.	(1995) 0.28 (2002) 0.67	3.57 1.49
Fally et al. (2010)	Brazilian states[+], 1999	1.45	0.108	9.26
b. Direct estimation[**]		distance decay parameter	$1/\sigma$	σ
Mion (2004)	Italian provinces, 1991~98	*170-190*	*0.17*	*5.9*
Brakman et al. (2004a)	German regions, 1995	0.2	0.27	3.7
Hanson (2005)	US counties, 1970, 80, 90	1.6~3.2	0.13~0.20	4.9~7.6
Kiso (2005)	Japanese prefectures, 1978-98	300~*1700*	*0.16~0.70*	*1.4~6.1*
Pires (2006)	Spanish regions, 1981-95	2.9	0.19~0.23	4.3~5.2
Fingleton (2006)	British regions[#], 2000	--	0.15	6.6
Niebuhr (2006)	European regions, 1985	11.9	*0.04*	22.5
Amiti and Cameron (2007)	Indonesia districts[+], 1998	28	0.11	9.3

Note: Estimates reported are extracted from each paper with selection based on (i) specification including both market and supplier access, (ii) panel estimation or IV estimation, and (iii) authors' most preferred specification; [+]Use of micro-firm/household data; [*] Average over years of the original estimates; [**] Selected estimates using non-linear least squares; [♥]OECD-countries, [♠]Non-OECD countries; distance decay effect measured per 1000 km; [#] wage equation is modeled as a spatial durbin model (with spatially lagged wage and spatially lagged market access as additional explanatory variables) and to construct market potential distance decay parameter, elasticity of substitution σ is set at 7.8 and 6.25 respectively; *Italic* indicates insignificant estimates.

The distance coefficients from the first-stage gravity equation are provided in upper panel of Table 5.1a. Most distance coefficients are higher than the Disdier and Head (2008) meta-analysis trade flow benchmark, 0.9. Bosker and Garretsen (2010) pay special attention to the specification of the trade cost and estimation strategy (direct or two stage estimation). Their analysis of 22 papers with 262 estimates suggests that the size and significance of the estimated market access coefficient is sensitive to the estimation strategy and choice of trade cost specification.[37] They find the market access coefficient from two stage estimation is generally significant while direct estimation in most cases gives insignificant market access coefficients. Niebuhr (2006) provides evidence that the direct estimation approach suffers from the attempt to stay as close as possible to the core geographical economics model. It pays off to regress a wage equation that is more loosely based on geographical economics.

An important problem in these estimates is to control for worker characteristics and differences in skill intensity. Redding and Schott (2003) show that the impact of market access on the wage premium for skilled workers in central regions is reinforced by human capital accumulation. Research failing to control for human capital accumulation suffers from endogeneity problems and will incorrectly attribute the influence of geographical economics factors to spatial wage differences. Fallah et al. (2011) analyze different skill groups in US metropolitan areas and conclude that better market access increases wages more for high skilled workers than for low skilled workers. Hering and Poncet (2010) find a similar result for Chinese micro-data. In contrast, Fally et al., (2010) do not find this effect when using Brazilian micro-data.

To summarize, most studies show that market (demand) access and supplier (cost) access have a significant positive impact on manufacturing wages. The benefits are localized, however, in view of the high distance coefficient. Agglomeration benefits in Indonesia for example reveal that only 10 percent of the market access and supplier access spread beyond 108 km and 262 km, respectively (Amiti and Cameron, 2007). This mirrors Redding and Venables's (2004, pp. 77-78) simulation results, which suggest that the "*gain from closer integration between low-income developing countries may be relatively small compared to those to be had from closer integration with high-income developed countries.*" In other words, despite the significant effect of market access on wages found in most geographical economics research, the actual influence is heavily discounted by distance. That implies at the sub-national level that it is predominantly local density that determines local factor prices. A familiar result from the field of urban economics. Fingleton (2006) empirically confronts geographical economics with urban economics in an artificial nested model and shows that density rules over market access.

---

[37] Use of the exponential cost function as in (5) makes it less likely to find a significant effect of market access.

5.3 Market access and factor mobility

The second line of research explores the influence of market access on the location choice of workers (looking for high real wages) and firms (looking for high profits). Within a geographical economics framework (transportation costs, returns to scale, and linkages) clustering of firms and workers becomes attractive since agglomeration provides demand access for firms and access to a large product range for workers. A process of cumulative causation then reinforces this attractiveness, potentially counter-balanced by the competition effect of new entrants. Research on factor mobility needs to take into account this feedback mechanism in which the attractiveness of each location is determined by the location decision of different economic agents. The *demand* or *backward linkage* tests whether firms are attracted to locations with large local demand. The *cost* or *forward linkage* examines whether workers are attracted to locations with high real wages. We analyze these linkages in the next two subsections.

5.3.1 Firm location decision — demand / backward linkage

Empirical work on the demand linkage has tested the determinants of multinational's location choice for their (footloose) foreign subsidiaries, focusing on market access and savings in transportation costs. With some exceptions, early work focuses on local demand and not the demand in nearby locations. See Table 5.2 for an overview of recent work. Head and Mayer (2004b) analyze the location choice of Japanese firms investing in Western European countries using a conditional logit model derived from Krugman (1992). Profitability in location i depends mainly on the costs in that location and its market access (including neighbouring countries). When deciding where to locate the multinational firms weigh these issues in a stochastic framework. They find that market access is important: the probability that a location is chosen rises by 3 to 11 percent when the market access variable rises by 10 percent – though it should be noted that a market access variable strictly based on the core geographical economics model is outperformed by a Harris' market potential. Since there are no intermediate inputs terms in the cost function, this study focuses on market access, not supplier access.

Amiti and Javorcik (2008) also incorporate supplier access by taking into account empirical inter-industry linkages. This leads to profits for location k as given in equation (6), where the industry and time fixed effects are denoted by ς . Amiti and Javorcik calculate the equilibrium number of foreign subsidiaries in each location and estimate using non-linear least squares. They find that both market access and supplier access are important factors for determining the number of foreign firms in each Chinese province. One standard deviation increase in market access raises the entry of foreign firms by 13 percent; one standard deviation increase in supplier access raises it by 20 percent. They also show that the *supplier* access effect from other provinces accounts for approximately 16 percent of the total supplier access effect. This is lower than the *market* access effect accounted for by other provinces which is round 32 percent of the total market access effect.

$$\ln \pi_{Ikt} = \beta^{I}(1-\sigma^{I})\ln w_{kt} + \gamma^{I}(1-\sigma^{I})\ln v_{kt} + \alpha_S SA_{Ikt} + \alpha_M MA_{Ikt} + \varsigma_I + \varsigma_t \tag{6}$$

Table 5.2 Overview of firm origin and choice locations studied since 2004

Study	Origin of the investors	Location choice	Year
Head and Mayer (2004b)	Japan	17 EU countries	1984-1995
Lu and Tokunaga (2007,08,09)	Japan (food industry)	8 East Asian locations	1985-2006
Disdier and Mayer (2004)	France	EU and CEECs	1800-1999
Mayer et al. (2010)	France	France and the world	1985-2002
Yamawaki (2006)	US and Japan	7 EU member states	1993
Pusterla and Resmini (2007)	World	4 CEECs	1990-2001
Basile et al. (2008)	World	8 EU countries	1991-1999
Amiti and Javorcik (2008)	World	Chinese provinces	1998-2001

CEEC = Central and Eastern European countries.

The third column of Table 5.2 gives an overview of the location choice analyzed, indicating that recent research focuses more on rapidly developing emerging markets. Pusterla and Resmini (2007) analyze Central and Eastern European countries (CEECs) and find that the location choice is mainly affected by the demand rather than cost factors. Other research finds different strategic decisions for different types of firms. Japanese affiliates, for example, are more export-oriented than American affiliates in the EU, China, and East Asia.[38] This suggests that the strategic decision of Japanese multinationals is relatively more affected by cost linkages (supplier access, see also Lu and Tokunaga, 2007, 2008, 2009). Foreign investors also have a tendency to follow other foreign investors in the same sector.[39] The existence of agglomeration economies in FDI may result from positive externalities such

[38] Basile et al. (2008), Yamawaki (2006), Amiti and Javorcik (2008), and Lipsey (2000).
[39] Wheeler and Mody (1992), Head et al. (1995), Head and Ries (1996), and Yamawaki (2006).

as information sharing, technology spillovers and greater availability of specialized inputs and labor. Research on Japanese firms in the US and Europe reveals that the agglomeration benefits are even larger when proximate plants are operated by other Japanese firms.[40] A similar network effect is found for French firms by Mayer et al. (2010).

5.3.2 Migration decision — cost / forward linkage

The main motivation for migration decisions is the real wage differential between locations. We discuss Crozet's (2004) approach, which combines a geographical economics model with Tabuchi and Thisse's (2002) discrete choice model of migration. There are three sectors: agriculture, services, and manufacturing. The agricultural sector produces a tradable and homogenous product under constant return to scale and perfect competition. It serves as the numéraire and uses a fixed supply of immobile farmers in a region as the sole input. The manufacturing and services sectors operate in a monopolistically competitive setting, producing varieties of products under increasing return to scale. Workers in the manufacturing and services sectors are mobile. The real wage ω_i in location i is the nominal wage w_i corrected for aggregate price indices $P_{mi,t}$ and $P_{si,t}$ of manufacturing and services goods: $\omega_{i,t} = w_{i,t} / (P^{\mu}_{mi,t} P^{\phi}_{si,t})$. Services are not tradable across regions. Migration costs are assumed to increase with the distance between home and the host regions as follows: $[d_{ij}(1+bF_{ij})]^{\lambda}$, where F_{ij} is a dummy variable equal to one if the two regions do not share a common border. When deciding to relocate migrants take the probability of finding a job and the migration costs into consideration, leading to the equation to be estimated:

$$
\begin{aligned}
(7) \qquad \ln\left(\frac{migr_{ji,t}}{\sum_{i' \neq j} migr_{ji',t}}\right) &= \frac{\mu}{\sigma - 1} \ln\left(\sum_{r=1}^{R} L^{m}_{r,t-1} (w_{r,t-1} d^{\delta}_{ij})^{1-\sigma_m}\right) + \\
&+ \alpha_1 \ln\left(L^{s}_{i,t-1}\right) + \alpha_2 \ln\left(\omega_{i,t-1}\right) - \lambda \ln\left(d_{ij}(1 + bF_{ij})\right) + controls
\end{aligned}
$$

Where L^{m}_{r} and L^{s}_{r} denote total manufacturing and services employment in region r, respectively. The first two components on the right hand side represent the attractiveness of region i, namely industrial activity in the (vicinity of the) region (market access) and the

[40] Smith and Florida (1994), Head et al. (1995, 1999), O'Huallachain and Reid (1997), Head and Mayer (2004b), Belderbos and Carree (2002), and Belderbos et al. (2000).

availability of services varieties. The third term reflects the expected wage in the region and the fourth term mobility costs. A list of control variables is added, such as a time trend and the size of a location. Crozet uses bilateral migration data from five European countries to support this model, see Table 5.3. The elasticity of substitution (σ_s) is lower for services (ranging from 1.41 for Italy to 1.93 for the Netherlands, for $\phi = 0.4$) than for manufactures (ranging from 1.3 in UK to 4.3 in the Netherlands). The transportation cost coefficients $\delta(1\text{-}\sigma_{\mathrm{m}})$ are significantly different between countries, where the low coefficients in UK and Spain suggest that workers are more sensitive to the market access differential than workers in the Netherlands, Italy and Germany.

Table 5.3 Comparable estimates of migration choice based on Crozet 2004 model.

Study	Location	year	σ_{m}	δ	α_1	α_2	λ	b
Crozet (2004)	Netherlands	1980-90	4.32	1.42	0.46	-0.45	1.02	0.51
	Italy		3.58	3.55	0.97	-0.06	0.31	9.04
	Germany		3.74	3.62	0.72	*-0.08*	0.92	0.86
	Spain		1.53	0.46	0.90	-0.39	0.76	1.41
	UK		1.30	1.54	0.73	-0.21	0.48	1.27
Pons et al. (2007)	Spain regions	1920s	2.81	1.79	0.82	2.06	1.76	-0.82
Paluzie et al. (2009)	Spain regions	1920-30	2.81	1.79	0.82	2.06	1.76	-0.82
		1960-70	3.29	1.98	0.97	2.23	1.05	-0.93
		2000-04	1.76	0.89	0.90	0.22	0.85	-1.21

Parameters extracted from each corresponding paper with $\mu = 0.4$ imposed. Numbers in *italic* are insignificant estimates. Coefficients are estimates based on equation 9, for σ_m as elasticity of substitution, manufacturing sector; δ as elasticity of trade to distance; α_1 as influence of local service supply ($\alpha_1 = \phi/(\sigma_s - 1)$); α_2 as influence expected wage; λ as distance elasticity of migration cost and *b* as influence of borders on migration.

The framework developed by Crozet is later adopted by Kancs (2005, 2011) to explain the migration flows between the Baltic States, Poncet (2006) for migration dynamics in China, and by Pons et al. (2007) and Paluzie et al. (2009) for migration flows in Spain. Kancs finds that mobility in the EU is low even after the EU enlargement, such that core-periphery patterns are less likely. The studies on Spain find that the forward linkages in the large industrial centers are largely offset by the high elasticity of migration costs. Paluzie et al. find a decreasing effect of distance and an increasing magnitude of the border effect over the last century. Interestingly, the attractiveness of industrial wages (α_2) has been decreasing over time while the attractiveness of the services sector (α_1) has been rising, indicating that the services sector has become increasingly important for explaining migration decision.

5.4 Freeness of trade and the degree of agglomeration

Geographical economics models are well equipped to analyze the impact of economic integration on the degree of agglomeration within a country. Economic integration increases the freeness of trade[41], the change in the freeness of trade affects market access and supplier access and thereby the location choice for firms and workers. This section investigates how the degree of agglomeration is affected by reductions in impediments to interregional trade in general (4.1), trade liberalization (4.2), and improvements in transport infrastructure (4.3). We highlight the empirical literature that is either based on calibration of a geographical economics model or determines the structural parameters of the model (producing counterfactual distributions of economic activity using simulations).

5.4.1 Impact of economic integration in general

An important result of core geographical economics models is that a reduction in trade costs, or equivalently an increase in the freeness of trade, affects the geographical distribution of economic activity. Models with a weak spreading force result in a *Tomahawk* relationship between freeness of trade and the degree of agglomeration (see Brakman and van Marrewijk, forthcoming). If freeness of trade is low spreading is the only stable equilibrium, while if freeness of trade is high agglomeration is the only stable equilibrium. In models with a stronger spreading force (Helpman, 1998; Puga, 1999; Tabuchi, 1998; Tabuchi and Thisse, 2002) the result is a *Bell-shaped* relationship between freeness of trade and the degree of agglomeration. The spreading equilibrium is stable at low *and* high freeness of trade levels. At intermediate levels there is at first a rising and then a declining trend towards agglomeration. These results are, however, based on models with a stylized geographic structure. Distance (and thereby freeness of trade) between any pair of regions is the same between all regions (equidistant regions). Economic integration thus leads to a uniform increase in freeness of trade between all pairs of regions. Beyond the three-region setting on the two-dimensional service of the earth, there is no simple geographic structure to substantiate equidistant regions. In other words, for four or more regions the equidistant world is an exclusively theoretical construction.

---

[41] Freeness of trade between *i* and *j* (φ_{ij}) is $\tau_{ij}^{1-\sigma}$.

To analyze economic integration within the European Union Bosker et al. (2010) *add geography* to a stylized geographical economics model. They analyze a regional-wage equation using panel data to estimate the structural parameters without imposing real wage equalization across regions. Instead, a region's manufacturing price index *P* is simplified by assuming two regions only: the region itself and all other regions. The iceberg transport cost function in the wage equation allows for economies of distance and a (country) border effect:

(8) $$\tau_{ij} = d_{ij}^{\delta}(1 + bB_{ij})$$

where d_{ij} is the great-circle distance between i and j; δ is the distance-decay parameter; and B_{ij} is a dummy variable equal to zero if regions i and j are in the same country and equal to one otherwise. The sample consists of 183 EU regions in the period 1992-2000. Three important parameters are directly estimated: the elasticity of substitution σ, the distance-decay parameter δ and the border effect b. The parameters σ and δ are statistically and economically significant with point estimates 7.122 and 0.102, respectively. The distance decay is relatively small compared to other empirical studies (see Table 5.1)[42].

For the simulations the authors ignore the services sector, use great-circle distances between the capital cities of any pair of regions, and use the actual distribution of employment and arable land as the initial distribution. They then show the impact of increasing economic integration on the equilibrium degree of agglomeration by varying the value of δ (lower δ indicates a higher degree of economic integration). Basic results from the geographical economics literature are confirmed. With perfect interregional mobility the simulations lead to complete agglomeration at the current level of integration ($\delta = 0.102$) and at higher levels. Perhaps not too surprising as the services sector is not included, leading to strong agglomeration forces. Without interregional labor mobility the current degree of agglomeration is higher than the counterfactual. The limitations of these exercises are clear. We should include the important services sector, the appropriate spatial level of analysis may

[42] Also when comparing the estimates with distance coefficients found in gravity equations explaining bilateral international trade flows, the distance decay is relatively small for interregional trade within the EU, in which $(1-\sigma)\delta = -0.624$.

not be the regional but the urban or local level[43], and amenities and spatial sorting of skills need to be taken into account. Clearly, more elements from urban economics need to be incorporated into geographical economics models.

5.4.2 Impact of trade liberalization

To analyze the impact of trade liberalization within a geographical economics model the literature answers three main questions as summarized in Table 5.4, namely (i) does liberalization lead to a pull of economic activity towards the border? (border effects), (ii) does it affect regional specialization?, and (iii) does it lead to convergence / divergence? The empirical literature as surveyed in Brülhart (2011) mainly focuses on the first question and finds significant border effects, showing that trade liberalization leads to a pull towards regions with an easy access to foreign markets. The study by Sanguinetti and Volpe Martincus (2009) on Argentina is the exception as it finds a positive employment gradient from Buenos Aires in the industries that face less tariff protection.

Regarding the second question: Faber (2007) shows that distance to the US border matters for regional specialization in Mexico, leading to higher specialization for comparative advantage industries closer to the border and higher specialization for comparative disadvantage industries away from the border. Volpe Martincus (2010) finds that in Brazil the nearer the region is to the Buenos Aires mainport the more specialized it is in industries with a high degree of openness, in contrast to the findings for Argentina by Sanguinetti and Volpe Martincus (2009).

Regarding the third question: the evidence points to a negative relationship between trade liberalization and regional convergence. In both Mexico and China regional GDP per capita diverged as trade openness increased. See, however, the World Bank (2009) for the opposite long-term effect. We discuss two studies in more detail to provide further insight, namely Redding and Sturm (2008) on the German division and Brülhart, Carrère and Trionfetti (2011) on the fall of the Iron Curtain. Redding and Sturm (2008) examine the impact of the East-West division in Germany on city growth in West Germany. The division after World

[43] Arzaghi and Henderson (2008) analyze location choice for start-ups in the advertising industry in Manhattan and find no benefits from information sharing anymore if the distance between the networking firms exceeds 750 metres.

War II led to a sharp decline in intra-German trade and a relatively large fall in market access for West German cities close to the border, especially for small border cities. The authors use a *difference-in-differences* approach to test differences in population growth. The pre-treatment period is from 1919 to 1939. The (division) treatment period is from 1950 to 1988. Twenty West German cities that are located within 75 km from the East-West German border belong to the treatment group. The other 99 cities are the control group. The following equation is estimated:

(9) $$PopGrowth_{ct} = \beta_1 treatGroup_c + \beta_2 (treatGroup_c * division_t) + t + \varepsilon_{ct}$$

where $PopGrowth_{ct}$ is the annualized growth rate of population in city *c* in period *t*, $treatGroup_c$ is a dummy variable equal to one if city *c* is a border city, $division_t$ is a dummy variable equal to one if period *t* is within the range 1950-1988, and t is a time dummy.

The results confirm the market-access hypothesis. Before World War II there is no statistically significant difference in the growth rate for the treatment group and the control group. After World War II, however, the annualized growth rate is significantly lower (3/4th percentage points) for the treatment group of cities close to the East-West German border than for the control group of other cities, particularly for small cities. The authors also perform simulations based on a geographical economics model to replicate their findings. Nakajima (2008) uses the same approach and finds similar results for Japanese cities relatively close to former colony Korea after the division between Japan and Korea after the Second World War. Brakman et al. (2012) use a similar procedure (without the simulations) to show that cities close to a border affected by EU integration experience a rise in population growth (for a period of about 30 years).

Brülhart et al. (2011) examine the impact of the Fall of the Iron Curtain in 1989 on the increase in wages and employment in municipalities in Austria. The relatively large increase in market access of municipalities close to the border with former communist countries (Czech Republic, Slovakia, Slovenia, and Hungary) is expected to lead to relatively high annual growth rates of wages and employment for these municipalities. The authors also use a difference-in-differences approach, where the pre-treatment period is 1975: I – 1989: IV and the treatment period is 1990: I – 2002: IV. The treatment group consists of the municipalities located within 25 km of the nearest border crossing with the four former communist countries. The control group consists of the other Austrian municipalities.

*Table 5.4 Overview of country studies on trade liberalization and internal geography**

4.a Boarder	*Country*	*Boarder effect?*	*Year (s)*
Sanguinetti and Martincus (2009)	Argentina	No, industries facing less protection locate further away from main port	1985, 1994
Brüllhart et al. (2011)	Austria	Yes, after the Fall of the Iron Curtain growth of employment and wage is the highest in regions bordering the Czech Republic, Slovakia, Slovenia and Hungary.	1975-2002
Henderson and Kuncoro (1996)	Indonesia	Yes, proximity to main ports and other metro areas becomes more important for location choice new manufacturing plants	1980-1985
Nakajima (2008)	Japan	Yes, lower growth of Japanese cities near to Korea after division Japan and Korea after Second World War	1925-1985
Hanson (1997)	Mexico	No, wage gradient from US border or Mexico city is unaffected	1965,1970, 1975,1980, 1985,1988
Hanson (1998)	Mexico	Yes, employment gradient from US border becomes negative	1980,1985, 1993
Chiquiar (2008)	Mexico	Yes, increase in real wage of unskilled the highest in regions specialized in manufacturing and receiving the highest share of FDI	1990-2000
Pernia and Quising (2003)	Philippines	Yes, growth of GDP pc the highest in regions with highest export ratio (special economic zones) and mainport	1988-2000
Redding and Sturm (2008)	W-Germany	Yes, division of Germany makes border cities grow less	1919-1988
Brakman et al. (2012)	EU	Yes, EU integration effect raises growth of border cities; not enough to counter negative general effect	1979-2010
4.b Specialization	*Country*	*Regional specialization?*	*Year (s)*
Sanguinetti & Volpe Martincus (2009)	Argentina	Yes, industries facing less protection locate further away from mainport	1985, 1994
Volpe Martincus (2010)	Brazil	Yes, industries with higher degree of openness locate nearer to border	1990-1998
Faber (2007)	Mexico	*Negative* employment gradient from US border for *comparative-advantage* industries, *positive* employment gradient from US border for *comparative-disadvantage* industries	1993,1998, 2003
4.c Convergence	Country	Convergence or divergence?	Year(s)
Kanbur and Zhang (2005)	China	Trade openness leads to an increase in regional inequality.	1952-2000
Chiquiar (2005)	Mexico	Divergence, beta coefficient turns positive in regional-growth regressions for the period 1985-2001	1970-2001
Rodriguez-Pose and Sanchez-Reaza (2005)	Mexico	Divergence in the period with the highest degree of openness	1980,1985, 1993,2000
Gonzalez-Rivas (2007)	Mexico	Divergence, trade openness affects beta coefficient in regional- growth regression positively	1940-2000

*Selection of studies largely based on Brülhart (2011, pp. 76-78).

Brülhart et al. (2011)'s findings support the hypotheses: the growth rate of the median wage for the treatment group is 0.267 percent higher than for the control group and the employment growth rate is 0.861 percent higher for the treatment group than for the control group. The authors subsequently have difficulty in replicating their estimates based on a geographical economics model, requiring in particular implausibly high income shares spent on housing (40 to 50 percent). They can resolve this issue for plausible parameter values if they extend the baseline model with taste heterogeneity (Tabuchi and Thisse, 2002) which introduces locational preferences (sentimental attachment to location) for individuals. Empirical analysis based on geographical economics models should thus take imperfect interregional mobility into account, especially when analyzing economies in which labor markets are relatively rigid.

5.4.3 Impact of improvements in transport infrastructure

Roberts et al. (2010) analyze the impact of the construction of the National Expressway Network (NEN) in China, designed to connect all cities with a population size of over 200,000 and to reduce regional inequality, using a geographical economics model. They first calculate the travel time between each pair of locations (Chinese prefectures) with and without the NEN using geo-referenced road information for China and information about the average speed on each type of road.[44] These travel times are used to determine the iceberg transport costs between any pair of regions.

The next step is to estimate a wage equation for the year 2007 that allows for regional differences in productivity and interregional productivity spillovers. Regional productivity is determined by region-specific observables like investment per worker and human capital and the spatial lag of investment per worker and human capital, and also by a region-specific stochastic unobservable. Combining this with guesstimates for other parameters to calculate regional manufacturing price indices they determine each region's market access.[45] They then use these estimates to re-calculate each region's manufacturing price index and market

---

[44] The types of roads are city street, local road, motorway, national highway, provincial highway, expressway, and whether they paved or unpaved.

[45] When calculating manufacturing price indices they assume productivity differences to be non-existent, which leads to a measurement error when calculating market access and for that reason market access needs to be instrumented.

access and simulate their model to arrive at a short-run equilibrium of the year 2007 with NEN. They also simulate a counterfactual short-run equilibrium of the year 2007 with iceberg transport costs based on the minimum travel times without the infrastructural improvements. The difference between these two short-run equilibria is the impact NEN has. They find that NEN has increased income levels in China, with the largest wage increases concentrated in the richer eastern part of China. As regional inequality measured by the standard deviation of regional income is not affected, the infrastructure improvements did not lead to income convergence.

5.5 Shock sensitivity and path dependency

A final prominent feature of geographical economics models is the existence of multiple equilibria. A large shock can then permanently move an economy from its initial equilibrium to a new equilibrium. Temporary shocks may have permanent effects. History matters. The literature uses quasi-experiments with large and temporary exogenous shocks to test if the economy will move back to the initial situation (mean-reverting process) or not.

Davis and Weinstein (2002) analyze the impact of US bombing during World War II on the population of Japanese cities. They assume that there is an initial stable equilibrium and that the actual population share may deviate from this equilibrium: $\log(s_{it}) = \Omega_i + \varepsilon_{it}$, where s_{it} is city *i*'s share in total population in period *t*, Ω_i is its inherent size and ε_{it} is the deviation from initial equilibrium of city *i* in period t. The deviation in period *t* depends on the deviation in the past period *t-1* and a shock v : $\varepsilon_{it} = \rho\varepsilon_{it-1} + v_{it}$, where parameter ρ is the autoregressive parameter representing the rate at which a shock dissipates over time. Consider the war shock to occur in period t. It will affect the relative change in population:

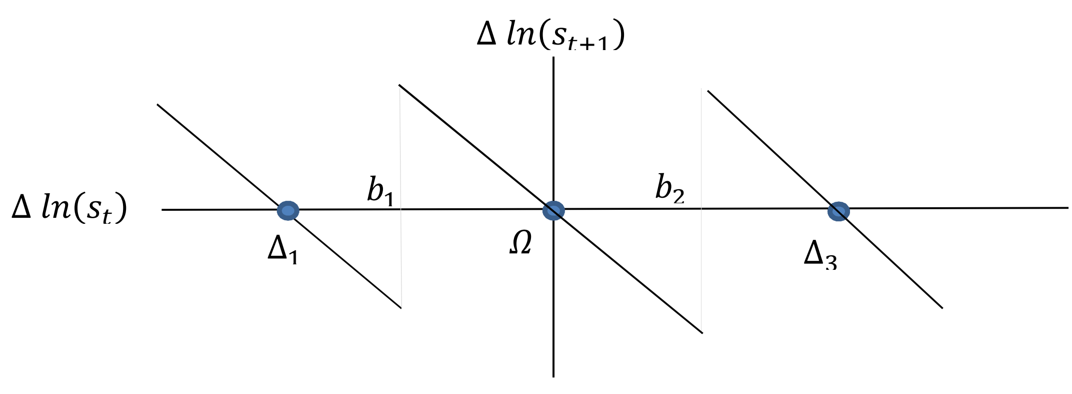
$$(10) \qquad \Delta\log(s_{it+1}) = (\rho - 1)v_{it} + (\rho - 1)\rho\varepsilon_{it-1} + v_{it+1}.$$

If ρ=1, then *s* follows a random walk and the war shock (and all other shocks) have permanent effects. If ρ=0 the war shock dissipates completely in the post-war period and there is a mean-reverting process. So, the test for the existence of a unique equilibrium or multiple equilibria is to obtain an estimate of parameter ρ. The authors find a mean-reverting process which points to the existence of a unique equilibrium of population distribution in Japan.

In contrast, Brakman, Garretsen and Schramm (2004) using a similar approach for the bombing of German cities in World War II find that the war shock there has permanent effects (multiple equilibria). Miguel and Roland (2011) analyze the impact of US bombing on Vietnam in the period 1965-1975. As they do not have data about the change in population at the local or regional level they focus on the impact on the current state of development for Vietnamese districts instead. They do not find a long-term impact of bombing intensity on local development indicators, with the exception of access to electricity (where heavily bombed districts have currently better access to electricity). Since the more heavily bombed districts received more government investment per capita in the post-war period their findings may be caused by government intervention.

Davis and Weinstein (2008) design a different test for the existence of a unique equilibrium or multiple equilibria. As illustrated in Figure 5.1, the approach provides a test for the existence of a unique equilibrium or multiple equilibria by estimating the threshold values of the basin of attraction of the initial equilibrium (b_1 and b_2 in the figure) and potentially other equilibria (Δ_1 and Δ_2 in the figure). If a negative war shock pushes $\log(s_{it})$ below $\Omega_i + b_1$ then city *i* moves to a new lower equilibrium in the post-war period. In the case of a positive shock pushing $\log(s_{it})$ above $\Omega_i + b_2$, city *i* moves to a new higher equilibrium in the post-war period. So Davis and Weinstein impose that each city will be in an equilibrium in the post-war period and that the basin of attraction of the initial equilibrium is identical for each city and for each of the other equilibria. Using threshold regressions they are able to falsify the existence of multiple equilibria in Japan, even at the city-industry level.

Figure 5.1 Two-period growth representation of a model with three stable equilibria



Source: Brakman, Garretsen, and van Marrewijk (2009).

The main problems with the Davis and Weinstein (2002, 2008) approach are that the point estimates are sensitive to the choice of the post-war period, the analysis is about the average city and not about the individual cities, and the approach is a static cross-section regression. Bosker et al. (2008) deal with these problems using data on 62 West German cities in the period 1925-1999 by testing for a unit root in the city's population share and estimating equation (20) by applying an Augmented Dickey Fuller test for each city.

$$(11) \qquad \Delta \log(s_{it}) = \psi_i + \zeta_i \log(s_{it-1}) + \sum_k \Delta \log(s_{it-k}) + \eta_{it}$$

Where ψ_i is a city-specific trend, the past relative changes in s_i are included to control for potential autocorrelation. The critical parameter is ζ_i. If it is significantly negative then s_i will be stationary and temporary shocks have no permanent effects. If it is zero then there is a unit root in s_i and temporary shocks have permanent effects. The authors also allow for a one-time deterministic shift in ψ_i to be decided by the data. They find that the temporary shock of WW II had permanent effects, which is evidence in favour of multiple equilibria and shock sensitivity.

Finally, Redding, Sturm and Wolf (2011) analyse the impact of the division of Germany for air travel, a network industry. They argue that the relocation of the airport hub in Germany from Berlin to Frankfurt is conclusive evidence for the existence of multiple equilibria. The argument runs as follows: in the airline industry operating a connection requires fixed costs and profitability of a connection depends on the number of passengers. The fundamentals of an attractive location for an airport hub are therefore local population size, economic activity, and bilateral distances to other locations. The creation of a new airport hub requires substantial sunk costs. These switching costs may prevent the airline industry to choose a different location for an airport hub. Only if the difference between the present value of profits at the new location and the present value of profits at the current location of the airport hub exceeds the sunk costs a switch to a new location will be made. So the higher the sunk costs of creating a new airport hub, the less important fundamentals become and the higher the scope for multiple equilibria.

The division of Germany isolated West Berlin which became unattractive as an airport hub. In the years immediately after World War II the US military chose the airport of Frankfurt as its main European air transportation terminal. During the Soviet blockade of West Berlin in

1948/49 the airport of Frankfurt became the main airport for the airlift to Berlin. This made Frankfurt a profitable location for the airline industry, which thus became the airport hub by chance and not because of superior fundamentals.[46] A cost-benefit analysis of relocating the airport hub from Frankfurt to Dusseldorf leads to a net present value gain of less than €1 billion, which is certainly not enough to warrant a relocation. The location of the airport hub in Germany is thus clearly a case of lock-in, path dependence, and multiple equilibria.

5.6 Concluding remarks

Our literature review, which focuses on the developments since 2004, in general finds strong evidence in support of four key implications of geographical economics models.

(i) *Market potential and factor prices*. Most studies show that market (demand) access and supplier (cost) access have a significant positive impact on manufacturing wages. The benefits are localized, however, in view of the high distance coefficient.

(ii) *Market potential and factor flows*. Market access is an important determinant of firm location decisions (demand / backward linkage), with extra agglomeration benefits for network industries. Similarly, market access and the availability of local services play an important role in migration decisions (cost / forward linkage).

(iii) *Trade costs and agglomeration*. The conclusions of the geographical economics theory literature can be replicated in a multi-region model that incorporates more complex geographical features. The empirical literature finds significant border effects, indicating that market access should play an important role in trade liberalization decisions. Trade liberalization also influences regional specialization patterns and in some cases seems to lead to regional divergence. A study on China suggests that infrastructure investments lead to higher income levels but not to income convergence within and between regions.

(iv) *Shock sensitivity*. With some notable exceptions, most of the literature investigating the complex questions on multiple equilibria, shock sensitivity, and path dependence do indeed find evidence for long-term economic effects of large exogenous shocks, such that history plays an important role in understanding the current economic playing field.

[46] The paper provides calculations to show that the differences in market access of Frankfurt compared to other locations is minimal while Frankfurt does not have the largest local market.

Taken together, these findings indicate that empirical research in the field has made some headway. There is clear empirical evidence that market potential (market access) affects income per capita and that changes in freeness of trade affect the spatial distribution of economic activity. The literature using quasi-natural experiments that reveals the impact of sudden changes in market access (e.g. through liberalization, division, or unification) also points to the importance of market access in the spatial distribution of economic activity. However, empirical analysis of core-periphery dynamics using the highly stylized core models of geographical economics with its reliance solely on pecuniary externalities, is also shown to be less fruitful than expected a decade ago.

Highly localized agglomeration rents, for example, suggests that an urban economics approach may be more valid. Moreover, sticking to highly stylized models with Dixit-Stiglitz monopolistic competition, iceberg transport costs, Cobb-Douglas or linear production functions and homogeneous firms, and perfect labour mobility may lead to inferior empirical results. One also needs to be aware of the sensitivity of the results to specific transport costs functions. The lesson we can draw from this is to relax some of the main assumptions and allow for spatial equilibria with partial agglomeration by, for example, introducing taste heterogeneity. We can also use a more eclectic approach by allowing for other types of agglomeration externalities through knowledge spillovers, networking between firms, sorting of skills across space and spatial selection of firms.

The fact that model outcomes can be replicated in a model featuring more than two non-equidistant regions has limited practical implications. To replicate the spatial distribution of economic activity and its evolution over time one should also relax the spatial equilibrium condition of real-wage equalization and focus more attention to non-pecuniary externalities. Especially in the services sector, the location decisions of firms seem to be driven by networking opportunities and knowledge spillovers. We should, therefore, include the important services sector and the appropriate spatial level of analysis may not be the regional but the urban or local level. Amenities and spatial sorting of skills also need to be taken into account. Clearly, more elements from urban economics need to be incorporated into geographical economics models. Brakman and van Marrewijk (2013), for example, recently showed that the ‘lumpy’ distribution of factors of production does not appear to affect international trade flows if the analysis is at the regional level, while it *does* have an impact if the analysis is at the urban level.

These suggestions are strengthened by the differences between nations and the increasing number of studies in developing and emerging countries (China). For Central and Eastern European countries the location choice is mainly affected by demand- rather than cost factors. Japanese affiliates are more export-oriented than American affiliates in the EU, China, and East Asia, suggesting stronger cost linkages. Foreign investors have a tendency to follow other foreign investors in the same sector. Japanese firms experience larger agglomeration benefits when proximate plants are operated by other Japanese firms. A similar network effect is found for French firms. The low worker mobility in the EU makes core-periphery patterns less likely, even though workers in the UK and Spain seem to be more sensitive to the market access differential than workers in the Netherlands, Italy and Germany. The attractiveness of industrial wages has been decreasing over time while the attractiveness of the services sector has been rising. Again, we need to incorporate urban aspects in geographical economics models, focus more attention on services sectors and networks, while allowing for (firm-, consumer-, and taste-) heterogeneity.

5.7 References

Amiti, M. and L. Cameron (2007), ‘Economic geography and wages’, *Review of Economics and Statistics*, **89**(1), 15-29.

Amiti, M. and B. Javorcik (2008), ‘Trade costs and location of foreign firms in China’, *Journal of Development Economics*, **85** (1–2), 129–149.

Arzaghi, M. and J. Henderson (2008), ‘Networking off Madison Avenue’, *Review of Economic Studies*, 75, 1011-1038.

Basile, R., D. Castellani and A. Zanfei, (2008), ‘Location choices of multinational firms in Europe: the role of EU cohesion policy’, *Journal of International Economics*, **74** (2), 328–340.

Belderbos, R.A., G. Capannelli and K. Fukao (2000), ‘Local procurement by Japanese electronics firms in Asia’, in Ito T. and A. Krueger (eds.), *The Role of Foreign Direct Investment in Economic Development*, Chicago University Press/NBER, Chicago.

Belderbos, R.A. and M. Carree (2002), ‘The location of Japanese investment in China: Agglomeration effects, Keiretsu, and firm heterogeneity’, *Journal of the Japanese and International Economics*, **16**, 194-211.

Bosker, M., S. Brakman, H. Garretsen and M. Schramm (2008), ‘A century of shocks: the evolution of the German city size distribution 1925-1999’, *Regional Science and Urban Economics, **38***, 330-347.

Bosker, M., S. Brakman, H. Garretsen and M. Schramm (2010), ‘Adding geography to the new economic geography: bridging the gap between theory and empirics’, *Journal of Economic Geography, **10***, 793-823.

Bosker, M. and H. Garretsen (2010), ‘Trade costs in empirical New Economic Geography’, *Papers in Regional Science,* **89** (3), 485-511.

Boulhol, H. and A. de Serres (2010), 'Have developed countries escaped the curse of distance', *Journal of Economic Geography*, **10**, 113-139.

Brakman, S., H. Garretsen and M. Schramm (2004a), 'The spatial distribution of wages: estimating the Helpman-Hanson model for Germany', *Journal of Regional Science*, **44**, 437-466.

Brakman, S., H. Garretsen and M. Schramm (2004b), 'The Strategic Bombing of German Cities during WWII and its Impact on City Growth', *Journal of Economic Geography*, **4**(2), 201-218.

Brakman, S., H. Garretsen, and C. van Marrewijk (2009), *The new introduction to geographical economics*, Cambridge University Press, Cambridge, U.K.

Brakman, S., H. Garretsen, C. van Marrewijk, and A. Oumer (2012), "The border population effects of EU integration," *Journal of Regional Science*, **52**(1), 40-59.

Brakman, S., and C. van Marrewijk (2013), 'Lumpy countries, urbanization, and trade', *Journal of International Economics* **89**, 252-261.

Brakman, S., and C. van Marrewijk (forthcoming), "Factor prices and geographical economics," in: C. Karlsson and M. Andersson (eds.) Chapter 28 in *Handbook of Research Methods and Applications in Economic Geography*, Edward Elgar.

Breinlich, H. (2006), 'The spatial income structure in the European Union - What Role for Economic Geography?', *Journal of Economic Geography*, **6**(5), 593-617.

Brülhart, M. (2011), 'The spatial effects of trade openness: a survey', *Review of World Economics,* **147**, 59-83.

Brülhart, M., C. Carrère and F. Trionfetti (2011), '*How wages and employment adjust to trade liberalization: quasi-experimental evidence from Austria'* (GREQAM Documet de Travail 2011-33). Marseille: Groupement de Recherche en Economie Quantitative d'Aix-Marseille.

Brülhart, M., M. Crozet and P. Koenig (2004), 'Enlargement and the EU periphery: the impact of changing market potential', *World Economy,* **27**, 853-875.

Chiquiar, D. (2005), 'Why Mexico's regional income convergence broke down', *Journal of Development Economics,* **77**, 257-275.

Chiquiar, D. (2008), 'Globalization, regional wage differentials and the Stolper-Samuelson theorem: evidence from Mexico', *Journal of International Economics,* **74**, 70-93.

Crozet, M. (2004), 'Do migrants follow market potentials?', *Journal of Economic Geography*, **4**(4), 439-458.

Cui, F. (2006), 'Geographic location and regional income inequality In China', mimeo London School of Economics.

Davis, D. and D. Weinstein (2002), 'Bones, bombs, and break points: the geography of economic activity', *American Economic Review,* **92**, 1269-1289.

Davis, D. and D. Weinstein (2008), 'A search for multiple equilibria in urban industrial structure', *Journal of Regional Science*, **48**, 29-65.

De Bruyne, K. (2003), 'The location of economic activity: is there a spatial employment structure in Belgium?', Mimeo Katholiek University of Leuven.

De Sousa, J. and S. Poncet (2011), 'How are wages set in Beijing?', *Regional Science and Urban Economics*, **41**, 9-19.

Disdier, A.C, and T. Mayer (2004), 'How different is Eastern Europe? Structure and determinants of location choices by French firms in Eastern and Western Europe', *Journal of Comparative Economics* **32**(2), 280–296.

Disdier, A.C. and K. Head (2008), 'The puzzling persistence of the distance effect on bilateral trade', *Review of Economics and Statistics*, **90,** 37-41.

Faber, B. (2007), 'Towards the spatial patterns of sectoral adjustments to trade liberalisation: the case of NAFTA in Mexico', *Growth and Change,* **38**, 567-594.

Fally, T., R. Paillacar and C. Terra (2010), 'Economic geography and wages in Brazil: Evidence from micro-data', *Journal of Development Economics*, **91**, 155-168.

Fallah, B., M.D. Partridge and M.R. Olfert (2011), 'New Economic Geography and U.S. metropolitan wage inequality', *Journal of Economic Geography*, **11**, 865-895.

Fingleton, B. (2006), 'The new economic geography versus urban economics: an evaluation using local wage rates in Great Britain', *Oxford Economic Papers*, **58**, 501-530.

Forslid, R. and G.I.P. Ottaviano (2003), 'An Analytically Solvable Core-periphery model', *Journal of Economic Geography*, **3**, 229-240.

González Rivas, M. (2007), 'The effects of trade openness on regional inequality in Mexico', *Annals of Regional Science,* **41**, 545-561.

Hanson, G. (1997), 'Increasing returns, trade and the regional structure of wages', *Economic Journal*, **107**, 113-133.

Hanson, G. (1998), 'Regional adjustment to trade liberalization', *Regional Science and Urban Economics,* **28**, 419-444.

Hanson, G. (2005), 'Market potential, increasing returns and geographic concentration', *Journal of International Economics*, **67**(1), 1-24.

Harris, C. (1954), 'The market as a factor in the localization of industry in the United States', *Annals of the Association of American Geographers,* **64**, 315–348.

Head, K. and J. Ries (1996), 'Inter-city competition for foreign investment: Static and dynamic effects of China's incentive areas', *Journal of Urban Economics,* **40**, 38–60.

Head, K., J. Ries and D. Swenson (1995), 'Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States', *Journal of International Economics,* **38**, 223–247.

Head, K., J. Ries and D. Swenson (1999), 'Attracting Foreign Manufacturing: Investment Promotion and Agglomeration', *Regional Science and Urban Economics,* **29**(2), 197–218.

Head, K. and T. Mayer (2004a), Empirics of Agglomeration and Trade, in Vernon Henderson and Jacques-Francois Thisse (eds.), *Handbook of Regional and Urban Economics*, Amsterdam: North Holland.

Head, K. and T. Mayer (2004b), 'The market potential and the location of Japanese investment in the European Union', *The Review of Economics and Statistics,* **86**(4), 959–972.

Head, K. and T. Mayer (2006), 'Regional wage and employment responses to market potential in the EU', *Regional Science and Urban Economics*, **36**(5), 573–95.

Head, K. and T. Mayer (2011), 'Gravity, market potential and economic development', *Journal of Economic Geography,* **11**, 281–294.

Henderson, J. and A. Kuncoro (1996), 'Industrial centralization in Indonesia', *World Bank Economic Review,* **10**, 513-540.

Hering, L. and S. Poncet (2009), 'The impact of economic geography on wages: Disentangling the channels of influence', *China Economic Review*, **20**(1), 1-14.

Hering, L. and S. Poncet (2010), 'Market access and individual wages: Evidence from China', *Review of Economics and Statistics*, **92**, 145-159.

Helpman, E. (1997), 'The size of regions', in David Pines, Efraim Sadka and Itzhak Zilcha (eds.), *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge: Cambridge University Press, pp. 33-54.

Kanbur, R. and X. Zhang (2005), 'Fifty years of regional inequality in China: a journey through central planning, reform and openness', *Review of Development Economics,* **9**, 87-106.

Kancs, D. (2005), 'Can we use NEG models to predict migration flows? An example of CEE accession countries', *Migration Letters*, **2**(1), 32-63.

Kancs, D. (2011), 'The economic geography of labour migration: competition, competitiveness and development', *Applied Geography*, **31**, 191-200.

Kamal, F., M.E. Lovely and P. Ouyang (2011), 'Deeper Global Integration and the Effect of Market Access on Wages in China', *International Review of Economics and Finance*, doi:10.1016/j.iref.2011.10.006

Kiso, T. (2005), 'Does new economic geography explain the spatial distribution of wages in Japan?' Mimeo University of Tokyo.

Knaap, T. (2006), 'Trade, location, and wages in the United States', *Regional Science and Urban Economics*, **36**(5), 595-612.

Krugman, P. (1991), 'Increasing Returns and Economic Geography', *Journal of Political Economy*, **99**(3), 483-49.

Krugman, P. (1992), 'A dynamic spatial model', NBER Working Paper #4219.

Krugman, P. R. and A. J. Venables (1995), 'Globalization and the Inequality of Nations', *Quarterly Journal of Economics*, **110**, 857-880.

Lin, S. (2005), 'Geographic Location, Trade and Income Inequality in China', in R. Kabur and A. J. Venables Spatial (eds.) *Inequality and Development*, Oxford University Press, London.

Lipsey, R.E. (2000), 'Affiliates of U.S. and Japanese Multinationals in East Asian Production and Trade', in T. Ito and A. O. Krueger (edt.) *The Role of Foreign Direct Investment in East Asian Development*, Chicago, Illinois: University of Chicago Press.

Lopez-Rodriguez, J. and J.A Faina (2006), 'Does distance matter for determining regional income in the European Union? An approach through the market potential concept', *Applied Economics Letters,* **13**(6), 385-390.

Lu, J. and S. Tokunaga (2007), ‘Market potential, agglomeration effects and the location of Japanese food industry in East Asia’, in *Annual Meeting of Agricultural Economics Society of Japan*, pp. 390–395.

Lu, J. and S. Tokunaga (2008), ‘Market potential and location choice for the Japanese food industry in East Asia: An approach of New Economic Geography’, *Studies in Regional Science* **38**(1), 109-119.

Lu, J. and S. Tokunaga (2009), ‘Supplier access and the location of Japanese food industry FDI in East Asia’, *Letters in Spatial and Resource Sciences* **2**(1), 1-10.

Ma, Alyson C. (2006), ‘Geographical Location of Foreign Direct Investment and Wage Inequality in China’, *World Economy,* **29**:8, 1031-1055.

Mayer (2009?)

Mayer, T., I. Mejean, B. Nefussi (2010), ‘The location of domestic and foreign production affiliates by French multinational firms’, *Journal of Urban Economics,* **68**, 115-128.

Miguel, E. and G. Roland (2011), ‘The long-run impact of bombing Vietnam’, *Journal of Development Economics, 96*, 1-15.

Mion, G. (2004), ‘Spatial externalities and empirical analysis: The case of Italy’, *Journal of Urban Economics*, **56**, 97–118.

Nakajima, K. (2008), ‘Economic division and spatial relocation: the case of postwar Japan’, *Journal of the Japanese and International Economies*, **22**, 383-400.

Niebuhr, A. (2006), ‘Market access and regional disparities/ New economic geography in Europe’, *Annals of Regional Science*, **40**, 313-334.

O’Huallachain, B. and N. Reid (1997), ‘Acquisition versus Greenfield investment: The location and growth of Japanese manufacturers in the United States’, *Regional. Studies,* **31**(4), 403–416.

Paillacar, R. (2007), ‘Market potential and worker heterogeneity as determinants of Brazilian wages’, Mimeo, Paris School of Economic.

Paluzie, E., J. Pons, J. Silvestre and D.A. Tirado (2009), ‘Migrants and market potential in Spain over the twentieth century: a test of the new economic geography’, *Spanish Economic Review*, **11**(4), 243-265.

Pernia, E.M. and P.F. Quising (2003), ‘Trade Openness and Regional Development in a Developing Country’, *Annals of Regional Science,* **37**, 391-406.

Pires, A.J.G. (2006), ‘Estimating Krugman's Economic Geography Model for the Spanish Regions’, *Spanish Economic Review,* **8**(2), 83-112.

Poncet S. (2006), ‘Provincial migration dynamics in China: boarders, costs and economic motivations’, *Regional Science and Urban Economics* **36**, 385-398.

Pons, J., E. Paluzie, J. Silvestre and D.A. Tirado (2007), ‘Testing the new economic geography: migrations and industrial agglomerations in Spain’, *Journal of Regional Science*, **47**, 289-313.

Puga, D. (1999), ‘The rise and fall of regional inequalities’, *European Economic Review,* **43**, 303-334.

Pusterla, F. and L. Resmini (2007), Where Do Foreign Firms Locate in Transition Countries? An Empirical Investigation’, *Annals of Regional Science*, **41** (4), 835–856.

Redding, S. and P.K. Schott (2003), 'Distance, skill deepening and development: will peripheral countries ever get rich?', *Journal of Development Economics,* **72**, 515-541.

Redding, S. and D. Sturm (2008), 'The costs of remoteness: evidence from German division and reunification', *American Economic Review,* **98**, 1766-1797.

Redding, S., D. Sturm and N. Wolf (2011), 'History and industry location: evidence from German airports', *Review of Economics and Statistics, **93***, 814-831.

Redding, S. and A.J. Venables (2004), 'Economic geography and international inequality', *Journal of International Economics*, **62**, 53–82.

Roberts, M., U. Deichmann, B. Fingleton and T. Shi (2010), 'On the road to prosperity? The economic geography of China's national expressway network', World Bank Policy Research Working Paper 5479, Washington, DC: World Bank.

Rodríguez-Pose, A. and J. Sánchez-Reaza (2005), 'Economic polarization through trade: trade liberalization and regional growth in Mexico', in Ravi Kanbur and Anthony Venables (eds.), *Spatial inequality and development*, Oxford University Press, Oxford, pp. 237-259.

Roos, M.W.M. (2005), 'How important is geography for agglomeration?', *Journal of Economic Geography*, **5,** 605–620.

Sanguinetti, P. and C. Volpe Martincus (2009), 'Tariffs and manufacturing location in Argentina', *Regional Science and Urban Economics,* **39**, 155-167.

Smith, D. and R. Florida (1994), 'Agglomeration and industry location: An econometric analysis of Japanese-affiliated manufacturing establishments in automotive-related industries', *Journal of Urban Economics*, **36**, 23–41.

Tabuchi, T. (1998), 'Urban agglomeration and dispersion: a synthesis of Alonso and Krugman', *Journal of Urban Economics*, 44, 333-551.

Tabuchi, T. and J.F. Thisse (2002), 'Taste heterogeneity, labour mobility and economic geography', *Journal of Development Economics*, **69**, 155–177.

Volpe Martincus, C. (2010), 'Spatial effects of trade policy: evidence from Brazil', *Journal of Regional Science,* **50**, 541-569.

Wheeler, D. and A. Mody (1992), 'International investment location decision', *Journal of International Economics*, **33**, 57–76.

World Bank (2009), *World Development Report 2009: Reshaping Economic Geography*, Washington.

Yamawaki, H. (2006), 'The Location of American and Japanese Multinationals in Europe,' *International Economics and Economic Policy*, **3**, 157-173.

Chapter 6 Gravity Models of Trade-based Money Laundering*

6.1 Introduction

Money laundering aims at disguising the illicit origin of money and other assets. The Chicago gangster Al Capone transformed illegal alcohol proceeds during times of the American prohibition into revenues from his laundry business, which gave the disguising process its name. Money laundering can take many forms: transferring criminal proceeds from one bank account to another repeatedly over the globe until its origin is untraceable, pretending to have won money through gambling, or using fake export and import bills to ship criminal proceeds outside or into a country.

The 9/11 terrorist attack in 2001 has enforced the US's sensitivity to illegal money flows, as authorities presume that such money transfers support international terrorist and criminal activities. Today financial transactions between the US and the rest of the world are closely monitored: banks have to report suspicious transactions and have to fulfill customer due diligence rules. However, Unger and den Hertog (2011) claim that similar to water which always finds its way, criminals also find new ways to escape anti-money laundering regulation. Zdanowicz (2004b) sees trade as one important 'backdoor' for launderers, who can use fake invoicing of exports and imports and other forms of trade-based money laundering (henceforth, TBML) to disguise illicit money flows. According to Zdanowicz, these new forms of abusing trade for laundering have not yet been addressed properly by the regulatory authorities.

TBML is a relatively unknown form of crime that is used to let illicit money pass borders unnoticed. The Financial Action Task Force (FATF, 2006) concludes that 'trade-based money laundering represents an important channel of criminal activity and, given the growth of world trade, an increasingly important money laundering and terrorist financing vulnerability. Moreover, as the standards applied to other money laundering techniques become increasingly effective, the use of trade-based money laundering can be expected to

---

* This chapter is based on Ferwerda, J., M. Kattenberg, H. Chang, B. Unger, L. Groot and J. Bikker (2013), "Gravity model of Trade-base money laundering", *Applied Economics*, 45(22): 3170-8182.
All references to the Appendix in this chapter can be found in the on-line published version of this paper.

become increasingly attractive'. As it is a quite recently discovered form of crime, not much research on TBML is yet available. Zdanowicz (2004a) estimates trade-related dirty money flows to and from the US, which amounts to about one fifth of its trade. He has created the only available dataset revealing an indication of the volume of this type of criminal activity.[47] To illustrate the dataset of Zdanowicz, Table 6.1 shows the top 10 countries in terms of incoming and outgoing TBML, in absolute and relative size, with data from 2004.

Baker (2005) makes clear that laundering is not only a US concern, but especially detrimental for developing countries. According to his findings for every dollar sent to developing countries for development aid, ten dollars flow back to developed countries in the form of money laundering and illicit capital flight. That trade can aggravate poverty in developing countries is an issue heavily discussed in the literature (see Bené *et al.,* 2010). Still lacking in this debate is the extent to which trade can be abused by criminals and tax evaders for money laundering purposes.

Table 6.1 Top 10 countries with trade-based money laundering from and to the US

	TBML out of US		TBML into US		TBML out of US		TBML into US	
	in billions of US dollars				*in percentage of trade*			
1	Canada	18.3	Japan	25.6	Azerbaijan	269	Barbados	26
2	Japan	14.1	Germany	25.5	Serbia & Montenegro	72	San Marino	25
3	China	13.8	Canada	21.6	Iran	70	Italy	24
4	Mexico	13.0	China	20.3	Cuba	62	Germany	23
5	Germany	11.8	UK	16.2	Kazakhstan	35	Monaco	23
6	UK	10.1	Mexico	14.7	Bulgaria	34	Switzerland	23
7	S Korea	7.2	France	9.6	Estonia	32	UK	20
8	France	5.5	Italy	9.4	Denmark	31	Austria	19
9	Taiwan	4.8	S. Korea	7.7	Barbados	26	France	18
10	Singapore	4.3	Taiwan	6.4	Philippines	24	Portugal	17

Source: dataset of Zdanowicz. The two lists in the left-hand panel show that the major trading partners of the US also have the most TBML with the US, but the two lists on the right hand reveal that in the rankings in relative terms other countries show up. The exports to Azerbaijan are so much undervalued, that the illegal money flow outnumbers the total trade data reported by a factor 2.7. Barbados is the only country that shows up in the top 10 of both outgoing and incoming TBML countries (measured as percentage of trade).

Since the introduction of anti-money laundering policy, high demand exists for estimates of money laundering to justify the burden that is put on public and private entities in charge of

[47] Lately, his method also has been applied by the Joint Research Center of the European Commission in Italy, to study TBML in Austria, Belgium and the Netherlands, but the results have not been published. Detailed data for individual Dutch transactions, similar to those of Zdanowicz, are not passed on for scientific research, because the Dutch customs claims that they would violate privacy protected under the Dutch Law.

chasing dirty money. Walker (1995), Schneider (2008), Unger (2007), and Bagella *et al.* (2009) estimate the amount of money laundering based on an economic or econometric model. Walker (1995) was the first to propose a prototype model to estimate money laundering worldwide. His so-called Walker model is based on the well-known gravity model, which is quite popular in trade economics (Walker and Unger, 2009; Brakman and Van Bergeijk, 2010). This gravity-style model describes the geographical allocation of proceeds of crime, which need to be laundered to cover their criminal origin. The share of proceeds transferred from country A to country B depends on the ’attractiveness’ of country B and the distance between the two countries. Unger (2007) revived the model for the Netherlands, by updating and refining the distance indicator. However, due to lack of observations of money laundering, the weights or parameters of the attractiveness factors in Walker’s model were never based on statistical estimates, but only plugged in with values of an inspirational guess. Although the outcome of the prototype model seems reliable when compared with other estimations (Walker and Unger, 2009), the actual specification of the model was never tested. So, the question that is still open is whether a gravity type equation can properly explain bilateral money laundering flows.

The outline of this chapter is as follows. Section 6.2 introduces the traditional gravity model and explains how different specifications on this model are used to estimate world-wide money laundering flows. Section 6.3 describes Zdanowicz’s dataset on TBML, investigates the relationship between these TBML flows and licit trade and introduces the variables used in the literature to predict money laundering. Section 6.4 tests the relevance of the prototype models of Walker and Unger for TBML. Section 6.5 presents our adapted gravity model to explain TBML. The final section draws conclusions.

6.2 Gravity model

6.2.1 The traditional gravity model

For over decades the gravity model has been successfully applied to flows of the most widely varying types, such as migration, buyers distributed across shopping centers, recreational traffic, commuting, patient flows to hospitals and interregional as well as international trade. The model is inspired by Newton’s universal law of gravity, which states that the attraction between two objects depends on the mass of these objects and (the inverse of) their squared mutual distance, apart from a constant. The model specifies that a flow from origin i to destination j is determined by supply conditions at the origin, by demand conditions at the destination and by stimulating or restraining forces relating to the specific flow between i and j. In a context of international trade the traditional gravity model usually has the following form:

$$(1) \qquad X_{i,j} = \beta_0 \; Y_i^{\beta_1} \; N_i^{\beta_2} \; Y_j^{\beta_3} \; N_j^{\beta_4} \; D_{i,j}^{\beta_5} \; P_{i,j}^{\beta_6}$$

where $X_{i,j}$ is the value of trade between countries i and j, Y_k is the Gross Domestic Product (*GDP*) of country k, N_k is the size of the population of country k, and $D_{i,j}$ and $P_{i,j}$ denote the distance between countries i and j and a possible special preference relationship, respectively.[48] The gravity model of bilateral trade has become the workhorse of applied international economics (Eichengreen and Irwin, 1998) and has been used in a number of contexts.[49] Some authors assume that the size of the population has no impact, thus $\beta_2 = \beta_4 = 0$, which renders the resemblance to Newton's Law of Gravity even more obvious.[50]

The empirical results obtained with the model have generally been judged as very good. Deardorff (1998) argues that the model is sensible, intuitive and hard to avoid as a reduced theoretical model to explain bilateral trade. Yet the gravity model has some theoretical imperfections. One is the absence of a cogent derivation of the model, based on economic theory. Several authors have tried to provide the model with such a theoretical basis, using models of imperfect competition and product differentiation, notably Anderson (1979), Bergstrand (1985), Helpman and Krugman (1985), Bikker (1987, 2010), and Anderson and Van Wincoop (2003), whereas Deardorff (1998) proves that the model is also consistent with the Heckscher-Ohlin trade theory under perfect competition.

However, none of these derivations generate the gravity model exactly as formulated in Equation (1).[51] This equation could only be approximated under a number of restrictive and unrealistic assumptions, as has been made clear by Bergstrand (1985). This is due to the

[48] For a complete overview of all the abbreviations , see Appendix 4.
[49] Linders (2006) finds 200 studies (actually a sample from a much larger set), and provides a selection in his Table 3.1. For an overview, see Anderson and Van Wincoop (2004).
[50] E.g. Tinbergen (1962), Pöyhönen (1963a, 1963b), Pulliainen (1963), and Bergstrand (1985).
[51] The most restrictive theoretical model of Anderson, Bergstrand, as well as Helpman and Krugman, is a gravity model with only GDPs as determinants. A less restrictive model has a different functional form (Anderson, Equation (16) or additional determinants (Bergstrand, Equation (14)).

absence of substitution in the basic gravity model of Equation (1). Substitution can be made plausible by an example from economic integration: the accession of Estonia to the European Union (EU) in 2004 has led to additional imports by other EU countries of wood, wood products and paper (the major export products of Estonia) – that is, gross trade creation. However, EU imports of wood products from other countries declined. This decline – trade diversion – is not described by the basic gravity model. Bergstrand (1985), Bikker (1987, 2010), Anderson and Van Wincoop (2003) and Redding and Scott (2003) extend the basic model with a substitution structure.

As TBML is closely related to trade flows, this paper will use the successful basic gravity model – Equation (1), extended with additional explanatory variables – for TBML flows stemming from the US.

6.2.2 Gravity model for money laundering

In earlier years, Walker (1995) already recognizes that the gravity model may be useful in order to explain laundered money flows between countries. His prototype model for money laundering assumes that the share of proceeds from crime generated in country *i* and sent to country *j* depends on both the mass and 'attractiveness' of country *j*, and the distance between the two countries:

$$(2) \qquad F_{ij} \Big/ \sum_i F_{i,j} = ((GNP_j / Population_j) * Attractiveness_j) \Big/ (\text{Distance})^2_{i,j}$$

The amount of money laundering flowing from country *i* to country *j* is denoted as F_{ij}. In this model, the flows are expressed into shares of countries *j* in the total outflows of country *i* (by dividing the flows by $\sum_i F_{ij}$, the total amount of money to be laundered in country *i*). If we compare this equation to Equation (1), we see that Mass of *j* is represented by $(GNP_j / Population_j) * Attractiveness_j$ and Mass of *i* by $\sum_i F_{ij}$ [52]. By the way, the interpretation of Equation (2) as gravity model is not from Walker himself but from Unger (2007). The first mass factor is per capita income, based on Gross National Product (*GNP*).

[52] Note that Mass of i is not exactly equal to ΣiFij.

The second is attractiveness, which is put forward as the sum of a number of weighted factors contributing to the quality of country *j* for money laundering:

$$Attractiveness_j = 3BS_j + GA_j + Swift_j - 3CF_j - CR_j + 15$$

where *BS* is banking secrecy and *GA* government attitude, *Swift* indicates countries with financial institutions that are member of Swift, *CF* refers to conflict and *CR* stands for corruption.[53] This equation became well known in the money laundering field as the Walker equation and has been used to estimate money laundering flows (see *e.g.* Walker, 1995, and Walker, 1999). The underlying idea is that since the gravity model can predict trade flows and many other flows so well, it may also be able to predict money laundering flows. Fundamental critiques on Walker's model concern its *ad hoc* nature and the fact that it is not empirically testable, in fact because data on money laundering flows F_{ij} are lacking.

Unger (2007) modifies the Walker equation by using the distance between the countries in the attractiveness factor, instead of its square. Empirical findings are that most gravity equations for trade come up with an estimate for the coefficient of distance of around -0.9 (Helliwell, 2000). Unger (2007) also redefines physical distance by including three 'cultural' factors: sharing a common language, having colonial ties or being major trading partners of each other. Moreover, in the attractiveness index she includes the 'membership of the Egmont group', a cooperation of national Financial Intelligence Units fighting money laundering, and 'financial sector size' (measured as deposits), as proxies for the fact that extended financial markets make it easier to launder criminal proceeds.[54]

[53] Note that corruption has a negative impact on laundering. Walker assumes that criminals do not like (excessively) corrupt countries, because corruption increases costs of laundering due to necessary side payments and bribes. On the other hand, a very low corruption level might make it difficult to find facilitators for laundering, increasing the transaction costs of laundering. The corruption–laundering literature is ambiguous about the sign. Chaikin and Sharman (2009) give an overview over the various theoretical links between corruption and money laundering. Dreher and Schneider (2010) find this ambiguity for the shadow economy also empirically: corruption reduces the shadow economy in high-income countries, but increases it in low-income countries.

[54] Further, she replaces GNP by GDP.

6.3 The data

In order to apply the gravity model to TBML flows, we use data for 199 countries, stemming from four different sources. The first dataset is from CEPII (www.cepii.fr), a French research center in international economics, which provides a database with determinants for gravity equations of trade flows. These variables are *Border dummy* (with value one when two countries share a border), *Common language dummy* (one when two countries have the same language), *Colony dummy* (one when one of the two countries was a colony of the other) and *Distance* (measured in kilometers 'as the crow flies' between the countries' economically most important cities). The second database on factors determining money laundering has been collected by John Walker for the estimation of his model (Walker, 1999). The attractiveness variables are *Corruption* (a simplified scale based on the Transparency international index of 1996, meaning that countries with a low score have a low corruption level), *Swift member* (a dummy variable with 1 for countries with financial institutions that are member), *Government attitude* (a score for the attitude of a government towards money laundering, going up with tolerance), *Conflict* (a score for a country's conflict status, with 1 for peace and 4 for a heavy conflict) and *Bank secrecy* (the score goes up with secrecy).[55] Tables A1-A2 in Appendix 1 provide greater detail.

The third database is constituted by the variables defined in Unger (2007). New attractiveness variables are *Egmont Group,* a dummy with value 1 for members, and *Financial sector size*, which measures the relative size of the financial sector as Financial Deposits (demand, time and saving deposits) as a share of GDP. Moreover, she replaces the physical distance for a measure of cultural and physical distance.[56] The other variables are roughly identical to those of Walker. Roughly, because the data have been constructed and for some variables scaled slightly differently, see Table A2 in Appendix 1 for details.[57] The

[55] The first two variables are from public sources, while the last three variables are from an unpublished report on potential international security threats to Australia, provided by the Australian Office of Strategic Crime Assessments.

[56] For a description of how this cultural and physical distance is constructed, see Unger (2007), page 78.

[57] We use Walker's definitions for the test of his model and for our traditional gravity and Unger's definitions for the test of her model.

fourth dataset consists of the 2004 estimates on TBML flows from the US by Zdanowicz. His dataset also provides the trade data (imports and exports) for the US in millions of dollars.[58]

6.3.1 Identifying trade-based money laundering

Zdanowicz (2009b) has scrutinized the US Merchandise trade database for suspicious transactions. It contains all import and export transactions data of the US with the rest of the world starting in 2004. He uses ten-digit product codes,[59] hence very detailed product specifications, and assumes that international trade transactions involving abnormally high or low prices, indicate money laundering[60] (for the thresholds, see Zdanowicz, 2009b). Examples are bottles of ketchup for 50 dollars, a football for 3000 dollars, and a Gucci watch worth US$ 100,000 dollar noted as a US$ 50 Swatch on the import bill. Table 6.2 presents aggregated estimates of TBML by origin.

In 2004, an amount of US$ 175 billion entered the US by means of undervalued imports, see Table 6.2. In that case, profits are generated in the US (that is, white money has been created) by selling expensive but fraudulently low-priced imported goods in the US. US$ 48 billion flew into the US by means of overvalued exports. In that case, profits are generated in the US by selling cheap but fakely high-priced US goods to foreign clients. US$ 56 billion left the US by means of overvalued imports and 112 billion dollar by undervalued exports. In these cases profits are similarly generated outside the US, respectively by importing overvalued goods and exporting undervalued goods. In total TBML amounts to 391 billion US$, or 17 percent of total US trade with other countries in 2004.

TBML flows may also be symmetric. Two partners in crime who both want to launder their criminal money, may set up a TBML *swap*, see the first row of Table 6.2. They sent

---

[58] Tables A1 and Table A2 in Appendix 1 provides descriptive statistics of these variables.
[59] Transactions are classified under approximately 8,000 different products. Every item that is exported is assigned a unique 10-digit identification code. The Harmonization Code System (HS-Code) is a system of progressively more specific identifiers for a commodity. For example, concentrated frozen apple juice is assigned a 10-digit identifier. This number is an aggregate of a series of codes starting with a broad category assigned a 2-digit identifier described as Preparations of Vegetables, Fruit, Nuts, etc. It is then assigned a 4-digit identifier described as fruit juices and vegetable juices, etc. The 6-digit identifier is described as apple juice. Information from http://exim.indiamart.com/product-classification/ on May 9, 2011
[60] These abnormal prices can also indicate typographical errors or transfer pricing by multinational firms. The typographical errors will not affect our results as long as they are randomly distributed, since we correct for the size of trade in our final estimation model. Illegal transfer pricing may be included in the TBML estimates of Zdanowicz.

undervalued trading goods to each other, both incurring effort and cost but also obtaining laundered money. An additional advantage may be that they both have profits as well as (tax-deductible) losses, so that they can avoid tax levies on their 'white' profits. An alternative swap consists of exchanging overvalued goods (swap type 2 in Table 6.2). This swap is even less bothersome as no need exists to sell (undervalued) goods and the overvalued goods may be disposed. Since these swaps do not transfer money internationally, they are less suited for sending terrorist money. Swaps may also be executed within one country. For international remittance of funds a single TBML transaction will do, or a (cross) diagonal combination of transactions in Table 6.2.

Table 6.2 Trade-based money laundering by origin in 2004 (billion US$)

	Imports into the US	Exports from the US	Combination
Undervalued	175 TBML into US	112 TBML out of US	Swap type 1
Overvalued	56 TBML out of US	48 TBML into US	Swap type 2

Source: own calculation based on Zdanowicz (2004a).

6.3.2 Relationship between licit trade flows and TBML

Our TBML dataset allows us to investigate the relationship between licit trade flows and TBML. We refer to licit trade when we consider trade as reported in the usual trade statistics. Of course, these flows are polluted by TBML where that occurs. The basic question here is whether TBML is more or less proportional to trade flows, or whether it is distinct from licit trade. We use the Lorenz curve, manifoldly used to measure income inequality, to visualize the relationship between trade and TBML. The calculations behind the Lorenz curves are explained in Appendix 2.

The diagonal in the Lorenz curve diagram reflects TBML which is strictly proportional to licit trade, which is the case if the respective ratios are unity. The values for the ratios determine the positions in the Lorenz curve diagrams: countries with a value below 1 will be located on the bottom-left hand. If the ratio is 1, the country will be at the point of the Lorenz curve where it runs parallel with the diagonal (so with a slope of 1). Finally, countries with ratios higher than 1 will be located on the upper-right hand corner. Figure 6.1 to Figure 6.4 ranks countries according to the ratio of their share of undervalued exports divided by their share of total exports. The horizontal axis measures the cumulative share of exports, while the vertical axis measures the cumulative share of TBML by means of undervalued exports. Canada with a share of 23 percent in US exports and a 9.9 percent share of total US TBML outflow through undervalued exports has a ratio of 0.43, hence relatively little TBML. Japan and Italy take intermediate positions. At the upper-far right end we find the Philippines, with a ratio of 3.52, and Denmark (not shown) with a ratio of 4.95 (see Table A5 in Appendix 1).

Figure 6.1 Lorenz curve of TBML out of US through undervalued exports versus exports

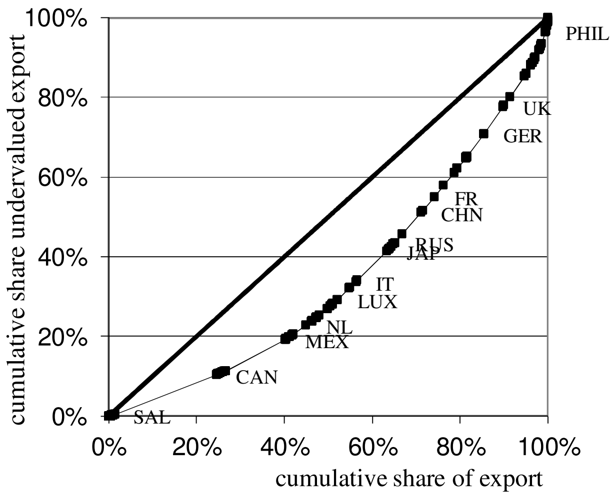


Figure 6.2 Lorenz curve of TBML out of US through overvalued imports versus imports

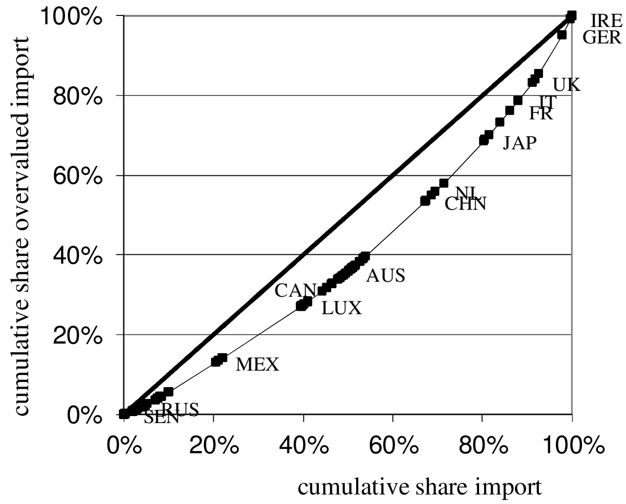


Figure 6.3 Lorenz curve of TBML into US through undervalued imports versus imports

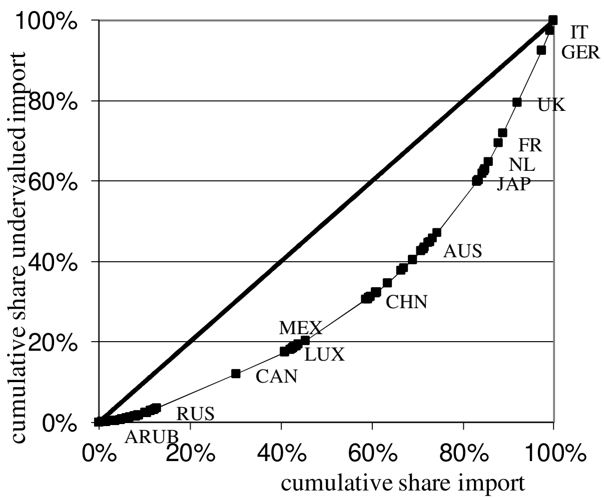
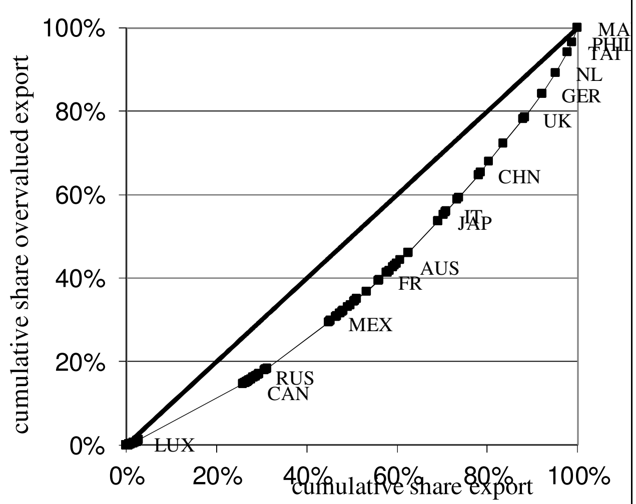


Figure 6.4 Lorenz curve of TBML into US through overvalued exports versus exports



Source: Own calculation based on Zdanowicz (2004a).

The Lorenz curves show that UK, Germany, France, Italy and also Malaysia always have over-proportional shares, while Canada and Mexico always have under-proportional shares of trade-based money laundering. Other countries have sometimes over proportional and sometimes under-proportional shares. The overall picture that emerges is that TBML flows seem to be strongly related to trade flows. This is shown by the rather ‘flat bellies’ of all the Lorenz curves. What we see is that the actual Lorenz curves, especially for overvalued imports (Figure 6.2, upper right panel) and overvalued exports (Figure 6.4, lower right panel), are close to the diagonal.

To conclude, this result suggests that overall trade and TBML are strongly related. Given the absence of clear theoretical principles explaining TBML flows and the fact that in the past licit trade flows have been successfully explained by gravity models, the next sections will investigate the explanatory power of the gravity model for TBML. If we go into greater detail, we observe that countries with a dubious reputation such as Barbados, Liberia, Monaco, Philippines and Switzerland have over-proportional TBML by a factor of at least three, indicating that TBML is not just a ‘normal’ part of licit trade (see Table A5, the FATF black list and the OECD grey list).

6.4. Testing the Walker and Unger model for TBML

6.4.1 Testing the Walker Model for TBML

We apply the prototype model proposed by Walker (1995) on TBML to test whether his specification is supported by our data. Therefore, we replace in Equation (2) the share of country *j* for money laundering from country *i* ($F_{ij} \big/ \sum_i F_{i,j}$) by the share of country *j* for TBML from the US ($TBML_{ij} \big/ \sum_j^n TBML_{ij}$):[61]

$$(3) \qquad \frac{TBML_{ij}}{\sum_j^n TBML_{ij}} = \frac{GDP_j}{Population_j} \frac{(\beta_1 BS_j + \beta_2 GA_j + \beta_3 Swift_j + \beta_4 CF_j + \beta_5 CR_j + \beta_6)}{D_{ij}^{\beta_7}}$$

[61] Note that our application of this Walker model differs from the standard gravity model for trade in the sense that we have data of flows to or from the US only, while normally flows from many origin countries to many destination countries are considered. Since in the Walker model only the characteristics of the destination country *j* matter, we use only the flows of TBML from the US to all other countries in the world. We cannot use the TBML flows to the US, because these flows all have the same destination country *j* (the US) and therefore have no variation in the model of Equation 3.

Table 6.3 presents the estimation results of Equation (3). Since the model is nonlinear in its parameters, we use the Nonlinear Least Squares (NLS) estimation procedure. Starting values for the model parameters, as required in the NLS's iterative numerical estimation procedure, stem from Walker (1995). Column 1 gives the initial Walker values as in Equation (2), while Column 2 gives the estimated NLS coefficients, which differ greatly from the Walker values for money laundering, where, for some variables, even the sign changed. The only coefficient that is significantly different from zero is the coefficient of distance (D). With an estimated value of 0.614 (significant at the 1% level), this coefficient is however not close to the value of 2, as suggested by Walker (and as in Newton's gravity law). The model is able to explain 43% of the variation in the share of total TBML across countries.[62]

Table 6.3 Testing the Walker (1995) model using estimates of TBML flows out of the US (2004)

	Walker's values for ML	NLS estimates of Equation (3) for TBML			
	Parameters α_i	*Coefficients β_i*	*Standard deviation*	*t-test H_0: $\beta_i=0$*	*t-test H_0: $\beta_i = \alpha_i$*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Bank secrecy	3	-0.022	0.015	-1.51	-201.47
Government attitude	1	-0.037	0.024	-1.52	-43.21
Swift membership	1	0.072	0.077	0.95	-12.05
Conflict level	-3	0.110	0.090	1.23	34.56
Corruption level	-1	-0.015	0.013	-1.20	75.77
Distance to US (D_{ij})	2	0.614^{***}	0.079	7.78	-17.54
Constant	15	0.052	0.074	0.71	-202
Observations		199			
Adjusted R^2		0.427			

Note: *** indicates significance at the 1% level.

To test whether the estimates differ significantly from the suggested values in Walker (1995), we assume that the errors are distributed normally and use the coefficient standard errors in Column 3 to compute *t*-values of the estimates in Column 5. Column 5 tests for each model variable the null hypothesis that the true beta coefficient is equal to the Walker values (denoted by α_i). For each of the model parameters, we reject the Walker value. We conclude

[62] A caveat is that Walker specifies his model for money laundering while we test the model using data of only one type of money laundering.

here the Walker (1995) model is unable to explain the estimates of TBML from the US to all other countries in the world in a satisfactory manner.[63]

6.4.2 Testing the Unger model for TBML

As a next step we use our dataset to evaluate the Unger model, a modification of Walker's model for money laundering as suggested by Unger (2007), which includes cultural factors for the constructed measure of cultural and physical distance (*CPD*) and two more variables for the attractiveness: Membership of the Egmont group and the size of the financial sector (FD). The equation that will be estimated is specified as:

$$TBML_{ij} \Big/ \sum_j^n TBML_{ij} = (GDP_j / Population_j) * (\beta_1 BS_j + \beta_2 GA_j + \beta_3 Swift_j + \beta_4 CF_j + \beta_5 CR_j + \beta_6 Egmont_j + \beta_7 FD_j + \beta_8) \Big/ CPD_{ij}^{\beta_9} \qquad (4)$$

Table 6.4 Testing the Unger model (2007) applied to trade-based money laundering (2004)

	Unger's values for ML	NLS estimates of Equation (4) for TBML			
	Parameters α_i	*Coefficients β_i*	*Standard deviation*	*t-test H_0: β_i=0*	*t-test H_0: $\beta_i = \alpha_i$*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Bank secrecy	3	-0.056	0.039	-1.45	-78.36
Government attitude	1	-0.082	0.054	-1.51	-20.04
Swift membership	1	0.014	0.117	0.12	-8.43
Conflict level	-3	0.021	0.019	1.10	159.00
Corruption level	-1	0.023	0.025	0.92	40.92
Egmont group	1	0.074	0.065	1.14	-14.25
Financial deposits	1	-0.002	0.009	-0.28	-111.33
Physical and cultural distance	1	2.848***	0.327	8.71	5.66
Constant	10	0.161	0.147	1.09	-66.93
Observations		199			
Adjusted R^2		0.428			

Note: *** indicates significance at the 1% level.

Table 6.4 presents the estimates of Equation (4). The parameter values in Column 1 are suggested by Unger (2007). Similar to the testing of the Walker specification, all coefficients are insignificant, except the coefficient of the cultural and physical distance (2.85), which is

[63] Please note that the failure of the Walker model to explain TBML well could come from the fact that the parameters of the Walker model have never been tested but resulted from an inspirational guess or because money laundering in general is explained significantly differently than TBML.

significant at the 1% level. Column 5 presents the *t*-test values on the null hypothesis that the Unger parameter values are 'true'. The results make clear that all variables are significantly different from what is suggested by Unger (2007). We conclude that we hardly find any empirical underpinning for the prototype models of Walker (1995) and Unger (2007), except that one of the key variables of the gravity model – distance – is apparently essential.

6.5 A gravity model for trade-based money laundering

Since the Walker model and the Unger model do not seem to explain TBML well, we use a traditional gravity model (Equation 1) expanded by the variables used in the attractiveness indicators of the Walker and Unger models to explain TBML.[64] This model has a straightforward log-linear nature instead of the combination of additive and multiplicative variables as in the prototype models of Walker and Unger.

Table 6.5 presents the results of the gravity model for TBML flows. One of the driving factors of TBML is licit trade itself: the larger the trade flow, the larger opportunities for fraud. Therefore, in the TBML model at the left-hand side of Table 6.5 trade has been added as explanatory variable. We call this the 'model with trade', as it explains the impact of trade separately from the other explanatory variables. The next paragraph will discuss the 'model without trade'. We consider the model with trade to be the best specification of a TBML gravity equation, because it allows us to distinguish between the impact of trade on TBML and the effect of other determinants.

Estimation results confirm that trade is an important determinant of TBML, as its coefficient is significant at the 1% level. The coefficient of around 1 indicates proportionality between trade en TBML, in line with the Lorentz curve analysis in Section 3.2. Except for trade and the border dummy (which in this case is basically a dummy for US trade with Canada and Mexico), many independent variables are insignificant or only significant at the 10% level. Therefore, we use Akaike's Information Criterion (AIC) to search for the best fit model specification. The model with the lowest AIC is presented at the right hand side of Table 6.5.

---

[64] Our dataset consists of estimations of TBML from the US to the rest of the world only, while normal gravity models use data of flows between all origins and destinations. To show that it is possible to estimate a gravity equation with data on only a part of the flows, we estimated a standard gravity equation for trade with our dataset in Appendix 3.

The AIC indicates to drop *Population*, *Common Language*, *Colonial Background*, *Swift membership*, *Financial deposits, Corruption level* and *Bank secrecy*. In this leaner specification, *GDP*, *Egmont Membership*, *Government Attitude* and *Distance* become significant at a 5% level. TBML flows can best be explained by GDP (which represents the mass of the countries), trade (which represents the mass of the flow in which TBML is hidden), distance (as is standard in gravity models, corrected for border countries) and two anti-money laundering policy variables (*Egmont-membership* and *Government attitude*). Surprisingly, the two anti-money laundering policy variables have the opposite sign of what was expected by Walker (1995) and Unger (2007). Membership of Egmont, *i.e.* a cooperation agreement between countries to fight money laundering together, leads to more TBML and countries with a government that have a hostile attitude towards money laundering have more TBML than those with a more tolerant attitude.

Table 6.5 Trade-based money laundering gravity model for US outflow (2004)

	Gravity model for TBML				Best fit model (AIC)	
	Model with trade		Model without trade		Structural model	
	Coef.	s.d.	Coeff.	s.d.	Coef.	s.d.
Trade	1.02***	0.09			1.03***	0.09
GDP	0.22*	0.18	0.85***	0.28	0.14**	0.07
Population	-0.05	0.10	0.06	0.24		
Border dummy	-2.60***	0.64	-0.13	1.61	-2.53***	0.57
Common language	0.33	0.35	0.42	0.53		
Colonial background	-0.20	0.61	0.66	1.14		
Swift membership	-0.15	1.20	0.00	1.69		
Egmont membership	0.48	0.33	1.42**	0.59	0.59**	0.28
Financial deposits	-0.05	0.32	1.05*	0.63		
Corruption level	0.20	0.28	-0.14	0.57		
Bank secrecy	0.11	0.19	0.40	0.39		
Government attitude	-0.44*	0.25	-0.64	0.50	-0.39**	0.19
Conflict	0.31*	0.19	0.24	0.31	0.29	0.19
Distance	-0.34*	0.19	-1.10***	0.31	-0.34***	0.14
Constant	-5.26**	2.24	5.42	4.53	-3.88**	1.54
Nr. of observations	199		199		199	
Adjusted R^2	0.882		0.669		0.885	

Notes: *, **, *** indicates significance at, respectively, the 10%, 5% and 1% level. We estimated with OLS and calculate robust standard errors.[65] All variables are expressed in logs.[66]

[65] To test for multicollinearity we perform a VIF-test on all three models presented in Table 6.5 The results for the model with trade, the model without trade and the best-fit model are respectively 2.82, 2.68 and 1.85. Since all VIF-values are far below the proposed cut-off points of 5 and 10, we conclude that we do not have a multicollinearity problem.

These relations can probably best be explained by the fact that TBML has been discovered quite recently, while the anti-money laundering policies at the moment are targeted almost completely on the traditional form of money laundering in the financial system. Our results suggest that money launderers use TBML as an alternative for traditional money laundering when the country they send their money to is fighting (the traditional form of) money laundering intensively. An alternative explanation is that countries which face more money laundering are more eager to undertake adequate actions to combat it.

Finally, the middle panel of Table 6.5 presents the estimation results of the 'model without trade'. Here, the underlying classical gravity model for international trade – with GDP, Population and Distance as its main determinants – is as if substituted in the 'model with trade'. As a result, TBML is explained directly from the traditional gravity variables GDP, Population and Distance as well as the typical money laundering variables. As common in trade models the GDP coefficient does not differ significantly from 1 and the distance measure coefficient does not deviate significantly from -1. Remarkably, in this specification, the Egmont membership coefficient is significant at the 5% level, while the Financial deposits coefficient is significant at the 10% level. However, the goodness-of-fit measure (adjusted R^2) makes clear that this model without trade is inferior to the model with trade (69.2% versus 88.2%). Apparently, the (complete) trade data are of great importance in explaining TBML.[67]

6.6 Conclusion

Combating money laundering is important for society. However, an evaluation of the effects of anti-money laundering policies is hampered by an enormous lack of data. That has motivated Walker and Unger to propose a prototype model for money laundering which has

---

[66] Note that we use only one variable for categorical variables (as explained in Table A2 in Appendix 1), instead of a dummy for each category. Using a dummy for each category does not alter the results significantly.
[67] In the reduced form model only the model value of trade (exclusively of the residual or error term) while in the structural model the complete value (including the error term) is used.

characteristics of the traditional gravity model. The model parameters were based on guess estimates. We use a dataset of Zdanowicz on trade-based-money laundering and find that TBML is indeed much more frequent in relative terms in US trade with countries which have a dubious status in terms of money laundering. With this dataset, we are able to test the Walker and Unger prototype models for money laundering when applied to TBML.

Our conclusion is that these models perform badly for this subset of money laundering. We replace the functional form of the Walker model, which is a mix of linear and multiplicative variables, by a multiplicative traditional gravity equation as frequently used for trade flows, and extend this model with explanatory variables from the Walker and Unger models. We are then able to explain the distribution of TBML between 199 countries and the US in a satisfactory manner. TBML appears to be highly related to licit trade, which permits this kind of money laundering to go unnoticed, as illicit proceeds hide in the large pool of licit exports and imports. Other explanatory variables are distance between the respective two countries, correction for bordered countries, GDP of the importing country, and whether the importing country fights money laundering. The latter is measured by membership of Egmont, *i.e.* a multilateral cooperation agreement to fight money laundering together, and by the attitude of the importing country towards money laundering. Hence, we can apply the traditional intuition of a gravity model – that is, explain a flow by two masses and the mutual distance – to TBML. One might expect that governments which agree to fight money laundering (through their Egmont Membership) and have a hostile attitude towards money laundering experience less TBML. However, our results suggest the opposite: countries, which have strict anti-money laundering regulation, experience more trade related money laundering. This may indicate that criminals have discovered a new way of laundering by using TBML to escape stricter anti-money laundering regulation of the financial sector.

6.7 References

Anderson, J.E. (1979), “A theoretical foundation for the gravity equation”, American Economic Review, **69:** 106-116.

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

Anderson J.E. and Van Wincoop, E. (2004), “Trade costs”, Journal of Economic Literature, **42**, 691-751.

Bagella, M., F. Busato and A. Argentiero (2009), “Money laundering in a microfounded dynamic model simulations for the US and the EU-15 Economies”, *Review of Law and Economics*, **5**, 879-902.

Baker, R. W. (2005), *Capitalism's Achilles heel, dirty money and how to renew the free market system*, John Wiley and Sons, Hoboken, New Jersey.

Bené C., Lawton R. and Alllison E.H. (2010), "Trade matters in the fight against poverty. Narratives, perceptions, and (lack of) evidence in the case of fish trade in Africa", *World Development,* **38**: 933-954.

Bergstrand, J.H. (1985), "The gravity equation in international trade: some microeconomic foundations and empirical evidence", *Review of Economics and Statistics*, **67**: 471-481.

Bikker, J.A. (1987), "An international trade flow model with substitution: an extension of the gravity model", *Kyklos*, **40**: 315-337.

Bikker, J.A. (2010), "An extended gravity model with substitution applied to international trade", pp. 135-162 in: Brakman, S., and P.A.G. van Bergeijk (eds), *The gravity model in international trade: advances and applications,* Cambridge University Press.

Deardorff, A.V. (1998), "Determinants of bilateral trade: does gravity work in a neoclassical world?" pp.7–28 in J.A. Frenkel (ed.), *The regionalization of the world economy*. London, Chicago: Chicago University Press.

Chaikin, D. and J.C. Sharman (2009), *Corruption and money laundering: a symbiotic relationship*, Palgrave Macmillan.

Dreher, A. and F. Schneider (2010), *Corruption and the shadow economy: an empirical analysis*, Public Choice 144, 215-238.

Eichengreen, B., and Irwin, D.A. (1998), "The role of history in bilateral trade flows", pp. 33-57 in: J.A. Frankel (ed.), *The regionalization of the world economy*, Chicago, University of Chicago Press.

FATF (2006), *Trade-based money laundering, Financial Action Task Force*, http://www.fatf-gafi.org/dataoecd/60/25/37038272.pdf.

Helliwell, J.F. (2000), *Language and trade, gravity modeling of trade flows and the role of language*, Canadian Heritage, Canada.

Helpman, E. and Krugman, P.R. (1985), *Market structure and foreign trade: increasing returns, imperfect competition, and the international economy*, Cambridge (USA).

Linders, G.J.M. (2006), "Intangible barriers to trade; the impact of institutions, culture and distance on pattern of trade", Tinbergen Institute Research Series nr. 371, Tinbergen Institute.

Pölyhönen, P. (1963a), "A tentative model for the volume of trade between countries", Welwirtschaftliches Archiv **90**: 93-100.

Pölyhönen, P. (1963b), "Toward a general theory of international trade", Economiska Samfundets Tidskrift, **16**: 69-77.

Pulliainen, K. (1963), "A world trade study: an econometric model of the pattern of the commodity flows of international trade in 1948-60", Economiska Samfundets Tijdskrift, **16**, 78-91.

Redding, S. and P.K. Scott (2003), "Distance, skill deepening and development: will peripheral countries ever get rich?", *Journal of Development Economics*, **72**: 515-541.

Schneider, F. (2008), "Money laundering and financial means of organized crime: some preliminary empirical findings", p. 309-330, in: *Global Business and Economics Review*, Interscience Enterprises Ltd, **10** (3).

Tinbergen, J. (1962), *Shaping the world economy: suggestions for an international economic policy*. Twentieth Century Fund, New York.

Unger, B. (2007), *The scale and impacts of money laundering*, Edward Elgar, Cheltenham, UK

Unger, B. and J. den Hertog (2011), Water always finds its way – Identifying new forms of money laundering, *Crime Law and Social Change* (A Special Issue on Black and Gray Finance, edited by S. Hsu; to appear in December 2011).

Walker, J. (1995), "Estimates of the extent of money laundering in and throughout Australia", report for the Australian Financial Intelligence Unit AUSTRAC.

Walker, J. (1999), "How big is global money laundering?", *Journal of Money Laundering Control* **3**: 25-37.

Walker, J. and B. Unger (2009), "Measuring global money laundering: the Walker gravity model", *Review of Law & Economics* **5**: 821-853.

Zdanowicz, J.S. (2004a), *U.S. trade with the world and Al Qaeda watch list countries - 2001: An estimate of money moved out of and into the US due to suspicious pricing in international trade*, http://business.fiu.edu/pdf/PrintJun2007/tfml.pdf.

Zdanowicz, J.S. (2004b), "Who's watching our back door?, Business Accents", *College of Business Administration magazine*, Florida International University, Fall 2004, Volume 1, pp. 26-29.

Zdanowicz, J.S. (2009a), "Money launderers wash billions through international trade", *Miami Herald*, May 11, 2009: http://www.miamiherald.com/154/story/1041027.html.

Zdanowicz, J.S. (2009b), "Trade-based money laundering and terrorist financing"*, Review of Law and Economics*, **5**: 855-878.

Chapter 7 Conclusion

Over the past decade, we see more and more developing countries become more deeply integrated in the world economy in an attempt to increase their economic growth rate. Globalization on the one hand promises market access and on the other hand enforces greater competition. The trade-off between the two forces has always made trade liberalization controversial. The reason for the heterogeneous impact on economic agents lies in the fact that the agents affected are heterogeneous. Hence we observe some benefits from the opportunities created by globalization while others suffer from the heightening threat to their survival. In this Dissertation I identify and characterize the producers who benefit from trade and model the proposed mechanism behind their success. The results of the investigation can be summarized to draw some main conclusions.

In chapter 2, I investigate the productivity heterogeneity among manufacturing and services sectors firms in 15 Latin American countries. By paying special attention to the shape of the productivity distribution and the correlation between its variation and the development stage, I confirmed the absence of a minimum threshold for firm viability at the country and sectoral level, as well as the absence of a productivity threshold between exporters and the non-exporting firms. The results challenge the current popular models for trade which imply such thresholds. A simple and meaningful remedy to the contradiction is shown in Chapter 3 by introducing heterogeneity in fixed costs of firms into the Melitz model. The modified model predicts a smooth bell shaped distribution and provides affirmative results suggesting the overall positive effect of trade liberalization on welfare (the increase in aggregate productivity). It also drew our attention to the assumptions we made regarding the underlying productivity and fixed cost distribution and its effect on the productivity distribution of viable firms and exporters in the equilibrium.

Chapter 4 goes a step further in providing the causal linkage between productivity heterogeneity among producers and their welfare by using micro-data of agricultural producers in Indonesia. These results support the self-selection mechanism; that is the more productive producers are, the more likely they are to invest to increase quality and so earn a higher profit. These farmers are also more likely to become producers supplying supermarkets and to have a higher household income. The empirical work reviewed in Chapter 5 shows the progress in the Geographical Economics literature and found strong evidence in support of the four key implications of geographical models. The literature shows that (i) market and supplier access have a significant positive impact on manufacturing wages; (ii) market access is an important determinant of firm location decision; (iii) trade liberalization influences regional specialization patterns and in some cases leads to regional divergence; and (iv) exogenous shocks have long-term economic effects. Chapter 6 estimates the bilateral flow of trade-base money laundering (TBML) activities using the gravity model. The results suggest that trade and GDP are strong positive drivers of illegal money flows, while the usual geographical indicators (border dummy and distance) significantly deter such flows.

In this dissertation, we minimize the use of a causal interpretation in the chapters where we were only able to characterize the heterogeneity found, in the case of Chapter 2. Having acknowledged the limitation of the cross sectional data, we rely on other well recognized estimation structural form, the Heckman self-selection model, to establish the causal interpretation in Chapter 4.

The linkage between productivity and welfare lies in its deterministic role in affecting the decision producers make for investments that affect their chance of entering a new market. I do not exclude other producer characteristics that may influence their investment. This study provides motivation for the effort and investment made in the area of microfinance. The lack of financial access limits the ability of individuals and enterprises to invest in their business and this limits their full potential to upgrade their production and/or escape poverty. Producers in the developing countries will also need better infrastructure such as roads and communications and better institutions that facilitate market efficiency such as enforcement of contracts, decreasing information asymmetry and the formation of a supporting service sector. These are all factors that have been suggested to be the precondition for trade to be beneficial.

For developing countries increasing their involvement in the global economy through trade, the trade-off between increasing competition and increasing market access is an ongoing battle. Understanding the heterogeneity among the affected producers and the constraints that are holding them back from reaping the benefits of access to foreign markets is of great importance to economists and policy makers alike. To conclude, I quote a speech Barack Obama made in Berlin in 2008:

"*This is the moment when we must build on the wealth that open markets have created, and share its benefits more equitably. Trade has been a cornerstone of our growth and global development. But we will not be able to sustain this growth if it favors the few, and not the many. Together, we must forge trade that truly rewards the work that creates wealth, with meaningful protections for our people and our planet. This is the moment for trade that is free and fair for all.*" -- Barack Obama

Source: Speech in Berlin, in Change We Can Believe In, p.268 Jul 24, 2008.

Nederlandse Samenvatting (Dutch Summary)

Ontwikkelingslanden integreren steeds meer in de wereldeconomie teneinde een proces van economische ontwikkeling op gang te brengen. Dit leidt aan de ene kant tot toegang tot nieuwe markten en aan de andere kant tot grotere concurrentie. De uitruil tussen deze twee krachten heeft handelsliberalisatie altijd controversieel gemaakt. De diversiteit van economische agenten, zoals verschillen in productiviteit tussen bedrijven in dezelfde sector, leidt ook tot diversiteit in de gevolgen van toenemende globalisering. Deze dissertatie schetst een beeld van die diversiteit voor producenten en voor hun geografische omgeving.

Hoofdstuk 2 analyseert de diversiteit in productiviteit voor producenten van goederen en diensten in vijftien latijns amerikaanse landen. Door aandacht te besteden aan de vorm van de productiviteitsverdeling en de relatie met de mate van economische ontwikkeling toon ik aan dat er geen “afkapwaarde” is voor levensvatbaarheid van bedrijven, noch voor het onderscheid tussen exporteurs en binnenlandse ondernemingen. In plaats daarvan, en in contrast met de huidige theoretische modellen, is er een graduele overgang van het ene type bedrijf naar het andere.

Hoofdstuk 3 levert een theoretische onderbouwing voor de in hoofdstuk 2 gevonden afwezigheid van een afkapwaarde door niet alleen diversiteit voor marginale kosten maar ook voor vaste kosten in het model van Melitz te incorporeren. Deze aanpassing geeft inderdaad de vereiste graduele overgang van het ene type bedrijf naar het andere. Bovendien geeft dit model een onderbouwing voor een positief welvaartseffect verbonden aan globalisering door middel van een stijging van de gemiddelde productiviteit van de actieve bedrijven.

Hoofdstuk 4 zet nog een stap verder in de verbinding tussen diversiteit en welvaart door de bedrijvigheid van producenten van pepers in Indonesië te analyseren. Het blijkt dat de meer productieve bedrijven ook bereid zijn te investeren in hogere kwaliteit, waardoor ze een grotere kans hebben dat hun producten voor een hogere prijs in supermarkten worden verkocht. Deze bereidheid leidt uiteindelijk ook tot hogere welvaart.

Hoofdstuk 5 richt zich op geografische diversiteit en geeft een overzicht van recent empirisch werk op het gebied van de geografische economie (ook wel nieuwe economische geografie genoemd). Het biedt een stevige onderbouwing van vier belangrijke implicaties van deze literatuur, namelijk (i) verbeterde toegang tot afzet- en aanbodmarkten leidt tot hogere lonen,

(ii) toegang tot afzetmarkten zijn een belangrijke factor in de lokatiebeslissingen van bedrijven, (iii) handelsliberalisatie heeft invloed op regionale specialisatie en leidt in sommige gevallen tot regionale divergentie en (iv) exogene schokken kunnen lange-termijn economische gevolgen hebben.

Hoofdstuk 6, tenslotte, richt zich ook op geografische diversiteit en analyseert de (illegale) stromen van witwas-activiteiten gebaseerd op internationale handelsstromen. De tot nu toe gebruikte witwas-modellen blijken duidelijk ontoereikend. In plaats daarvan ontwikkeld het hoofdstuk een robuust model gebaseerd op waarneembare handelsstromen en omvang van inkomen (leidend tot een toename van witwas-handelsstromen) en geografische indicatoren zoals buurlanden of fysieke afstand (leidend tot een daling van witwas-handelsstromen).

Curriculum Vitae

Han-Hsin Chang (1986) was born in Taiwan, where she completed her high school at the National Experimental High School in 2004. Between 2004 and 2008, she completed her bachelor education in Business Management at the National Sun Yat-sen University. During this period, she went to Europe for a year as an exchange student. This experience strengthened her motivation to switch to study economics. From 2008 to 2010, she pursued her studies at Utrecht University School of Economics and graduated with an M.Phil. degree in Multidisciplinary Economics.

In September 2010, she became a PhD candidate at Utrecht University School of Economics, where she completed this dissertation. During her PhD, Han-Hsin visited the University of Adelaide in Australia as a visiting research student for three months. As of September 2013, she works as a research associate (consultant) with the AIEN group in the Office of Regional Economic Integration (OREI) in the Asia Development Bank (ADB), Philippine.

Han-Hsin Chang is interested in a research position that provides solutions and opportunities to improve the welfare of the poor. She is passionate, ambitious, and disciplined, with good analytical skills. She works hard, with a particular strength for critical thinking and coordinating ideas from discussions. She is keen to work in an international context and can quickly adapt to a dynamic work setting.

Teaching experience

Monetary Economics, Mathematics, Econometrics

Publication

- Firm heterogeneity and development: Evidence from Latin American countries (2013), *Journal of International Trade & Economic Development*, 22(1): 11-52, with Charles Van Marrewijk.
- Gravity Models of Trade-Based Money Laundering (2013), *Applied Economics*, 45(22): 3170-3182, with Ferwerda, J. M. Kattenberg, B. Unger, L. Groot and J. Bikker.
- Firm and sectoral heterogeneity in mark-up variability (2013), *Journal of International Trade and Development*, 22(1): 157-178, with Saara Tamminen.
- Empirical applications of geographical economics models, (forthcoming) in *Handbook of Research Methods and Applications in Economic Geography* with C. van Marrewijk and M. Schramm.

TKI Dissertation Series

USE 001 **Bastian Westbrock** (2010): *Inter-firm networks: economic and sociological perspectives.*

USE 002 **Yi Zhang** (2011): *Institutions and International Investments: Evidence from China and Other Emerging Markets*.

USE 003 **Ryan van Lamoen** (2011): *The Relationship between Competition and Innovation: Measuring Innovation and Causality.*

USE 004 **Martijn Dröes** (2011): *House Price Uncertainty in the Dutch Owner-Occupied Housing Market*

USE 005 **Thomas van Huizen** (2012): *Behavioural Assumptions in Labour Economics: Analysing Social Security Reforms and Labour Market Transitions*

USE 006 **Martijn Boermans** (2012): *International Entrepreneurship and Enterprise Development.*

USE 007 **Joras Ferwerda** (2012): *The Multidisciplinary Economics of Money Laundering*

USE 008 **Federico D'Onofrio** (2013): *Observing the country: a history of Italian agricultural economics, 1900-1930.*

USE 009 **Saraï Sapulete** (2013): *Works Council Effectiveness: Determinants and Outcomes.*

USE 010 **Britta Hoyer** (2013): *Network Formation under the Threat of Disruption.*

USE 011 **Coen Rigtering** (2013): *Entrepreneurial Orientation: Multilevel Analysis and Consequenses.*

USE 012 **Beate Cesinger** (2013): *Context and Complexity of International Entrepreneurship as a Field of Research.*

USE 013 **Jan de Dreu** (2013): *Empirical essays on the governance of financial institutions*.

USE 014 **Lu Zhang** (2013): *Industrial Specialization: Determinants, Processes and Consequences.*

USE 015 **Matthias Filser** (2013): *Strategic Issues in Entrepreneurship and Family Firm Research.*

USE 016 **Mikko Pohjola** (2013): *A Compilation of Studies on Innovation in Firms: Capabilities, Strategies, and Performance.*