UNIVERSITEIT GENT

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

Multinational Activity and Firm Performance

Victoria Purice

2015

Supervisor:
Prof. Dr. Bruno Merlevede

SUBMITTED AT GHENT UNIVERSITY, FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION
IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR IN ECONOMICS

Multinational Activity and Firm Performance

Victoria Purice

Department of General Economics

Ghent University

A thesis submitted for the degree of

Doctor of Philosophy at Ghent University

October 2015

This book was typeset using $\LaTeX$.

Multinational Activity and Firm Performance

Victoria Purice

Doctoral Jury

Prof. dr. Marc De Clercq	Dean-President, Ghent University
Prof. dr. Patrick Van Kenhove	Academic Secretary, Ghent University
Prof. dr. Bruno Merlevede	Promotor, Ghent University
Prof. dr. Steven Brakman	University of Groningen
Prof. dr. Richard Kneller	University of Nottingham
Prof. dr. Glenn Rayp	Ghent University
Prof. dr. Koen Schoors	Ghent University
Prof. dr. Hylke Vandenbussche	University of Leuven

"Nobody ever figures out what life is all about,
and it doesn't matter. Explore the world.
Nearly everything is really interesting if you
go into it deeply enough."

Richard Feynman

To my grandmothers

Acknowledgments

If only I had a penny for every time someone said "Doing a PhD is a lesson in persistence". And so it is. About 4 years ago when I started this journey I could not have imagined what a roller-coaster writing a dissertation would be. Looking back at it all now, I can honestly say it was one of the best times of my life. I have learned so much even at times when it did not feel like learning, and I had a lot of people around me that have helped me to become a better researcher and a better person. To all of them I wish to express my gratitude here.

First, I would like to thank my promotor, Bruno, for the opportunity of doing a PhD with him at Ghent University. I was lucky enough to get on a project that fit both of our interests and that I could continue some of the work I have done previously. Bruno, thank you very much for giving me freedom to explore the things that interest me but also for being there when I got stuck and could not see any more. Your calm and patience when things did not go the way I expected, or when I discovered errors in my code, have really helped me pull through. Moreover, thank you for the yearly kayaking/barbecuing session at your home. I must say I enjoyed the food and playing with Ortiz much more than the physical exercise, but that is solely my fault and not yours. I will also fondly remember going to ETSG conferences together and the avid dinner discussions on economics, region and US politics. Hopefully some of this will one day spillover in our future work, as I hope we will continue to cooperate and write cool papers together.

I would also like to thank Koen Schoors, who is my academic grandfather seeing as he was Bruno's promoter. Koen, you are the reason I came to UGent and met Bruno. I want to thank you for believing in me and giving me this chance. Moreover, I very much appreciate your input when reading my papers as well as your contributions to internal seminars. Your ability to understand and ask the right question never ceases to amaze me. I don't know how you do it, but I hope to one day be at least half as good at distilling one's idea as you are.

I would also like to thank Glenn Rayp for being in my guidance committee. Glenn, thank you so much for reading my work and for your invaluable comments. I very much appreciate your views, be it on economics, history or philosophy. Unfortunately I never got the chance of having you as a teacher, but I wish I could have taken your course on the history of economic thought, a much needed overview for economists everywhere.

I would of course like to thank the external members of my reading committee, mainly Professor Steven Brakman, Professor Richard Kneller and last but not least Professor Hylke Vandenbussche. Thank you very much for agreeing to read and give comments on my work, comments which have made the papers better. I appreciate the fact that you devoted your time despite extremely busy academic schedules.

I would very much like to thank Professor Rudi Vander Vennet from Financial Economics. Although we were in different departments, I am very grateful for his support of the SMYE. I have also thoroughly enjoyed the too few, but very enlightening conversations, be it on the city of Ghent, its history and rich culture, or the financial crisis and how to build a more resilient banking system.

Next I would like to thank my friends and colleagues at the department of General Economics. Firstly, the people in our research group, Matthijs and Karo. You guys know better than anyone how much we struggled with the data, how at times it felt like we would never have it finished. I'm glad that those feelings were temporary and that we managed in the end. To the newcomers of the group, Angelos and Bram, good luck with your PhD and make all that data work count! I'd like to thank Jolyn for being such a sunny person and for bringing so much fun to the group. I really enjoyed spending time together in Goettingen! I would also like to thank Matthijs for joining on the early morning trips to Brussels for the "Mas-Colell" course, for going together to conferences and for being the Dutch island in our department. Since you left it's not the same. I would also like to thank Ilse and Delphine. Together with them and Karo we had for a very long time the "women's office". Girls, it was so good to have you around, to give each other birthday cards, to celebrate good news and share the not so good ones. Karo, I know I'm older, but you've been such an example of strength and calm when finishing your PhD, inspiring me and others as well. Ilse, good luck with the postdoc and thank you for being such an example to us all, women or not,

on how to do research. Elien, I had so much fun taking the Econometrics course with you in Leuven. Thanks for joining me and for the honest conversations on the train and otherwise. I want to send you the best thoughts and to wish you the best of luck in finishing your PhD! I would like to thank Sam, Stijn and Ben! You guys have been amazing. Thanks for the laughter, for helping with my rats (you know who you are), for organizing events on the terrace. I'll miss your daily ball game and the occasional cartoon on the white board. To Ruben, the chairman of the "Friday drinks" and possible the most optimistic person of the department, I know it's the end, but it's also a new beginning and we will make the most out of it! I cannot even begin to express my gratitude to Eva and Martine. Thank you for everything you do, day in and day out. You are the glue that holds the department together and knowing that we can always knock at your door is very comforting. Thanks to Mustafa for always inquiring how I was and for sharing the ups and downs of an academic career. Finally, a big thank you to Kevin, with whom I had endless discussions at lunch and from whom I learned a lot about money, banks, foreign policy, gardening and wine. I will miss our conversations dearly.

I would also like to thank the people from Financial and Social Economics. Many of them have long been gone, but still I would like to thank Julio, Arnoud, Frederick, Glenn, Valerie, Xing, Yannick, Corinna, Garo and Dries. Thanks for sharing your PhD experiences with me, for joining for the Friday drinks or the occasional pizza/burger. It felt good knowing that these PhD struggles are just a normal thing and that everybody goes though them one time or the other. I will miss our joint internal seminars and all the comments and support we showed each other.

I want to say a special thank you to the SMYE2015 team. Ewoud, Matthijs, Joris and Eva, thanks a lot for agreeing to take part in this amazing adventure. I still remember being all nervous before the Nobel Prize announcement and not being able to believe it when it came out. I think we did a fantastic job and I feel very lucky to have been able to work with you! Still looking forward to that Caribbean trip though...

As I thanked my second Belgian family, UGent, I would also like to mention my first one. Therefore special thanks to my former colleagues at the College of Europe in Bruges. They were the first who introduced me to the inner workings of our European institutions but also gave me the flavour of what it means to be Belgian and European.

Eric, Jacques, Jessie, Renée - thank you for your support and for being such great colleagues. Thanks as well to Chiara and Federico, two other ex-pats in this Europeanized world, for sharing their experiences and struggles over a beer at De Garre. Professor Hine, thank you for encouraging me to do a PhD and for the DG Trade trips to Brussels where we got to see how policy is made.

I would like to thank my, by now, long-time friend and co-author Jaap Bos. Jaap has been there even before the PhD and was a pillar of support for both me and Martien. Talking about technical efficiency, how to best smoke salmon or what is the most appropriate camping gear, you are surely going to learn something new from him every time. I hope that now we can see each other more often despite the four hour train ride.

A special thank you to Abe and Marco, two people that became our friends under the most unexpected circumstances, but who remain so dear to me and Martien even if we don't see each other very often.

Thank you Ilse from Illie Mangaro for the reception and for all the hard work, but also the fun memories during the SMYE. You were such a rock during those three days and we would not have been able to succeed without you. The food was great and I trust that your new adventures will prove just how good you are!

My family, who are more often than not far away but who always gave me strength and pushed me towards becoming a better person, even at great costs for them. Mama, tata, vă mulţumesc pentru tot ajutorul vostru şi pentru că m-aţi pus pe mine pe primul loc si m-aţi susţinut in orice moment, chiara dacă vouă nu v-a fost uşor. This book is dedicated to my grandmothers who never got the opportunities that I had and that made me realize how lucky I was to be born at such a privileged time in history. I will forever be grateful to them.

I would like to thank my family in law who have welcomed me into their home and have treated me like one of their own. Jan, Grazyna, Monica, Aaldert and little Mathias, thank you for being there in good and bad times, at graduations but also when we had to move heavy mattresses, tables, fridges and you name it. I hope we can be there for you as well and that we can spend more quality time together once we are in the same country.

Last but not least, I would like thank my husband Martien who has been the best possible companion in this period, not only because he had to go through the same

process a year ago, but also because he was the sweetest and kindest person around. He not only read my papers, gave me comments, made suggestions, but also made sure that our household did not collapse while I was immersed in writing this book. Martien, I know you said that I was responsible for 50% of your PhD. I think in my case you are responsible for a much larger share, as without your constant support and example I know I would have never been here in the first place. And finally, while I love going on conference city trips and discovering with you the Basilica of San Clemente and the Science Museum, I really think we need to take a proper vacation, and soon!

Victoria Purice

Ghent, September 2015

Table of Contents

List of Figures

List of Tables

1

Introduction

1.1 General Introduction

Achieving economic growth requires both a healthy real economy, with firms innovating and constantly improving their technologies, as well as a well-working financial system that can provide credit to these firms (King and Levine, 1993). With most capital restrictions being lifted in the last several decades and the financial sector being allowed to allocate resources to the best performers, the world has seen an unprecedented deepening in globalisation, trade and multinational activity. The collapse of the Soviet Union and the emergence of transition economies in Eastern Europe has proved to be a perfect opportunity for firms to expand their operations but also to bring with them previously unavailable technology and know-how. This in turn has resulted in the fragmentation of production chains and the emergence of European and global production networks (see e.g. Timmer, Erumban, Los, Stehrer, and de Vries, 2014).

However, financial and economic deepening has come at a cost. Living in a more interconnected world, financial and economic shocks are being transmitted more easily and can have serious repercussions to almost any corner of the world. The existence of multinational banks and multinational firms has facilitated the capital flow in good times but also its withdrawal when countries experience economic downturns. These direct effects are especially clear in the case of banks, where parent banks have called back capital from their subsidiaries abroad, potentially damaging the local economies (de Haas and van Lelyveld, 2010; Cetorelli and Goldberg, 2012; de Haas and Lelyveld, 2014). In the case of manufacturing firms, where the data is just becoming available, we do know that multinationals are more 'footloose' in times of crises (see e.g. Godart, Görg, and Hanley, 2011), they will divest from location that are further away (Landier, Nair, and Wulf, 2009; Abraham, Goesaert, and Konings, 2014), and if they do help their affiliates in the crisis, it is mainly those that have important production ties with the MNE (Alfaro and Chen, 2012).

Notwithstanding these negative externalities, multinationals have been a source of growth and productivity improvements. We know that these firms are generally larger, older and more productive (see e.g. Helpman, Melitz, and Yeaple, 2004). Studies have shown that they create jobs, bring in new capital and have better technologies than local firms, prompting governments to promote policies which will attract more of these MNEs to locate in their countries. Their impact can be direct, via mergers and acquisitions, where a domestic firm is purchased. However, more often than not, the improvements brought by MNEs are indirect and unintended. These so called spillovers occur via replication, demonstration or due to labour mobility from MNEs to domestic firms and vice-versa. Moreover, their impact differs depending on whether they take place between firms that are active in the same industry (horizontal) or in industries that are linked via supplier/buyer chains (vertical).

The direct impact of MNEs are mostly studied via mergers and acquisitions, as an MNE becomes the parent of a domestic firm and thus can directly control its operations. Recent studies by Arnold and Javorcik (2009) and Damijan, Rojec, Majcen, and Knell (2013) suggest that these productivity improvements are substantial, and that in a reversed way, if an affiliate is sold off and becomes again domestic, its productivity will decrease (see e.g. Javorcik and Poelhekke, 2014). Moreover, for a large set of ten transition countries, Damijan, Rojec, Majcen, and Knell (2013) find that productivity premia of foreign owned firms are as large as 9% and are due to direct technological transfers from parent to affiliate. They suggest that horizontal spillovers have been increasing in importance and that they might become even more significant than the vertical ones.

The indirect impact of MNEs on local firms have been extensively analysed since Caves (1974). The results are however mixed, largely due to the multitude of channels via which spillovers can manifest themselves as well as the indirect way we measure this phenomenon (Görg and Greenaway, 2004). Nonetheless, a new strand in the literature has tried to disentangle the effects of horizontal spillovers from the vertical ones (Schoors and van der Tol, 2002; Javorcik, 2004). These studies have found that while horizontal spillovers are difficult to measure and predict in a consistent way, the backward spillovers are often positive due to the interest of the MNE in helping its suppliers to deliver quality inputs. This therefore translates in higher productivity improvements for domestic suppliers of multinationals.

Finally, with the onset of the Global Financial Crisis, the interplay between banks and real economic activity has become more evident. On the one hand we know that, especially in Europe where banks play a major role in financing the economy and ultimately economic growth, bank credit is relevant for firm survival (see Beck, Degryse, de Haas, and van Horen, 2014). On the other hand, a bank's balance sheet is a map of which sectors are dominating the economy, and could be an indicator of future growth. Globalisation and access to international markets, aided by the emergence of complex financial products, allowed banks to expand in size. This however has not proportionally translated into more loans to the real economy (see e.g. Bezemer, 2014). A recent paper by Rousseau and Wachtel (2011) sheds light on how increasingly large financial sectors negatively impact growth. They suggest that the increase in credit loans have resulted in the weakening of the banking system. Moreover, Cecchetti and Kharroubi (2015) suggest that the growth of large banks and the decrease in more costly relationship banking has led to an increase in the importance of collateral based loans. However, this also resulted in lower productivity and less risky projects being financed, therefore leading to a slower aggregate productivity growth as well as harming innovation and R&D. It is therefore important to know what makes for a stable and efficient financial system.

Against this background, the thesis will investigate several questions related to the direct and indirect impact of MNEs on local firms and the role of financial linkages for this impact. First, we are interested in how the indirect effect of MNEs' presence are measured and whether including previously unexplored factors could explain the mixed results seen in the literature. Second, given that multinational networks are becoming increasingly important in the global economy, we investigate the direct impact of MNE parents on the performance of their affiliates. Moreover, as the emergence of multinational firms has promoted the expansion of large multinational banks as well (see e.g. Poelhekke, 2015), we will look at whether financial development in the country of the parent has a beneficial impact on its affiliates. Lastly, to mirror the developments on the firm level side, we look at how the banking system has changed over time, what these changes meant for financial stability and how this is reflected into the real economy via the balance sheets of these banks.

1.2 Data

To be able to answer these questions, we require a significant amount of firm level data both at the balance sheet level as well as on the ownership level. Most of the data used to answer these questions comes from Amadeus, a European firm-level database collected by Bureau Van Dijk Electronic Publishing. From Amadeus we construct two new datasets *AUGAMA*, a panel of European firms for the period 1996-2011, and *EUMULNET*, a European Multinational Network data set. In Chapter 2 we document the process of building these datasets from the raw Amadeus data for 26 European countries. Moreover, we show that the data sets adequately approximate the structure of the European economy across countries, regions, and industries as portrayed by data from Eurostat (Structural Business Statistics) and Cambridge Econometrics.

1.3 Indirect impact of MNEs on local firms

Chapters 3 and 4 of this dissertation both tackle indirect effects of MNEs on domestic firms. More specifically, Chapter 3 investigates additional factors that might impact how spillovers are transmitted between firms. Previous literature often finds mixed results of spillovers, especially in the case of horizontal ones. This however might be explained by certain assumptions made when measuring the effect of multinational presence. First, many papers assume that once MNEs enters a country, all domestic firms will benefit from spillovers. However, this is not realistic, especially in bigger countries, where firms located further away might be impacted less by the presence of the multinationals compared to firms close-by. This is in part due to the fact that distance plays a significant role in how knowledge and technical improvements of firms are transmitted (see e.g. Giroud, 2012). Second, often it is also assumed that this indirect effect will be almost instantaneous. This is again highly unlikely, as learning from multinationals might take time (see e.g. Merlevede, Schoors, and Spatareanu, 2014). The third chapter of this dissertation therefore combines the effect of distance and time for looking at how foreign direct investment impacts the productivity of local firms. We indeed find that for vertical spillovers distance is crucial, as local suppliers of foreign firms will benefit if located at short distances from MNEs. A positive effect on local competitors via horizontal spillovers is not significantly affected by distance but requires the presence of mature foreign firms, as channels such as labour mobility and competition work over a longer time horizon.

Having seen how important distance is for spillovers transmission, the fourth chapter looks at the role of borders in how domestic firms are affected by multinational presence. So far the literature has assumed that spillovers are bound to the country of MNE's location, but especially in the European Union, where we have seen a deeper economic integration taking place, borders might be less of a barrier for spillovers to materialize. To this end we use a panel of firms from Central and Eastern European countries, where previous studies have confirmed the existence of country-wide spillovers. We analyse whether borders are a strong enough barrier for FDI spillovers from for example Austria and Germany to reach neighbouring countries such as Slovakia or Poland, or whether borders are porous enough to allow for productivity improvement to cross them. We find that national borders constitute an almost insurmountable barrier for horizontal and vertical spillovers from MNEs to the domestic buyers of their products. In the case of vertical spillovers from MNEs to local input suppliers, national borders significantly dampen cross-border spillover effects, but the size of the impact of the border seems related to the regional and economic integration in the border area, as proxied by the existence of the Schengen agreement.

All in all, our findings explain to some extent the mixed results found so far in the literature concerning the horizontal spillovers, which require time to materialize. Moreover, we shed light on how vertical spillovers spread from MNEs to their domestic suppliers. Mainly, we find evidence of importance of proximity as well as the existence, albeit weaker, of cross-border spillovers in Central and Eastern European countries. This points towards a greater integration of the European internal market, where national borders play an ever smaller role in maintaining economic barriers between countries and allow for cross-country technology diffusion.

1.4 Direct impact of MNEs on their affiliates and the role of financial linkages

Chapter 5 explores the direct effects of MNE presence. Using the EMULNET database developed in Chapter 2, we revisit the question of whether being owned by a multinational is bad for affiliate survival or not, as MNEs are often found to close their affiliates more easily than domestic firms (see e.g. Bernard and Jensen, 2007). Moreover, the crisis has shown that MNEs will choose to keep some affiliates at the cost of closing down others, as those with stronger vertical ties have done better. Therefore, this chapter investigates what motivates MNE decision to divest and which characteristics would best predict the exit of their affiliates. We not only consider traditional affiliate characteristics such as age, size and productivity, but include firm-level parent characteristics as well. Moreover, having just experienced a significant financial crisis, which started in the banking sector and ultimately reached the real economy, this chapter explores differences in affiliate performance through the cycle, that is both before and after the crisis, to identify which characteristics of firms or parents improve their chances of survival. Finally, since credit constraints of the parent might play a role, we extend previous literature by investigating the role of financial development of the parent's country in predicting affiliate exit.

We find that once we control for parent performance, affiliates of multinationals and more diversified parents decreases affiliate exit. Moreover, we find some evidence that exit is less likely if affiliates are owned by parents from countries with more developed financial markets, as proxied by domestic credit to consumers. While most factors are robust to the crisis years, the role of MNE's country financial development in improving affiliate survival is less so. On the contrary, investment in affiliates seems to be pro-cyclical, with the beneficial effect of having a parent with easier access to capital breaking down during the crisis years.

1.5 Bank size and the real economy

The recent Global Financial Crisis originated in the US, with devastating consequences for both global trade and economic development (see e.g. Bems, Johnson, and Yi, 2012). Given how important having a healthy financial system is for economic and productivity growth, in Chapter 6 we look at the impact of banks on the real economy. We do so by analysing the evolution of the size and correlation of returns on assets of US banks, as that captures the wave of concentration and deeper interconnectedness that has emerged in the US but also at the global level. Partly due to a wave of liberalisation and deeper economic integration, we can track the evolution of ever larger banks and relate that to an increase in systemic risk.

Our results suggest that the largest banks are too big for the system to be able to absorb potential high increases in correlation between these banks, therefore posing a risk for the entire system. Moreover, a closer look at their balance sheets suggest a shift in the business models of larger banks towards non-interest income generated by banking services and away from financing the real economy. We conjecture that the financial sector would benefit from scaling down and pursuing more traditional bank business models.

1.6 Suggestions for future research

This dissertation takes a view on the direct and indirect impact of multinational firms on other firms, the role of financial linkages as well as the impact of banks on the real economy. We identify important channels via which firms can absorb technological improvements, learn from each other as well as what factors will impact survival of firms in times of economic prosperity or crisis. We find that financial development and internal credit channels between MNEs and their affiliates play a role in affiliate survival but that, unfortunately, this relationship breaks down during the crisis. As the literature on the role of finance for economic performance of firms is underdeveloped, future research should focus on understanding better the role of banks in financing firms, as well as how internal capital markets of large multinational networks function. Moreover, we can explore further the channels of technological transmission and how to improve this process in order to benefit most from innovation and economic growth.

2

Multinational Networks, Domestic, and Foreign Firms in Europe*

2.1 Introduction

This chapter in detail documents the build of a large pan-European firm-level data set, *AUGAMA* ('Augmented Amadeus'), with the aim to serve as a reference for future work. We extensively document our 'augmentation' that overcomes drawbacks -from an academic point of view- related to the way the data-provider, Bureau Van Dijk Electronic Publishing (BvDEP), issues the database. The advantage of our data is that it covers cross-country comparable firm-level data for 26 European countries in a single dataset. This allows for cross-country research at the firm level while maintaining representativeness that is comparable to other recent efforts (see for example CompNet (CompNet, 2014)). Our approach also improves the data with respect to exit and entry patterns. The chapter further documents the build of a dataset of European multinational networks, *EUMULNET*. This dataset uses raw data from the Amadeus database by BVDEP as well. Both databases cover the following European countries: Austria, Bel-

*This chapter is based on a paper co-authored with Bruno Merlevede, Karolien Lenaerts and Matthijs De Zwaan.

gium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Spain, Finland, France, United Kingdom, Greece, Croatia, Hungary, Ireland, Italy, Latvia, Lithuania, the Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, and Slovak Republic. We document representativeness across countries, industries, and regions. We then enrich the database with a measure of total factor productivity using the Wooldridge-Levinsohn-Petrin methodology (see Levinsohn and Petrin (2003), Wooldridge (2009), and Petrin and Levinsohn (2012)).

In the final part of the chapter we present a set of empirical applications that makes use of the datasets to test a number of results from the theoretical literature regarding the productivity of multinational firms vis-a-vis domestic firms. Specifically, in line with (Markusen, 1995), we show that foreign firms are more productive than their domestic counterparts. We show that the TFP premium of foreign over domestic firms on average amounts to 48% in *AUGAMA*. Theoretical work further suggests that only the most productive domestic firms will engage in foreign direct investment (FDI) (see Helpman et al. (2004) for horizontal and Antràs and Helpman (2004) for vertical investment), because only the most productive firms are able to cover the costs associated with this investment. Using *EUMULNET* we show these predictions to hold for European multinationals. For a matched sample of foreign and domestic firms (following Alfaro and Chen (2012)) we find that foreign firms grow faster than domestic firms both pre and post crisis.

The chapter is organised as follows. Section 2.2 documents the construction of the AUGAMA database. In Section 2.3 we consider overall representativeness of the data, while in section 2.4 the focus is on representativeness of the data in the industry and region dimension. In Section 2.5, we introduce the EUMULNET data. Section 2.6 discusses the estimation of total factor productivity and in Section 2.7 uses estimated *TFP* in different applications comparing total factor productivity (growth) of domestic and foreign firms. Section 2.8 concludes this chapter.

2.2 Database

2.2.1 The Amadeus database: basic data source

Raw data are taken from the Amadeus database provided by Bureau van Dijk Electronic Publishing (BvDEP). The Amadeus database is a pan-European database that comprises financial information on public and private companies. BvDEP gathers data from different local, i.e. country-specific, data providers and assembles them into a single database using a comparable format. Additionally, BvDEP assembles further information from firms' annual reports, media coverage, and other sources. It is not clear whether this is done for all firms and countries, but a bias towards large and listed firms seems likely.[1] The Amadeus database is available in different flavours depending on the application of some thresholds for firms to be included. Our data originate from the *'full'* version where no thresholds are applied and all available firms are included in the database. The database contains both consolidated and unconsolidated accounts. We only consider firms that report unconsolidated accounts. This involves discarding less than 1% of firms (that only report consolidated accounts).

The Amadeus database provides the balance sheet and the profit and loss account of firms, information is not available at plant or establishment level. Financial information is aggregated up to a format that is comparable across European countries. Bearing in mind cross-country differences (in Europe) in terms of accounting formats, detailed items that potentially are available for specific countries are not included in Amadeus (e.g. the social balance sheet in Belgium). In addition to the financial information the database also provides us with information on the firm's main activity, its location, its date of incorporation, its ownership structure, and its affiliate structure (if the firm has any). The database further provides us with an unique firm ID that allows to link firms across different versions (*cf. infra*). A firm ID consists of 2-digit country initials followed by a number which is typically a VAT number or a registration number.

[1] If some of the larger firms are missing, it could be due to the reporting standards in the country of incorporation. (e.g In the United Kingdom and Ireland limited companies are not legally required to report any form of accounts, allowing these to protect their finances from the public view.) Moreover, in this database we only look at unconsolidated accounts, which might overlook some of the larger corporations reporting their accounts in a consolidated way only.

2.2.2 Augmenting Amadeus using multiple issues

BvDEP regularly updates the Amadeus database. Aside from the continuously updated online version, a physical DVD/BLURAY of the database is released monthly. We use the following issues of the DVDs to create our data set: 60, 72, 84, 96, 108, 119, 132, 144, 156, 168, 180, 192, 204 (September issue of years 1998–2010) and issues 210 and 220. By making use of a 'time series' of DVDs we overcome a number of issues that arise from the use of a single issue of the database. One may identify the following three limitations of a single issue of the database.

The first limitation is that a single issue only includes at most the last ten available years of financial information for an individual firm. We start from financial information that is available from the most recent issue of the database (i.e. issue 220). We then work our way back to earlier years using information from previous versions, starting with the second most recent version and so on.[2] For any given financial item and calendar year our rule is to prefer information from an issue as recent as possible. This procedure allows us to obtain a maximal time span of 18 years for an individual firm (i.e. years 1995-2012).[3]

Second, firms that go out of business are dropped fairly rapidly from the Amadeus database.[4] Our time series approach therefore allows for better tracking of exit and short-lived entry over the period covered. Firms report their date of incorporation which we use to determine entry and a firm's age. We define the year of exit as *'sample exit'*, i.e. the last year a firm reports basic financial information without showing up in later years in the database. We take potential changes in firm IDs into account by applying the ID changes listed on BvDEP's dedicated website to earlier versions of the database.[5] Additionally, we check the help files of individual issues of Amadeus to control for systematic changes in firm IDs.[6] The updating of firm IDs from earlier

[2] The most recent issue is version 220 in the current version of our data set. However, we did use version 228 to fill out missing financial information for the year 2012 for firms not yet reporting balance sheet and profit and loss account in the 220 issue.

[3] We have 18,732,383 observations available for 3,649,965 firms to estimate total factor productivity (*TFP*). The average time span is 6.1 years.

[4] In recent versions of the database, a larger set of firms exiting the market more than two or three years earlier seems to be available for some countries. This is not the case for earlier versions of the database.

[5] In Germany, Austria and Italy a firm's ID will change if it moves to another region, while in Spain this will happen if the firm changes legal form.

[6] For example, in Belgium the firm ID is based on the VAT number. Recently, the administration added an additional zero in front of the existing 9-digit VAT numbers. For data retrieved from older issues we added the additional zero that was introduced in the official VAT number to the ID to have a comparable

versions avoids treating changes in firm IDs as exit. Moreover, aside from structural changes, a very small amount of ID changes recorded are due to mergers and acquisitions. Since we will update the ID of the firm to its most recent identifier, it is possible that these firms are dropped out due to a conflict in financial information belonging to the same identifier (e.g. acquirer and acquiree). Nonetheless, as according to Bureau van Dijk this happens very rarely, we do not keep track of these firms. Therefore, an exit in the data would only take into account those firms that stopped operating, due to bankruptcy or other non M&A related reason.

Third, BvDEP updates individual ownership links between legal entities rather than the full ownership structure of firms. The ownership information in a specific issue of Amadeus therefore often consists of a number of ownership links. A single issue of the database only contains the most recent information on ownership links and therefore does not allow to track changes in ownership structure.[7] Our 'time series' approach remedies this limitation as it allows us to construct a time series of foreign ownership. We use ownership information to separate foreign firms from domestic firms (in Amadeus the variables are the following: *'shareholder ID', 'shareholder name', 'shareholder direct %', 'shareholder total %', and 'shareholder country'*). We focus on direct shareholder links to determine whether a firm is foreign-owned or not.[8] BvDEP updates individual ownership links rather than the entire ownership structure, therefore each ownership link in a given issue has a reference date which may differ even up to a couple of years. Because ownership link information is updated irregularly, there is not necessarily for each firm-owner-*year* combination information available. Further, the BvDEP ownership manual suggests that the date of an ownership link is not always updated when it is verified at a later point in time. We therefore assume that all reported ownership information is valid at the moment when the specific version of Amadeus is issued and assign all ownership information of a given version (that sums to 100% in the vast majority of cases) to the year of the issue.[9] In line with a

ID across different issues. For Romania we detected an even more drastic change from the Chamber of Commerce number to the VAT number as a basis for a firm's ID.

[7]More recent versions do contain some history on ownership links, but not all the way back to the late 1990s.

[8]For part of the firms an ultimate owner is also recorded, but this is often only the case for larger firms. Quite often ultimate owners are individuals. For example, Lakshmi Mittal or the Mittal family are sometimes recorded as ultimate owner of Mittal steel affiliates in Europe.

[9]In Merlevede et al. (2014) we experiment with a dataset at the firm-*owner*-year-level for Romania with the available information on ownership *links* from Amadeus. There, we fill out missing firm-owner-

common definition applied by e.g. the OECD and the IMF, we require that at least 10% of shares are owned by a single foreign investor for a firm to be considered foreign.[10] Foreign owners are owners with a known nationality that differs from the host country nationality. If the country of origin is not known (in some countries not all small firms report ownership) we consider these owners as domestic. We also keep track of the percentage of shares owned by foreign firms. This allows us to separate majority from minority foreign-owned firms or to apply a more stringent definition and only consider firms as foreign if more than 50% of shares is foreign-controlled.

2.2.3 Industry classification

The raw Amadeus data provide us with a primary 4-digit code in the European NACE classification of activities. NACE stands for *"Nomenclature Statistique des Activités économiques dans les Communautés Européennes"*. Our time series approach to the database implies that we have annual industry codes for firms. We deal with potential variation in industry codes by creating different versions of our industry classification used in the data. If a firm reports the same code in the first and last year of information, we use that code throughout the entire period. For firms where only the fourth or the third (and fourth) digit are different between the first and last year's code we use the code of the last year. If a code is a clear outlier in the firm's time series, that code-year is ignored. For the remaining firms that show more 'bumpy' patterns (this is only a limited number of cases) we consider three alternatives. First we simply use the most recent code; second we use the most frequent code; and third we use the most recent code but allow for 'structural breaks'. A structural break is defined as a firm reporting two different codes with a single break and the less frequent code appearing in at least three years (versions) of the raw data. The first alternative serves as our basis, while the others are available for robustness tests. As indicated above, this issue comes up

year-entries under restriction that the full ownership structure cannot exceed 100%. In case of time gaps between entries for the same owner-firm combination but with a different share-size we assume that changes show up immediately in the database. We then fill out the gaps with the older information. In the end, this more elaborated but very cumbersome procedure (the majority of owners have no ID number and need to be matched by names that tend to show lots of small variations across versions) does reveal only marginal differences with our current approach. We therefore apply the more straightforward procedure of assigning all ownership information (i.e. the ownership structure) to the year of the issue from which the information is retrieved.

[10] A firm of which only 5% of shares are owned by one or more foreign firms is considered a domestic firm. We observe 29,208 ownership changes from domestic to foreign or vice versa.

only in a limited number of cases.

Our firms are classified according to revision 1.1 of the NACE nomenclature. Revision 2 of the classification became the standard classification near the end of our sample period. Because most of the firm-level information and most of our other data (*cf. infra*) refer to NACE revision 1.1, we opted to convert the industry codes of young firms at the end of the sample for which we only observe a NACE revision 2 code to revision 1.1 codes (because they are only included in the later issues of the database). For older firms we have a revision 1.1 code from the earlier versions. To convert the codes we use a conversion table obtained from Eurostat. In the conversion table most old codes match in multiple new codes (or vice versa). We deal with this issue in two alternative ways. One way is deterministic in the sense that we start from the available one-to-one matches and exclude these from multiple matches where these codes are mentioned. Transforming this into an iterative procedure results in a large number of one-to-one matches. For the remaining many-to-many matches we obtain a single match by preferring manufacturing over agriculture and services. Our second approach randomly matches a revision 2 code with one of the possible revision 1.1 matches from the conversion table. Random matching is performed firm by firm, not industry by industry. The deterministic approach serves as our basis, while the randomisation is available as robustness check. Note that for the vast majority of firms we have an original NACE revision 1.1 code from the Amadeus database.

Although a 4-digit code is available, we mostly rely on 2-digit (or slightly more aggregated) industry classifications for practical implementation (*TFP* estimations for example). In Table 2.A.1 in the Appendix a list of the industries used is provided.

2.2.4 Location of firms: Region classification

In *AUGAMA* regions are defined using the EU's NUTS-classification ('Nomenclature of territorial units for statistics') which is a hierarchical system for dividing the economic territory of the EU. In Table 2.A.3 in the Appendix we list the main criteria and an example of the NUTS-structure. Firms are assigned to a region on the basis of their zip code, which is available in Amadeus. Eurostat provides tables mapping zip codes into NUTS regions for most countries. For other countries we rely on national data sources (the Eurostat website provides contact details). When a zip code is not available for a firm, we try to map the '*region*'-variable reported in Amadeus to the NUTS-classification. As with the industry classification above, our time series approach to the database implies that we have annual zip codes for firms. We deal with potential time variation by creating different versions of the regional classification. If a firm reports the same zip code in the first and last year of information, we use that code throughout the entire time period. When a code is clearly a one-time outlier in the firm's time series, that code-year is ignored and replaced. For the limited number of firms that change zip code, we consider three alternatives. First we simply use the most recent zip code to assign the firm to a region; second we use the most frequent zip code; and third we allow a 'structural break' for firms reporting two different zip codes with a single break that implies that the less frequent code appears in at least three years of the raw data. The first alternative serves as our basis, while the others are available for robustness tests. As with the industry classification above, the large majority of firms consistently reports the same zip code throughout the different versions.

We account for changes in the NUTS classification itself by means of conversion tables retrieved from the Eurostat website. In our data we use the 2006 vintage of the NUTS-classification. Where other data sources use a different version of the NUTS classification (e.g. the Cambridge Econometrics data described below), we reclassify the NUTS codes in the data. In most cases, codes change because of slight modifications of the area definition, such as border shifts. Since these shifts are small and unlikely to have important economic consequences, we rename the new code back to the old, and merge with our data. In other cases, regions are split or merged. Where a one-to-one correspondence between codes is straightforward to establish, we do so. Where it is not (as for example in the case where two old regions are split into more than two new regions), we distribute data values proportionally over those regions. In most cases, regions are redefined at the NUTS 3 level, and have no impact at the NUTS 2 distribution. Moreover, since such changes always involve bordering regions, regional differences are never very pronounced. Eurostat discusses changes in the NUTS classification in its publications on 'Regions in the European Union' (e.g. European Commission, 2011), and provides spreadsheets for changes between differences NUTS vintages at its website.[11]

[11] `http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/history_nuts`

2.2.5 Deflation and currency

The data retrieved from Amadeus are downloaded in units of national currency. In order to make cross-country comparisons, these data are converted to Euro. Because our price deflators refer to national currency, we first deflate the data in national currency to obtain unit equivalents and then convert them to Euro using 2005 exchange rates.[12] By making use of the 2005 exchange rate we avoid the exchange rate movements driving our results in cross-country comparisons (see Gal, 2013). We choose 2005 as reference year because it corresponds to the middle of our sample and we prefer to avoid the later years when most of European economies experienced the the Global Financial Crisis. We believe this choice provides the least unbiased TFP estimates. For Euro-zone countries earlier data in the old national currency are converted using the Euro-entry conversion rate. For countries adopting the Euro more recently (e.g. Slovenia) financial information dating before Euro adoption was converted to Euro by BvDEP using concurrent exchange rates. These data were converted by multiplying them with the ratio of the Euro-entry conversion rate and the concurrent exchange rate.

Our main data source for output deflators is the EU KLEMS database. These deflators have been incorporated and updated by Eurostat. We use EU KLEMS data up to the year 2005 and then continue with Eurostat data. For the last three years of our sample NACE revision (rev.) 1.1 price deflators are no longer available (nor are NACE rev. 2 price deflators for the earlier years). We therefore apply the percentage change of a corresponding NACE rev. 2 series to the later years of the NACE rev. 1.1 series (both revisions do report similar broad categories such as e.g. food processing). We define our capital deflator as the average of the following five NACE rev. 1.1 industries: machinery and equipment (29); office machinery and computing (30); electrical machinery and apparatus (31); motor vehicles, trailers, and semi-trailers (34); and other transport equipment (35) (see Javorcik, 2004). We calculate an intermediate input deflator as a weighted average of output deflators where country-time-industry-specific weights are based on intermediate input uses retrieved from input-output tables. We obtain most input-output tables from Eurostat.[13] For countries for which Eurostat does not yet provide input-output tables, we use input-output tables from the World Input-Output

[12] For comparison with Eurostat's Structural Business Statistics (SBS) database ($cf.\ infra$), data in national currency are converted to Euro with annual exchange rates rather than 2005 fixed exchange rates.

[13] For most countries we have IO-tables for 1995, 2000, and 2005.

Database which are slightly more aggregated in terms of industries (Eurostat tables are at NACE 2-digit level) (Timmer, 2012).

Value added is double deflated, i.e. real value added is calculated as output deflated with an output deflator minus intermediate use deflated with the calculated intermediate input deflator.[14] Note that making use of industry-level deflators has some implications for our measure of total factor productivity (*cf. infra*).

2.2.6 Variable definitions

We define the following variables. Output Y is measured as *'operating revenue turnover'*, real output y is obtained by deflating Y by producer price indices of the appropriate NACE industry (*cf. supra*). Value added VA is defined as output minus intermediates M, i.e. operating revenue minus *'material costs'* (from the Amadeus database).[15] Real value added va is double deflated and defined as real output y minus real material costs m. The latter are defined as material costs deflated by the intermediate input deflator defined above. Labour L is the *'number of employees'* (end-of-period). Capital K is measured by *'tangible fixed assets'*, real capital k is obtained by deflating K by the capital deflator defined above. The age of a firm is calculated on the basis of its *'date of incorporation'*. We have information on the number of months accounts refer to. We use this information to convert flow variables (operating revenues, material costs, and thus value added) to twelve month equivalents as far as the number of months is not below 6 or above 24. Outside these boundaries variables are set to missing. End-of-period variables such as tangible fixed assets and the number of employees are unchanged. The number of non-12 months accounts is very small and generally below 0.5% for a country-industry-year cell. We define labour productivity as operating revenues divided by the number of employees and estimate a measure of total factor productivity (*TFP*). We prefer total factor productivity to labour productivity as the latter does not control for intermediate inputs usage and capital intensity differences across firms (Gal, 2013).

The strength of *AUGAMA* is that it provides information for firms from many

[14] For Croatia we do not have detailed prices, nor IO-tables.

[15] Amadeus does contain value added figures for some countries that are either obtained directly from the data-provider or are calculated using an accounting definition, but gauging from the manual, the definition differs across countries.

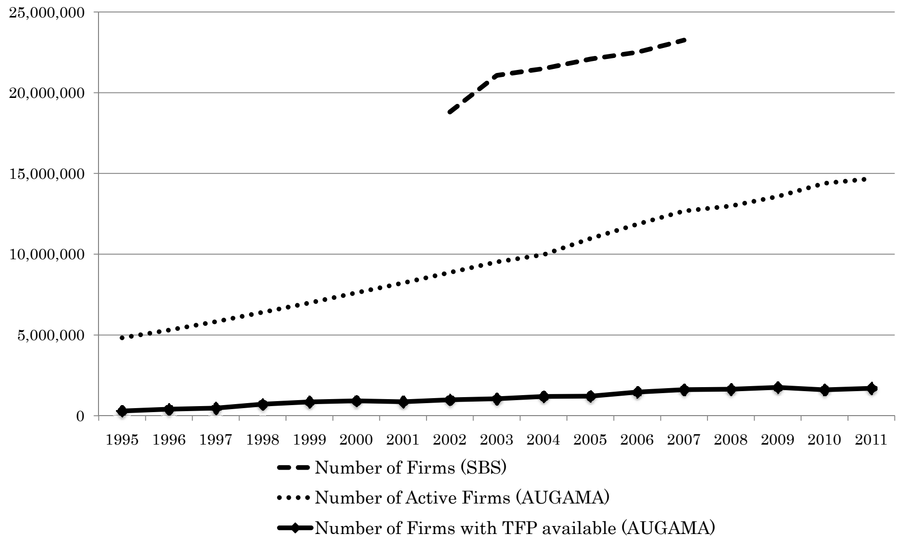
countries and industries for more than 15 years including the Great Recession and Euro crisis period. These are important features as the CompNet Task Force notes that *"firm-level analysis in Europe is hampered by a lack of sufficient and comparable data across countries"* (CompNet, 2014). There are however some caveats one needs to bear in mind when interpreting results obtained using *AUGAMA* in empirical analysis. Regarding the measurement of real output, we have no data on output quantities but can only observe output expressed in terms of revenue. As indicated above, output is deflated with industry-level price deflators as we do not have firm-level price deflators. This implies that we are only able to consider *TFPR* and not *TFPQ* (total factor productivity in terms of revenue rather than quantity). According to Syverson (2011) this approach is satisfactory when differences in prices only reflect differences in product quality. When differences in prices also reflect differences in market power, measured efficiency/productivity of firms will also reflect market power (Syverson, 2011). On the input side we have information on the total number of employees, but not on the total number of hours worked nor on other employee characteristics, such as skill levels.[16] Estimated productivity levels should therefore be thought of as including labour quality and capacity utilisation (Gal, 2013). The Amadeus database provides data on the total stock of *'tangible fixed assets'*, but more detail is not available. Changes in (capital) capacity utilisation can thus not be accounted for. Note that these issues are not specific to our data set, but are faced by many micro-level data and studies (e.g. CompNet, 2014, or Gal, 2013).

2.2.7 Basic data cleaning

First, negative values of the number of employees, tangible fixed assets, operating revenue, sales, material costs, and value added (defined as the difference between turnover and material costs) are set to missing. We then calculate growth rates of the aforementioned variables and replace observations associated with growth rates below

[16] Total wage costs, *'costs of employees'*, are reported in Amadeus, and in principle could be used as a quality adjusted labour input. The variable is filled out less frequently, however, and more importantly it is also prone to cross-country differences in the regulatory framework (e.g. social security contributions). Therefore it is likely to be a good reflection of actual labour costs, but cross-country comparisons in terms of 'quality' are not recommendable. Further, it is not always clear from the Amadeus manual whether the definition of *'costs of employees'* is similar across countries (e.g. whether management compensation is included or not). For this reason, we restrict ourselves to the number of employees as labour input.

Figure 2.1: Number of firms in Europe according to Eurostat SBS (Structural Business Statistics) and *AUGAMA* (1995–2011)



Notes: *AUGAMA*: all active firms and all firms with *WLP-TFP* available are displayed. Countries included are AT, BE, BG, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LT, LV, NL, NO, PL, PT, RO, SE, SI, and SK (HR is not available in SBS).

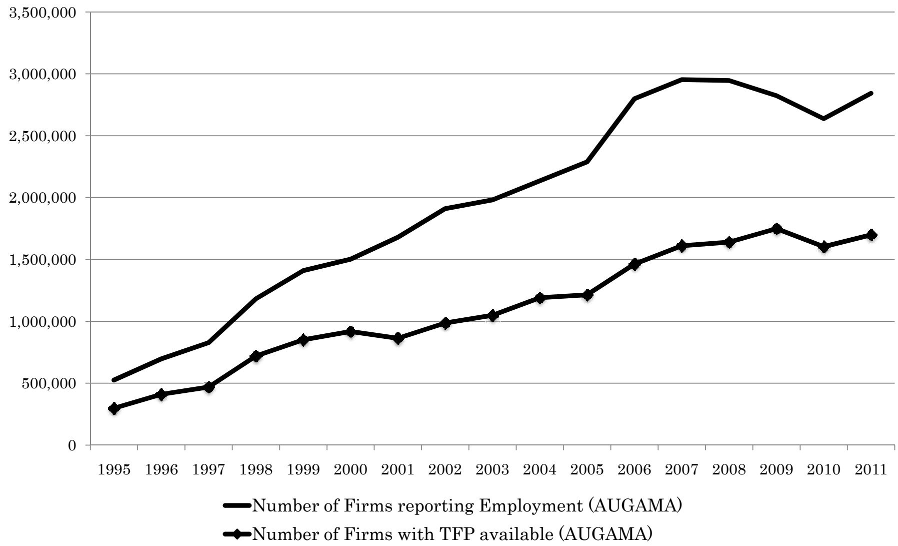
(above) the 1st (99th) percentile with missing values (cf. CompNet (2014)). Further cleaning, e.g. after *TFP* estimation, is done in function of specific applications of the data for specific research questions.

2.3 Representativeness

In this Section we discuss *AUGAMA* in terms of coverage and representativeness. Our main comparison base is the Structural Business Statistics database (SBS) provided by Eurostat.[17] SBS data in NACE revision 1.1 are available for the period 2002–2007 (2003–2007 for some countries). We use this period to infer the representativeness of our data set. We consider firms in the 'business economy' (mining, manufacturing, construction, and services, excluding financial services), i.e. NACE 2-digit codes 10 to 74, excluding 65 to 67. Table 2.A.1 in the Appendix lists all the 2-digit industries

[17] We think of the SBS data as providing the population of firms. Do note that SBS data are retrieved from surveys with incomplete coverage of the population of firms for some countries, which might result in ratios above 100% (CompNet, 2014). Furthermore, it is not always clear whether SBS data consider only companies or a larger set of business entities which also includes sole proprietors. Nonetheless, we think a comparison is warranted since its statistical unit, the enterprise, is defined as "an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resource" (see European Council (1993)), thus providing a good estimate of European economic activity.

Figure 2.2: Number of firms reporting employment and number of firms with *WLP-TFP* in *AUGAMA* (1995–2011)



Countries included are AT, BE, BG, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LT, LV, NL, NO, PL, PT, RO, SE, SI, and SK (HR is not available in SBS).

included in *AUGAMA*. In order to get a first broad overview Figure 2.1 plots the number of firms over time.[18] The SBS data count more than 20 million firms in Europe (some countries do not report numbers in 2002, hence the *'jump'* in 2003). We see a gradual increase in the number of active firms recorded in Amadeus from about five to fifteen million firms from 1995 tot 2011. In 2007 this accounts for 55% of SBS firms, up from 45% in 2003. These numbers are smaller when we only count firms that report the variables necessary to obtain a measure of total factor productivity based on the Wooldridge-Levinsohn-Petrin estimator, *WLP-TFP* (see Wooldridge, 2009 and Petrin and Levinsohn, 2012; $cf.\ infra$). From Figure 2.2 one can infer that the number of firms for which we are able to estimate TFP also steadily increases to more than 1.5 million observations from 2007 onwards (from 5% (2003) to 6.9% (2007) of the number of SBS firms). The number of firms that at least report employment is considerably higher (12.7% of SBS firms in 2007).

Table 2.1 shows further numbers illustrating representativeness in columns two to five. The entries in Table 2.1 are based on *AUGAMA* corrected for outliers following the procedure described above. The percentages shown are averaged over industry-time cells by country. A country-industry-time cell in this case is defined by the host country,

[18]Countries included are: AT, BE, BG, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LT, LV, NL, NO, PL, PT, RO, SE, SI, and SK. HR is not reported in SBS.

Table 2.1: *AUGAMA* versus Eurostat Structural Business Statistics (SBS): Representativeness

	SBS 2002–2007	AUGAMA as share of SBS 2002–2007				SBS 2007		AUGAMA 2007	
						share of firms in		share of firms in	
	# firms	# firms	# empl.	turnover	costs of empl.	manufacturing	services	manufacturing	services
AT	269,426	11.0%	46.5%	40.8%	34.6%	10.0%	90.0%	22.2%	77.8%
BE*	333,564	42.3%	85.0%	81.3%	87.7%	9.1%	90.9%	14.2%	85.8%
BG	221,116	17.6%	60.3%	83.5%	66.4%	11.8%	88.2%	15.2%	84.8%
CZ	871,067	11.4%	74.7%	76.0%	72.5%	16.9%	83.1%	19.7%	80.3%
DE	1,714,904	14.6%	32.4%	42.0%	36.1%	11.0%	89.0%	21.0%	79.0%
DK	198,369	18.4%	41.9%	36.7%	45.8%	8.7%	91.3%	14.2%	85.8%
EE	38,270	86.9%	98.6%	97.7%	62.1%	12.2%	87.8%	16.2%	83.8%
ES	2,499,620	36.9%	71.3%	75.1%	72.1%	7.9%	92.1%	17.2%	82.8%
FI	186,972	28.2%	49.1%	47.1%	39.9%	12.0%	88.0%	17.3%	82.7%
FR	2,158,887	23.6%	62.1%	63.9%	62.9%	10.2%	89.8%	14.0%	86.0%
GB	1,571,916	10.0%	80.0%	65.8%	67.1%	8.9%	91.1%	22.0%	78.0%
GR	694,183	12.7%	52.3%	50.4%		11.2%	88.8%	27.2%	72.8%
HU*	551,119	8.5%	35.2%	38.7%	35.7%	10.4%	89.6%	19.2%	80.8%
IE	87,175	12.6%	26.0%	31.2%	8.6%	4.7%	95.3%	18.8%	81.3%
IT	3,790,324	15.0%	55.8%	58.8%	57.8%	13.1%	86.9%	30.7%	69.3%
LT	88,187	22.8%	52.5%	61.9%		11.6%	88.4%	20.4%	79.6%
LV	60,581	18.7%	54.9%	57.1%	7.0%	11.1%	88.9%	16.5%	83.5%
NL**	497,613	5.2%	54.5%	42.9%	39.0%	8.6%	91.4%	18.4%	81.6%
NO	198,926	38.6%	72.5%	65.8%	72.5%	12.0%	88.0%	11.9%	88.1%
PL	1,452,512	6.5%	46.8%	50.3%	27.9%	13.3%	86.7%	30.9%	69.1%
PT	711,778	34.0%	30.9%	34.9%	33.6%	10.9%	89.1%	16.0%	84.0%
RO	389,286	67.6%	87.7%	36.1%	34.6%	12.6%	87.4%	15.8%	84.2%
SE	514,925	32.2%	64.5%	78.6%	79.2%	10.9%	89.1%	13.6%	86.4%
SI	91,065	24.2%	80.8%	80.2%	79.6%	17.8%	82.2%	30.6%	69.4%
SK	42,525	40.7%	78.4%	89.3%	89.2%	14.3%	85.7%	21.4%	78.6%

For BE and HU the period is 2003–2007; ** For NL some outlier firms in levels were removed after manual inspection. Firms are assigned to a country-year-industry cell, for which aggregates are calculated. The ratio of each *AUGAMA* cell with the corresponding SBS cell is determined. Table entries are then obtained by averaging over industry-year. Only firms that report the indicated variable are considered in the analysis. This results in differences in coverage across countries , but also across variables within countries.

Table 2.2: *AUGAMA* versus Eurostat Structural Business Statistics (SBS): Firm Size Distribution

	SBS (2002–2007 average)				AUGAMA (firms with employment) (2002–2007 average)				AUGAMA (firms with TFP) (2002–2007 average)			
	1–19	20–49	50–249	250+	1–19	20–49	50–249	250+	1–19	20–49	50–249	250+
AT	94.3%	3.7%	1.7%	0.3%	52.1%	21.6%	20.4%	5.9%	17.8%	14.4%	42.8%	25.1%
BE	96.3%	2.5%	0.9%	0.2%	87.6%	8.1%	3.5%	0.7%	36.3%	30.2%	27.5%	6.0%
BG	94.9%	3.1%	1.7%	0.3%	75.3%	12.3%	10.0%	2.4%	60.4%	19.2%	16.6%	3.9%
CZ	97.5%	1.4%	0.8%	0.2%	69.1%	14.6%	13.0%	3.3%	65.8%	15.8%	14.7%	3.8%
DE	92.6%	4.4%	2.3%	0.5%	69.3%	15.6%	12.2%	3.0%	40.7%	19.5%	28.9%	10.9%
DK	93.6%	4.2%	1.9%	0.3%	80.2%	11.8%	6.6%	1.4%				
EE	91.0%	5.7%	2.9%	0.4%	86.1%	8.8%	4.5%	0.6%	85.8%	9.2%	4.4%	0.6%
ES	96.7%	2.3%	0.8%	0.1%	86.6%	9.5%	3.4%	0.5%	84.7%	10.8%	3.9%	0.6%
FI	96.2%	2.3%	1.2%	0.3%	88.9%	7.2%	3.2%	0.7%	88.3%	7.7%	3.3%	0.6%
FR	96.2%	2.5%	1.0%	0.2%	84.7%	9.7%	4.6%	1.1%	83.4%	10.4%	4.9%	1.2%
GB	94.6%	3.3%	1.7%	0.4%	40.3%	18.3%	31.0%	10.5%				
GR	98.5%	0.8%	0.3%	0.1%	68.1%	21.7%	9.1%	1.2%				
HU	96.0%	1.5%	0.7%	0.1%	74.1%	13.1%	10.0%	2.9%	73.4%	13.4%	10.2%	2.9%
IE	91.2%	5.1%	2.7%	0.5%	42.8%	22.4%	28.6%	6.2%				
IT	98.2%	1.3%	0.5%	0.1%	75.6%	15.2%	8.1%	1.1%	74.3%	16.0%	8.5%	1.1%
LT	93.2%	4.6%	2.6%	0.3%	52.1%	24.8%	19.8%	3.2%				
LV	91.8%	5.2%	2.7%	0.4%	64.2%	19.2%	14.4%	2.2%	47.1%	16.9%	27.7%	8.4%
NL	94.7%	3.4%	1.6%	0.3%	43.6%	19.1%	30.2%	7.0%	55.8%	16.6%	21.5%	6.1%
NO	99.0%	2.7%	1.2%	0.2%	87.6%	8.4%	3.3%	0.6%	85.8%	9.7%	3.8%	0.7%
PL	97.6%	1.3%	0.9%	0.2%	35.9%	21.7%	33.3%	9.2%	36.2%	22.9%	32.7%	8.2%
PT	97.1%	1.8%	0.8%	0.1%	90.7%	6.0%	2.8%	0.5%	89.2%	7.0%	3.3%	0.5%
RO	93.6%	3.7%	2.3%	0.5%	88.9%	6.4%	3.9%	0.8%	88.1%	6.8%	4.1%	0.9%
SE	97.2%	1.8%	0.8%	0.2%	92.1%	5.2%	2.2%	0.5%	91.4%	6.1%	2.2%	0.3%
SI	96.2%	2.1%	1.3%	0.3%	78.0%	10.8%	8.5%	2.7%	76.5%	11.6%	9.1%	2.8%
SK	89.0%	4.9%	4.7%	1.1%	66.2%	11.4%	17.3%	5.2%	63.1%	12.0%	19.1%	5.8%

Comparison of the share of firms that employ 1 to 19 employees, 20 to 49 employees, 50 to 249 employees or at least 250 employees in SBS and AUGAMA. The share of firms that fall within a specific size category is based on the average number of firms in period 2002–2007 for the SBS and AUGAMA data. In the case of AUGAMA, only firms for which the number of employees or *WLP-TFP* are available are considered.

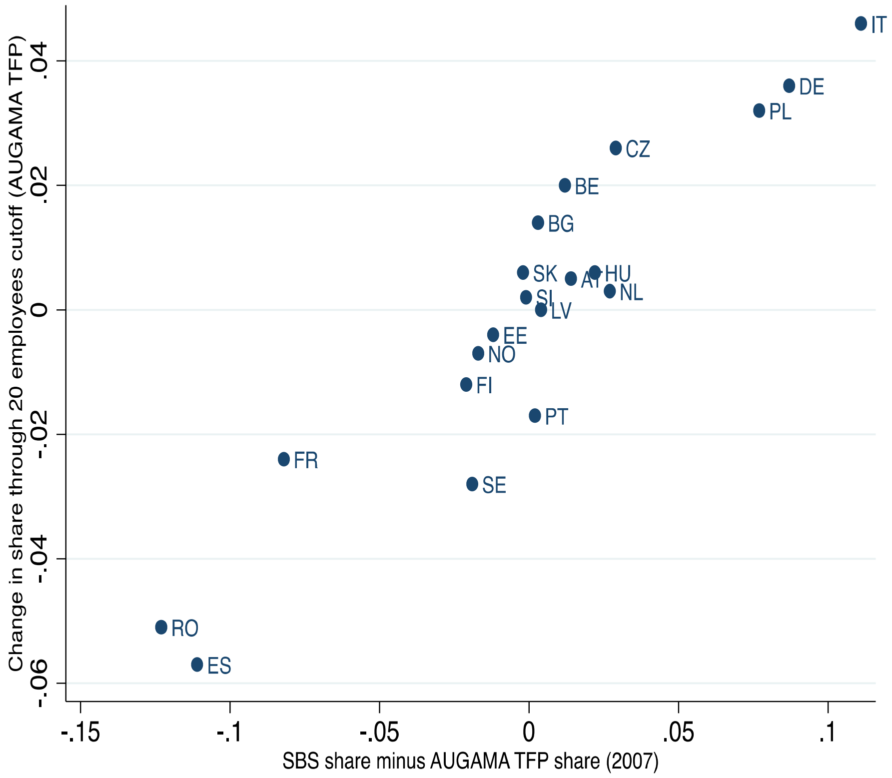
a broad NACE category (SBS does not report finer detail, see Table 2.A.1), and the year of observation. Based on the information from *AUGAMA*, a firm is assigned to a cell. After assigning firms to cells we calculate cell aggregates (total number of firms, employees, total turnover, and total labour costs) and create the ratio with the corresponding aggregate from the SBS data. The numbers in Table 2.1 are obtained by averaging over industry-time cells by country. Note that calculations are based on all firms that report the indicated variable and that 'coverage' in this respect may differ not only between countries, but also within countries across variables. Table 2.1 reveals that coverage in terms of the number of firms ranges from a low of 5.2% for the Netherlands to a high of 86.9% in Estonia. On average over countries we observe a quarter of SBS firms. The coverage in terms of total employment, labour costs, and turnover is higher and indicates that *AUGAMA* typically includes larger firms. Averaged over countries, *AUGAMA* accounts for about 60% of employment and turnover and 53% of wage costs. The last four columns of Table 2.1 reveal that our data set is slightly biased towards manufacturing firms in comparison to what is reported in SBS statistics (count of firms), but the discrepancy falls within very reasonable margins.[19] Table 2.2 shows the distribution of firm size according to both SBS data and *AUGAMA*. *AUGAMA* is generally biased towards larger firms (especially firms with between 20 and 249 employees). In most countries this bias increases when we focus on firms for which *WLP-TFP* is available, but not to a large extent.[20]

The dispersion across countries in representativeness in terms of the number of firms found, and to a lesser extent in terms of the number of employees, turnover and wage costs (see Table 2.1) is mainly due to differences in the coverage of small firms. When we focus on a sample of firms with on average more than 20 employees (as in e.g. Gal, 2013 and CompNet, 2014), the sample becomes more balanced in cross-country terms. Figure 2.3 plots the difference between a country's share in the number of firms according to SBS data and its share in the number of firms according to *AUGAMA* in 2007, against the resulting change in the share in the number of firms according to *AUGAMA* when a cut-off of at least 20 employees (on average by firm) is applied

[19] To account for these firm, industry and country misalignments from the population a reweighing of the sample could be used, by which under-represented firms are given a larger weight.

[20] For AT, BE, DE, and LV we record larger changes. In Belgium, for example, this is due to the fact that most smaller firms file an abridged account and are not obliged to report material costs.

Figure 2.3: The impact of a 20 employee cut-off on the distribution of firms across countries: Eurostat Structural Business Statistics (SBS) and *AUGAMA* in 2007

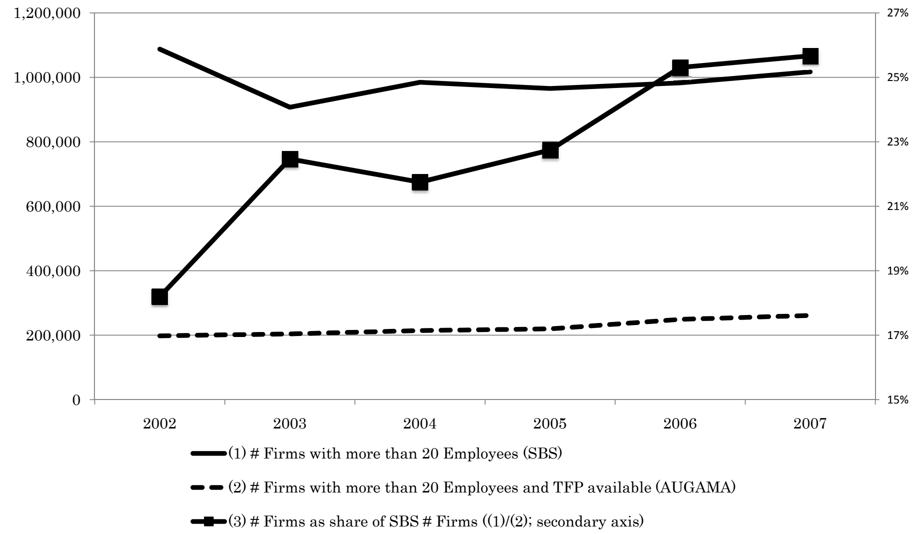


Difference between a country's share in number of firms in SBS compared to *AUGAMA* plotted against the change in a country' share in *AUGAMA* when a 20 employee cut-off is applied.

to the data. There is a clear relation between a country's share falling short of the SBS share (more to the right on the horizontal axis) and an increase in its share when applying the cut-off (more to the top on the vertical axis). This brings the cross-country distribution closer to the SBS data. Figure 2.4 shows the number of firms with more than 20 employees for SBS and the number of firms with more than 20 employees and *WLP-TFP* available for *AUGAMA*.[21] By 2007 our TFP-sample accounts for a quarter of the number of firms with more than 20 employees. Tables 2.A.5 and 2.A.6 in the Appendix list the annual number of observations in the TFP sample for all firms and foreign firms separately.

[21] DK, GB and IE are not included in the figure because one of the variables for TFP-calculation is missing for all firms. GR and LT are included but only have a very small sample of TFP firms, either limited in the time dimension (GR) or limited in coverage (LT).

Figure 2.4: Number of firms when applying a 20 employee cut-off in Eurostat Structural Business Statistics (SBS) and *AUGAMA* (2002–2007)



AUGAMA: only firms with *WLP-TFP* available are considered. Ratio of number of firms in *AUGAMA* and SBS displayed on secondary axis on the right.

2.4 Representativeness of economic activity across industries and regions

In this Section we focus on representativeness of *AUGAMA* in terms of the distribution of economic activity across industries and across regions (within individual countries and within 'Europe'). Representativeness across industries is presented as the correlation of total industry activity recorded in *AUGAMA* with the total activity reported in SBS statistics. Industries are defined as 21 'broad' industries grouping sometimes several NACE 2-digit industries (see Table 2.A.1 in Appendix) as this is the level of aggregation reported in the SBS statistics. The period considered is generally 2002–2007, which is determined by the availability of SBS statistics in NACE revision 1.1. As indicators of activity we consider output, employment, wage costs, and material costs. Table 2.3 contains the pairwise correlations for all countries in the sample. The correlations are always positive and statistically significant. The majority of correlations is well above 0.75. This especially holds for output and employment. For wage costs and material costs correlations are generally somewhat lower, but still statistically significant. Table 2.3 clearly indicates that the distribution across industries found in *AUGAMA* is very well aligned with the distribution derived from SBS statistics.

We compare the regional distribution of economic activity found in *AUGAMA*

Table 2.3: Pairwise correlation coefficients of economic indicator totals for broad industries as calculated from *AUGAMA* and recorded in Eurostat SBS

	Output	Employment	Wage costs	Material costs
AT	0.76	0.87	0.65	0.64
BE	0.97	0.97	0.79	0.98
BG	0.94	0.87	0.82	0.88
CZ	0.97	0.98	0.82	0.95
DE	0.92	0.91	0.77	0.93
DK	0.92	0.94	0.81	-
EE	0.96	0.91	0.64	0.82
ES	0.98	0.99	0.71	0.99
FI	0.77	0.88	0.74	0.55
FR	0.95	0.96	0.84	0.91
GB	0.97	0.94	0.69	-
GR	0.80	0.84	-	-
HR	-	-	-	-
HU	0.75	0.81	0.74	0.72
IE	0.81	0.90	0.26(a)	-
IT	0.86	0.96	0.82	0.67
LT	0.81	0.83	-	-
LV	0.97	0.92	0.41	0.54
NL	0.80	0.71	0.50	0.83
NO	0.94	0.98	0.63	0.93
PL	0.90	0.95	0.68	0.83
PT	0.75	0.51	0.50	0.65
RO	0.98	0.99	0.88	0.89
SE	0.97	0.98	0.85	0.65
SI	0.96	0.98	0.71	0.96
SK	0.95	0.89	0.88	0.90

Period 2002-2007, except for BE and HU 2003-2007. All correlations are statistically significant at 5% level except (a) where 0.26 is not significant (if we exclude the year 2002, we obtain a statistically significant correlation of 0.47).

with the regional distribution we derive from the Cambridge Econometrics Regional Database.[22] We consider both within-country and Europe-wide regional representativeness of economic activity. The Europe-wide regional distribution might be affected by differences in coverage across countries (see Table 2.1 above) while the within-country regional distribution is not. For country-by-country or country-specific analysis within-country distribution comparisons are relevant, whereas for Europe-wide regional analysis the Europe-wide distributions' comparison should be considered. Economic

[22]This database has been used in academic research by Becker et al. (2010, 2012), among others.

activity is measured by the total number of employees and the total amount of generated value added in a specific region. Because before we detected a tendency towards better representativeness of larger firms in countries with a lower overall coverage in terms of the total number of firms, we may expect the distribution of economic activity to be less affected as larger firms will account for the bulk of economic activity.

We consider two samples of industries: i) the business economy, i.e. NACE 2-digit codes 10 to 74 (see Appendix), and ii) 'broad' manufacturing, i.e. NACE 2-digit codes 10 to 42. In case of the former Cambridge Econometrics data also cover financial services (NACE 2-digit codes 65 to 67) which we do not consider in our data (*cf. supra*). As far as the financial services sector's regional distribution is more or less in line with the distribution of other activity this should not have sizeable implications. Nevertheless, one should bear this in mind interpreting comparisons below. Therefore we also consider a more narrow definition labelled 'broad' manufacturing[23] which is the most detailed level provided in the Cambridge Econometrics database.

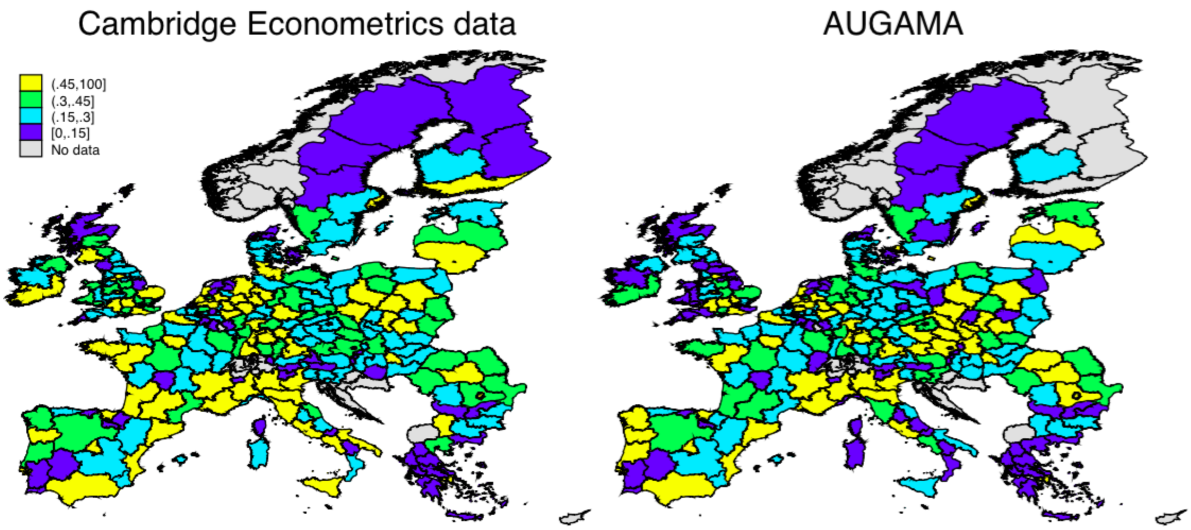
As a first indication we plot a map for the year 2005 in Figure 2.5 representing NUTS 2-digit regions' share in total *Europe-wide* employment according to the Cambridge Econometrics database (left panel) and *AUGAMA* (right panel). In Germany, the Netherlands, and the UK the shares of the NUTS 2-digit regions seem smaller for *AUGAMA* than for the Cambridge Econometrics data. In Figure 2.6 we plot similar maps for German NUTS 3-digit regions' share in total *German* broad manufacturing employment. The comparison of both panels suggests that *AUGAMA* very reasonably approximates the distribution of German economic activity for this sample.

In Tables 2.4 and 2.5 we quantify the information visualised in the maps by calculating the amount of economic activity (employment) that needs to switch region to align the distribution of regional economic activity obtained from *AUGAMA* with the distribution according to the Cambridge Econometrics data. In Table 2.4 numbers are obtained as the sum of absolute values of a region's share in total European employment according to the Cambridge Econometrics data minus its share according to *AUGAMA* divided by two.[24] Over time the number is quite stable. For the sample

[23] 'Broad' manufacturing includes Mining (10-14) and Utilities (41-42) in addition to 'narrow' pure manufacturing (15-37).

[24] There are a few regions where *AUGAMA* records zero (no) activity. Deleting or retaining these regions from/in both databases prior to calculation does not affect conclusions.

Figure 2.5: NUTS 2-digit regions' share in total European employment in 2005 according to the Cambridge Econometrics database and *AUGAMA*



Countries included: AT, BE, BG, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LT, LV, NL, PL, PT, RO, SE, SI, and SK.

Figure 2.6: NUTS 3-digit regions' share in total German employment in 2005 according to the Cambridge Econometrics database and *AUGAMA*

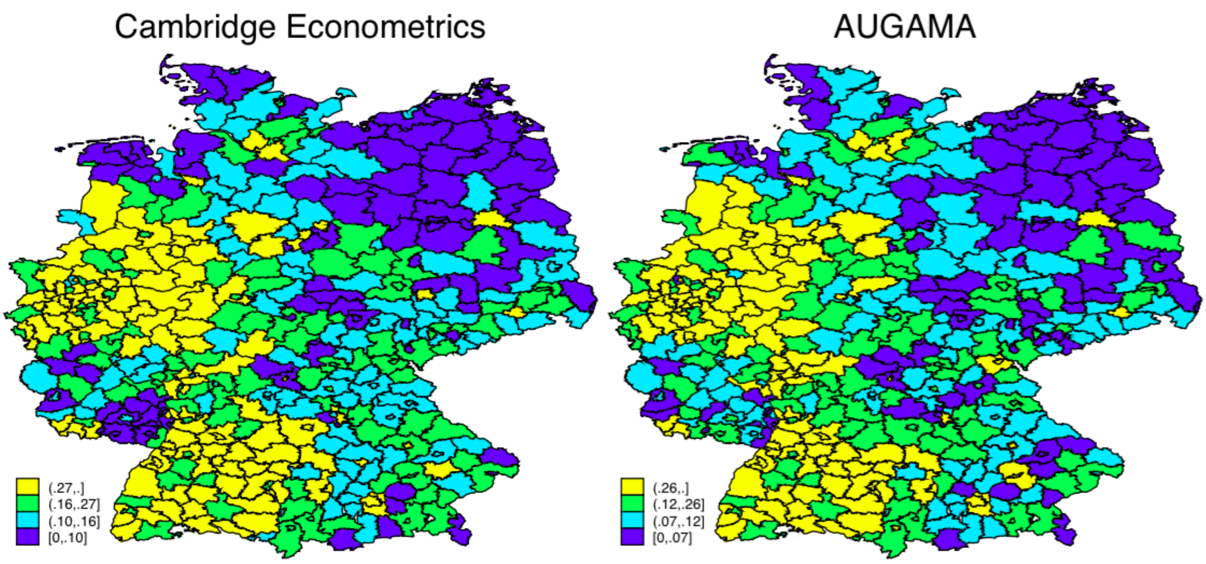


Table 2.4: Annual share of European employment that needs to switch region for the *AUGAMA* regional distribution to equal the Cambridge Econometrics distribution

	All available		*TFP Sample*	
	Business economy	Broad manufacturing	Business economy	Broad manufacturing
2003	24.5	22.1	28.3	28.1
2004	23.6	21.5	28.6	27.4
2005	22.9	20.7	27.2	25.4
2006	20.4	18.6	24.1	22.5
2007	19.6	17.2	25.0	23.1
2008	20.5	17.5	25.4	23.3
2009	23.2	19.6	26.3	24.0
2010	21.6	18.6	27.9	25.7

Numbers are obtained as the sum of absolute differences between Cambridge Econometrics and *AUGAMA* region shares divided by two.

of *TFP* firms about 25% of employment needs to switch regions across Europe for the distributions obtained from the Cambridge Econometrics data and *AUGAMA* to be equal. This suggests that to a large extent the distribution of activity in *AUGAMA* follows the distribution of overall economic activity. For the 'broad manufacturing' industries numbers are generally smaller than for the business economy sample. Table 2.5 suggests that the bulk of this number refers to cross-country movement of employment. Table 2.5 lists the amount of activity that needs to switch regions *within* a country to match both distributions for the year 2005. For most countries this number is much smaller than the average in Table 2.4. Only for Bulgaria, the Netherlands, and Portugal the number is higher; only in Bulgaria this is the case for all samples considered.

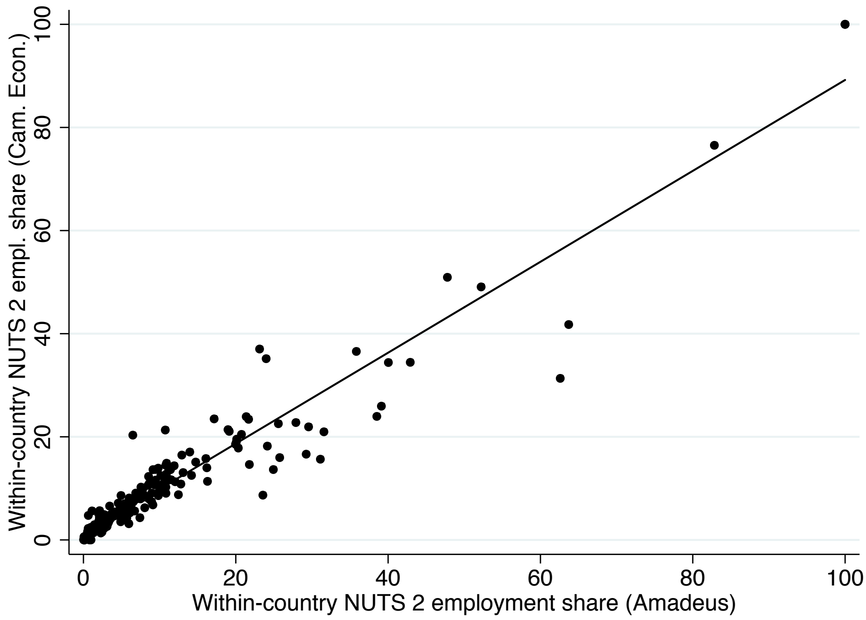
Figure 2.7 is a scatter plot of within-country shares in employment of NUTS 2-digit regions according to the *AUGAMA* database (horizontal axis) and the Cambridge Econometrics database (vertical axis) for the year 2005. The Figure suggests a high correlation between both measures. This is confirmed in Table 2.6 where pairwise correlations (also for 2005) are listed for NUTS 3-digit regions' share in within-country economic activity (measured either as employment or value added). Using all regions in all countries (i.e. row headed with 'Europe') the correlation varies between 0.71 and 0.87 according to the sample and activity measure. The correlation is always statistically significant. If we calculate the correlation for the 24 countries as a whole (in the NUTS classification these can be thought of as *'0-digit'* regions) we obtain a

Table 2.5: Share of employment that needs to switch region within-country for the *AUGAMA* regional distribution to equal the Cambridge Econometrics distribution

	All available		*TFP Sample*	
	Business economy	Broad manufacturing	Business economy	Broad manufacturing
AT	14.5	10.3	20.5	17.5
BE	12.0	9.8	15.9	12.2
BG	34.6	27.0	36.4	29.7
CZ	12.6	7.0	12.4	7.1
DE	20.1	19.9	23.2	26.2
EE	15.1	14.9	14.9	17.1
ES	11.9	8.0	10.8	8.0
FI	1.7	1.6	1.5	1.6
FR	17.5	23.3	13.6	23.0
HU	9.1	12.1	11.2	11.8
IT	20.2	17.0	20.0	17.1
LV	13.7	9.2	17.6	17.5
NL	6.7	15.6	26.9	30.8
PL	15.0	11.2	19.5	15.9
PT	31.4	24.5	25.6	24.5
RO	18.7	15.8	18.0	15.7
SE	13.2	8.2	7.0	8.7
SI	5.5	6.5	5.8	6.5
SK	14.4	9.8	14.9	10.4

Numbers are obtained as the sum of absolute differences between Cambridge Econometrics and *AUGAMA* region shares divided by two.

Figure 2.7: Within country share in employment of European NUTS2 regions according to the *AUGAMA* database and the Cambridge Econometrics database in 2005



Countries included: AT, BE, BG, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LT, LV, NL, PL, PT, RO, SE, SI, and SK.

correlation coefficient of 0.96 for the sample used in the first column in Table 2.6 .

In the other rows of Table 2.6 we show the correlations over NUTS3 regions by country. For most countries these correlations are high and statistically significant, also for the countries with a larger number of regions.[25] Only in the cases of Bulgaria, Lithuania, and the Slovak Republic, the regional distribution of *AUGAMA* is less well aligned with the one from the Cambridge Econometrics database. For Bulgaria and the Slovak Republic this seems especially driven by the construction and services sectors as the correlation becomes positive and statistically significant if we only consider broad manufacturing. Though positive and significant, the correlation remains smaller than that for most other countries. Overall, we conclude that *AUGAMA* captures the regional distribution of economic activity fairly well. Bearing in mind some smaller caveats highlighted by the above numbers, the use of *AUGAMA* for regional firm-level analysis seems warranted.

[25] The conclusions from Table 2.6 are unchanged when we recalculate correlations after excluding -by country- the most concentrated region that often looks like an 'outlier' compared to the other regions (typically the most concentrated region is also the home of the capital).

Table 2.6: Pairwise correlations of within-country shares of NUTS3 regions in economic activity according to *AUGAMA* and Cambridge Econometrics

	Bus. Econ.		Broad manuf.		Bus. Econ.		Broad manuf.		Bus. Econ.		Broad manuf.	
	L	#reg	L	#reg	L	#reg	L	#reg	VA	#reg	VA	#reg
Europe	0.803*	1239	0.753*	1255	0.872*	975	0.832*	1049	0.723*	975	0.709*	1049
AT	0.985*	35	0.974*	35	0.964*	24	0.941*	34	0.986*	35	0.857*	35
BE	0.965*	44	0.895*	44	0.942*	39	0.860*	44	0.972*	44	0.874*	44
BG	0.154	28	0.496*	28	0.200	17	0.448*	28	0.346	28	0.533*	28
CZ	0.923*	14	0.737*	14	0.923*	14	0.744*	14	0.970*	14	0.864*	14
DE	0.871*	429	0.730*	429	0.858*	392	0.788*	425	0.758*	429	0.601*	429
DK	0.936*	11	0.935*	11	-		-		-		-	
EE	0.977*	5	0.945*	5	0.965*	5	0.884*	5	0.993*	5	0.984*	5
ES	0.973*	52	0.991*	52	0.979*	52	0.990*	52	0.968*	52	0.984*	52
FI	0.983*	6	0.985*	6	0.987*	6	0.983*	6	0.979*	6	0.972*	6
FR	0.907*	96	0.623*	96	0.903*	96	0.628*	96	0.913*	96	0.771*	96
GB	0.664*	120	0.402*	122	-		-		-		-	
GR	0.991*	46	0.993*	51	-		-		-		-	
HU	0.979*	20	0.902*	20	0.966*	20	0.871*	20	0.989*	20	0.908*	20
IE	0.993*	8	0.926*	8	-		-		-		-	
IT	0.954*	107	0.947*	107	0.955*	107	0.947*	107	0.956*	107	0.943*	107
LT	0.589	8	0.596	8	-		-		-		-	
LV	0.995*	6	0.991*	6	0.948	3	0.662	6	0.827*	6	-0.167	6
NL	0.983*	40	0.841*	40	0.886*	38	0.747*	39	0.858*	40	0.706*	40
PL	0.908*	66	0.790*	66	0.894*	66	0.751*	66	0.966*	66	0.776*	66
PT	0.950*	16	0.830*	25	0.967*	14	0.837*	25	0.982*	28	0.894*	28
RO	0.917*	41	0.868*	41	0.924*	41	0.871*	41	0.939*	41	0.863*	41
SE	0.968*	21	0.983*	21	0.993*	21	0.961*	21	0.989*	21	0.946*	21
SI	0.997*	11	0.996*	11	0.996*	11	0.995*	11	0.994*	11	0.986*	11
SK	-0.568*	8	0.757*	8	-0.543	8	0.752*	8	-0.156	8	-0.022	8

L stands for employment, VA for value added, #reg refers to the number of regions. The 'Business Economy' contains all NACE 2-digit sectors from 10 to 74, with the exception of financial services, 'broad manufacturing' contains all all NACE 2-digit sectors from 10 to 42. 'All unconsolidated' refers to the sample of firms that report the number of employees in unconsolidated accounts. 'TFP sample' refers to all firm-year observations for which we are able to obtain a *WLP-TFP* measure.

2.5 *EUMULNET* - A European Multinational Network data set

In addition to our 'regular' panel of European firms, *AUGAMA*, we have also created a separate data set on European multinational networks based on Amadeus (*EUMULNET* henceforth). *EUMULNET* contains parent-affiliate combinations for which we have information both on the side of the parent and its affiliate from Amadeus. For every firm, Amadeus contains information on whether or not the firm has any affiliates. For firms with affiliates, Amadeus provides a list of affiliates with a limited amount of further information. The basis for *EUMULNET* is that affiliates available as separate entries in Amadeus are identified by their unique ID number. For these affiliates we are able to retrieve the full information set from their own entry in Amadeus rather than being limited to the information provided through the parent's entry in the database. *EUMULNET* is then the dataset of those parent-affiliate combinations for which both firms are listed as separate entries in Amadeus with full information. For affiliates not listed in Amadeus, we do have information on their existence and country of operation from the parent's entry in the database. There are two more potentially useful variables in the parent's entry in the database: operating revenue turnover of and employment at the affiliate.[26] However, this information from the parent's entry is not always filled out. Therefore this information is not very useful for further analysis compared to the full entry information. We do retain a variable indicating whether the parent has extra-Europe affiliates or not and focus on parent-affiliate combinations where both firms are listed in Amadeus with full information. This also implies that the resulting data set is limited to the the European network of the parent should it also own non-European affiliates.

To create *EUMULNET* we use the following procedure. First, we extract parent-affiliate ID number combinations (plus the actual share owned by the parent in the affiliate) from every issue of the database.[27] This creates a time series of parent-affiliate

[26]These variables became available only in later versions of the database. Furthermore, with respect to the timing of the information it is also unclear what the calendar year is, since the variable refers to the *latest available* year.

[27]We limit ourselves to European ID numbers. For some affiliates there is a non-European ID number that refers to other Bureau Van Dijk products. This however applies to a very small number of firms. We also do not consider affiliates in Russia and Ukraine at this point.

links. We then restrict our attention to those combinations where the parent owns at least 50% of the affiliate at some point in time. We then replace '*direct*' parents that are found to be affiliates themselves with the '*ultimate*' parent as detected in Amadeus. In the resulting '*ultimate*' parent-affiliate-year data set we then fill out the *AUGAMA* information both on the parent and affiliate side. For earlier/later years when the link does not exist, we do fill out information for parent and/or affiliate from *AUGAMA* when available.[28] Our final data set forms a traditional panel data set in the affiliates-year dimension with full information on the parent side attached to each affiliate-year entry (as such duplicating parent-year information when the parent has multiple affiliates).

Table 2.7 lists the annual number of links of more than 50% between a parent and its affiliate that both have an Bureau Van Dijk ID number. From column (1) one can infer that the number of links we retrieve considerably increases over time, which is influenced by increased coverage over time. For about 4.1% of these links we are able to obtain a *TFP* measure (*cf. infra*) for both parent and affiliate. When we consider the evolution over time of this subset of links in column (2), we still observe an increase in links, but from 2002 onwards, and even more so from 2004 onwards, the number of links is fairly stable. From 2002 onwards between 10 and 17% of these links is between a parent and affiliate in a different European countries (see column (3)).

Table 2.8 focuses on the cross-country distribution for the year 2007. The first column lists the number of parents with a given nationality in the data set. This number is affected by cross-country differences in coverage, but only to a limited extent since we do not require any financial information to be provided by a firm to obtain these numbers. Most parents are found in the UK and the Netherlands, followed by Germany and France. The second column shows the number of affiliates owned by these parents (irrespective of the host country). Across countries, parents on average own between 1.4 and 2.9 affiliates. When we restrict the data to those parent-affiliate combinations for which *WLP-TFP* is available on both sides of the link we retain 34,847 observations in 2007, about 15% of these affiliates is located abroad.[29] For Denmark, the UK, Ireland

[28] Occasionally a link is not reported in the year t issue of the database, while it is in the $t-1$ and $t+1$ issues, we then assume the link to exist in t as well.

[29] These numbers should be interpreted with care as they are partly driven by differences in both pure coverage and data quality (i.e. reporting variables necessary to obtain *WLP-TFP*) across countries. For example, for the Netherlands and Germany we only retain 0.4% and 3.5% of reported links because of

Table 2.7: Total number of parent-affiliate links over time

		parent-affiliate links with...	
	All links	...double TFP	...of which abroad
	(1)	(2)	(3)
1997	24,385	1,221	150
1998	49,795	3,385	455
1999	93,637	7,664	1,021
2000	144,031	13,314	1,444
2001	272,807	19,306	2,064
2002	469,312	23,674	2,983
2003	542,621	24,806	3,378
2004	726,771	29,280	3,899
2005	808,268	32,038	4,168
2006	835,149	34,895	4,719
2007	898,022	34,847	5,127
2008	1,016,984	34,099	5,537
2009	1,139,099	38,969	6,485
2010	1,125,608	38,275	6,504
2011	1,197,820	46,191	6,604

Only parent-affiliate links where the parent owns at least 50% of the affiliate at some point in time are considered. Column (1) shows all links that fulfil these requirements. Column (2) presents the number of these links for which *TFP* is available and column (3) shows links for which *TFP* is available with an affiliate abroad.

and Latvia we are not able to compute a measure of total factor productivity because a necessary variable is missing for all firms.[30] For Greek firms we were able to obtain total factor productivity, but only for a limited number of firms for a limited number of years (late 1990s, early 2000s; not 2007). The last three columns of Table 2.8 focus on the number of affiliates located in the country indicated by the row heading. The correlation with the number of affiliates owned by parent firms from the country is fairly large (abstracting from the requirement of *WLP-TFP* availability), indicating that a lot of these affiliates are typically domestically-owned. The overall share of foreign-owned affiliates is 15% like the share of affiliates owned abroad before.

Finally, Table 2.9 considers the distribution of affiliates per parent for the sample without *WLP-TFP* restrictions (column (1) in Table 2.8) for the year 2007. The general conclusion from the Table is that a small number of parents owns a disproportionally large share of affiliates. Columns (1) and (2) reveal that 61% of parents owns a single affiliate, while another 19% owns two affiliates. In total 94% of parents owns five or less affiliates. Columns (3) and (4) reveal that parents owning affiliates in a foreign country are exceptional: 91.4% of parents does not engage in cross-border investment. 5.3% of parents owns a single foreign affiliate, 3.3% owns two or more foreign affiliates. In columns (5) to (8) we consider the number of affiliates rather than the number of parents. The 65% of parents with a single affiliate account for 26% of the total number of affiliates. 35% of affiliates is owned by parents that own more than five affiliates. The distribution of foreign affiliates looks fairly similar. Foreign affiliates typically belong to multi-affiliate parents. More than 70% of foreign affiliates are owned by parents that have at least two affiliates, 38.4% of foreign affiliates are owned by parents that have six or more affiliates.

poor reporting of financial information. This also accounts for the fact that a large share of the affiliates is located abroad, i.e. in a country with better reporting of financial information.

[30] UK firms do report value added in Amadeus but not material costs.

Table 2.8: Cross-country breakdown of parents and affiliates for the year 2007

	# parents	# affiliates owned	# parent–affiliate links		# affiliates in country	# parent–affiliate in country	
			with double TFP…	…of which abroad		with double TFP…	…of which foreign-owned
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AT	5,783	14,063	319	187	14,261	272	140
BE	16,429	30,271	1,344	362	27,373	1,378	396
BG	3,931	12,803	866	1	13,011	903	38
CZ	463	648	164	27	1,363	318	181
DE	38,438	110,832	3,863	1,351	111,305	2,857	345
DK	25,168	44,179	-	-	41,185	-	-
EE	715	1,134	110	3	1,635	241	134
ES	20,945	47,127	7,659	465	49,564	7,998	804
FI	3,584	8,764	1,985	302	8,552	1,756	73
FR	35,244	87,867	5,776	739	86,972	5,915	878
GB	82,929	227,790	-	-	232,883	-	-
GR	1,113	2,195	-	-	2,408	-	-
HR	442	818	365	10	1,073	427	72
HU	506	789	206	36	1,655	354	184
IE	4,834	11,376	-	-	11,791	-	-
IT	8,848	24,335	5,859	840	22,406	5,498	479
LT	247	364	-	-	643	-	-
LV	146	225	3	1	563	7	5
NL	86,083	171,895	703	510	165,703	256	63
NO	17,333	36,105	1,713	111	37,347	1,740	138
PL	1,803	3,586	228	11	6,471	613	396
PT	2,862	6,161	953	53	6,962	1,152	252
RO	2,733	4,043	1,238	3	6,470	1,620	385
SE	25,745	50,453	1,446	84	46,162	1,478	116
SI	73	134	39	27	77	23	11
SK	33	65	8	4	187	41	37

Column (1) show the number of parents in each country, with columns (2)–(4) providing information about their affiliates and the parent-affiliate links (when *TFP* is available, affiliate located abroad or not). Columns (5)–(7) hold information on the number of affiliates in each country and the parent-affiliate links (when *TFP* is available, domestically-owned or foreign-owned).

Table 2.9: Distribution of the number of affiliates per parent for the year 2007 (without *TFP* restrictions).

	# parents with				total # affiliates owned			
	X affiliates		X foreign affiliates		all		foreign	
X	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	234,360	60.6%	20,573	5.32%	234,360	26.1%	20,573	26.2%
2	72,427	18.7%	5,252	1.36%	144,854	16.1%	10,504	13.4%
3	30,261	7.8%	2,456	0.64%	90,783	10.1%	7,368	9.4%
4	15,712	4.1%	1,385	0.36%	62,848	7.0%	5,540	7.0%
5	9,164	2.4%	891	0.23%	45,820	5.1%	4,455	5.7%
6	5,778	1.5%	591	0.15%	34,668	3.9%	3,546	4.5%
7	3,859	1.0%	409	0.11%	27,013	3.0%	2,863	3.6%
8	2,726	0.7%	319	0.08%	21,808	2.4%	2,552	3.2%
9	1,984	0.5%	239	0.06%	17,856	2.0%	2,151	2.7%
10	1,507	0.4%	157	0.04%	15,070	1.7%	1,570	2.0%
>10	8,652	2.2%	906	0.23%	202,942	22.6%	17,532	22.3%
0	-		353,252	91.41%	-		-	
Total	386,430		386,430		898,022		78,654	

Columns (1)–(4) show the number of parents who own a certain number of affiliates (domestic or abroad), columns (5)-(8) show the number of affiliates owned (by domestic or foreign parents, also considering the number of (other) affiliates this parent owns).

2.6 Total Factor Productivity

2.6.1 Estimation framework

This Section is devoted to the estimation of total factor productivity (*TFP*). As input choices of firms are likely to be based on their productivity, the estimation of total factor productivity will be biased if the endogeneity of inputs is not addressed (Griliches and Mairesse, 1995). A number of alternative estimation procedures have been suggested in order to tackle this issue. The most popular alternatives are the semi-parametric approaches developed by Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP), in which a proxy is introduced to handle the endogeneity bias. Olley and Pakes (1996) use investment as a proxy. Levinsohn and Petrin (2003) argue that investment is lumpy and does not react smoothly to productivity shocks and propose to use material inputs as a proxy instead. In a more recent contribution, Ackerberg et al. (2008) (ACF) present an alternative semi-parametric procedure that deals with potential collinearity issues in Olley and Pakes (1996) and Levinsohn and Petrin (2003). Wooldridge (2009) shows a method to implement OP/LP in a GMM framework with several advantages over ACF: i) estimators are more efficient; ii) the first stage of the algorithm contains identifying information for the parameters on the variable inputs, and iii) fully robust standard errors are easy to obtain. In short, Wooldridge (2009) derives two equations with the same dependent variable (output) and fixed and variable inputs as explanatory variables. The difference between both equations is the approximation of unobserved productivity which provides a different set of instruments for identification of the production function parameters. We use the implementation of Petrin and Levinsohn (2012) of this methodology (referred to as *WLP-TFP* henceforth).

2.6.2 Estimation and coefficients

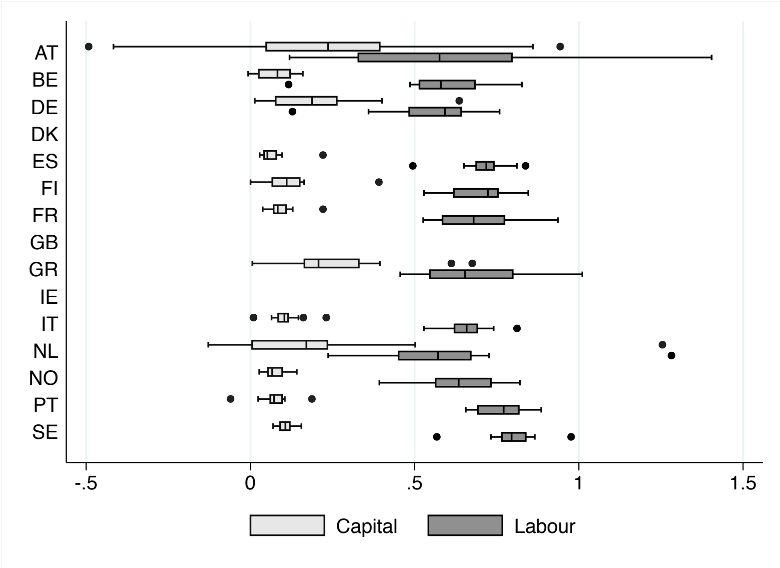
The production function to be estimated is given in its logarithmic form in (2.1) with ω_{it} the unobserved productivity shock known to the firm but not to the researcher and va double deflated value added (*cf. supra*). The sum of the constant term, β_0, and ω_{it} captures Hicks-neutral TFP. ϵ_{it} is a standard i.i.d. error term incorporating unanticipated shocks and measurement error. As indicated above, we use the GMM-approach advocated by Wooldridge (2009) as implemented by Petrin and Levinsohn (2012). The trade-off we face is between allowing β_l and β_k to vary maximally across countries and industries and retaining enough data points to estimate β_l and β_k.

$$\ln va_{it} = \beta_0 + \beta_l \ln l_{it} + \beta_k \ln k_{it} + \omega_{it} + \epsilon_{it} \tag{2.1}$$

We first estimate equation (2.1) by country-industry pair using all available years. Industries are defined as 21 '*broad*' NACE aggregates capturing one or more NACE 2-digit categories (listed in Table 2.A.1). Figures 2.8, 2.9, and 2.10 show box plots[31] of the coefficients by country and by industry respectively. As one can infer from the box plots in Figures 2.8, 2.9, and 2.10, for multiple countries we obtain capital and labour coefficients that fall outside the unit interval. Moreover, for many country-industries,

[31] In a box plot (see Tukey, 1977), the vertical line within the box indicates the median, while the edges of the box represent the 25th and 75th percentiles. The whiskers of a box indicate the upper and lower adjacent values. These are calculated as follows. Let $x_{[25]}$ and $x_{[75]}$ be the 25th and 75th percentiles of for an ordered variable x. Define U as $x_{[75]} + 1.5(x_{[75]} - x_{[25]})$, the upper adjacent value is then x_i such that $x_i \leq U$ and $x_{i+1} > U$. Define L as $x_{[25]} - 1.5(x_{[75]} - x_{[25]})$, the lower adjacent value is then x_i such that $x_i \geq L$ and $x_{i+1} < L$. Values falling outside of this range are indicated by dots.

Figure 2.8: Boxplot of capital and labour elasticities for the old EU15+ countries.



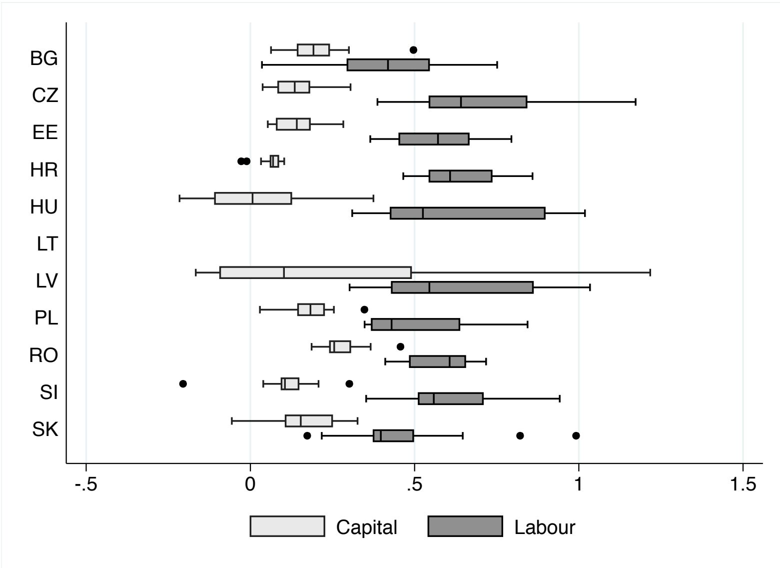
WLP-TFP estimations by country-industry pair; countries indicated on vertical axis, (values of) capital and labour elasticities on horizontal axis.

the capital coefficient is not significant at conventional levels, even when it falls within the unit interval (this is also the case in CompNet (CompNet, 2014)). Gal (2013) deals with this issue by not calculating *TFP* for firms in industries where either the capital or labour coefficient falls outside the unit interval.

We proceed by estimating production functions by industry, but aggregating over countries. When estimating equation (2.1), we restrict β_l and β_k to be the same across countries but allow β_0 to be country-specific (capturing for example country-specific technology levels or management skills). We realise that this is a strong assumption[32], but we prefer to do so because this results in sensible estimates for capital and labour coefficients as shown in Figure 2.11. This allows us to obtain *TFP* for the largest possible set of firms. Furthermore, specifically for multinational (foreign) firms (12% of observations in the dataset) a European production function might be as relevant as the 'local' production function. Our analysis in the next Section is therefore based on *TFP*-values obtained using the estimation results visualised in Figure 2.11. In Figure 2.12 we present a box plot of log *WLP-TFP* by country. The Figure is based on the sample of manufacturing firms with more than 20 employees. The period considered is 2003-2010. Countries are ranked on the basis of the 75th percentile of the *TFP*-distribution. The

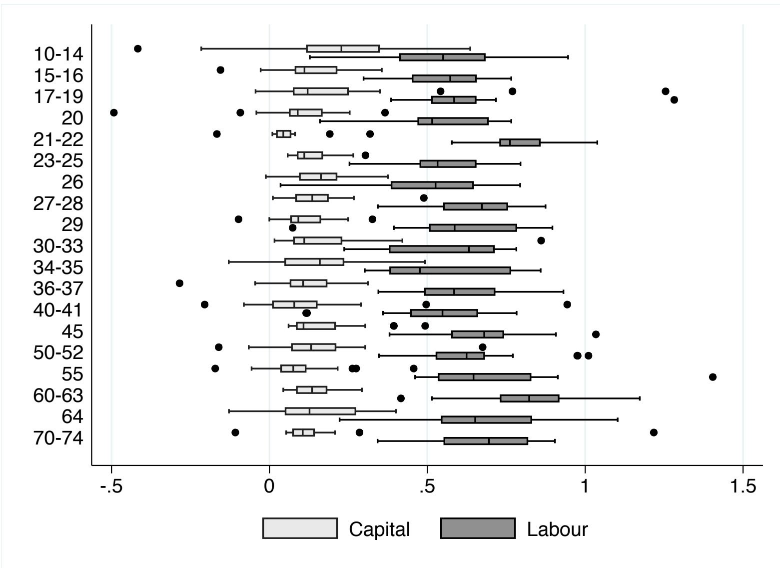
[32] Differences in labour market institutions do exist for example.

Figure 2.9: Boxplot of capital and labour elasticities for the CEEC10+ countries.



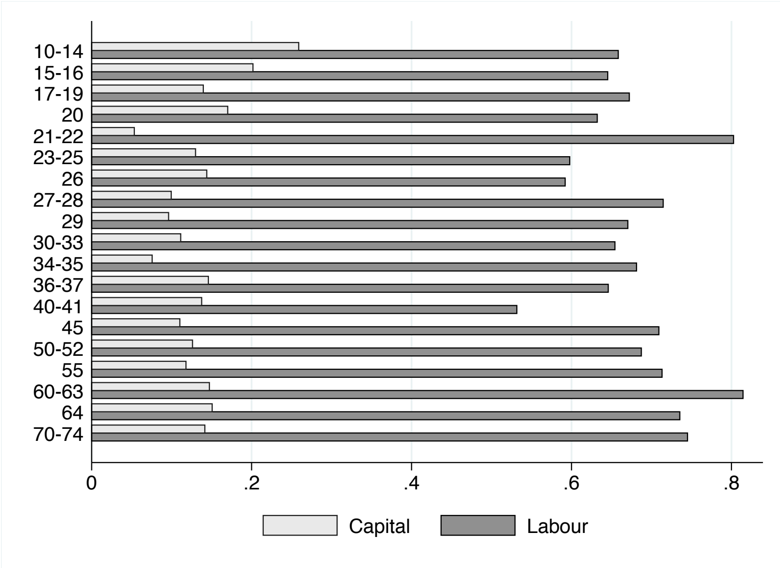
WLP-TFP estimations by country-industry pair; countries indicated on vertical axis, (values of) capital and labour elasticities on horizontal axis.

Figure 2.10: Boxplot of capital and labour elasticities for 21 broad NACE categories.



WLP-TFP estimations by country-industry pair; industries indicated on vertical axis, (values of) capital and labour elasticities on horizontal axis.

Figure 2.11: Capital and labour elasticities for Europe-wide production functions by 21 broad NACE categories.



WLP-TFP estimations by industry (aggregated over countries) including country dummies; industries on vertical axis, (values of) capital and labour elasticities on horizontal axis.

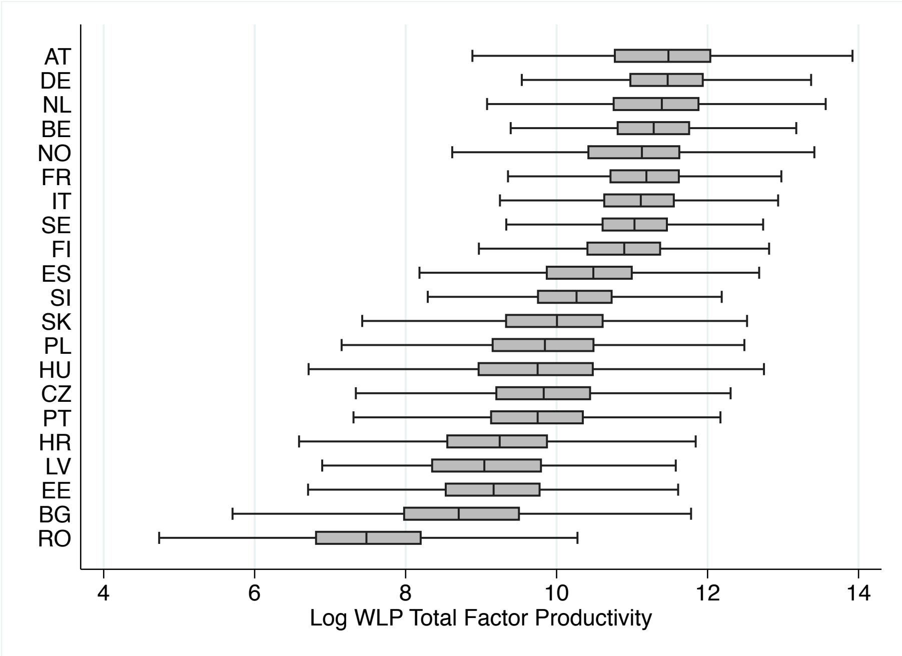
ranking is in line with what one would expect. The old EU-15 countries are generally more productive than the new members from Eastern Europe. Among the old EU-15 Spain and especially Portugal are among the countries with less productive firms.

2.7 Total Factor Productivity and Foreign Ownership

In this Section we analyse *TFP* differences between multinationals, foreign, and domestic firms. In order to take a first look at the productivity levels of the foreign and domestic firms across Europe, we plot the distributions of their *TFP*-levels in Figure 2.13.[33] We find that the distribution for foreign firms is clearly to the right of that for domestic firms. Figures 2.14 and 2.15 display the productivity distributions for domestic and foreign firms in manufacturing and services respectively. In both cases, the productivity distribution for foreign firms is to the right of that for domestic firms. For firms in services industries, the distance between both distributions seems larger (*cf. infra*).

[33]The period considered is 2003-2007, i.e. we exclude both the earlier years where coverage is more unbalanced across countries and the later years to eliminate potential crisis effects. Only firms with on average at least 20 employees are considered. This leaves us with 1,345,454 observations that are used in the Figure. 166,969, i.e. 12,4%, of observations refer to foreign firms.

Figure 2.12: Boxplot log WLP Total Factor Productivity by country.



Sample of manufacturing firms with more than 20 employees; 2003-2010; countries ordered by the 75th percentile of the *TFP*-distribution.

Figure 2.13: *WLP-TFP* distributions for domestic and foreign firms in Europe.

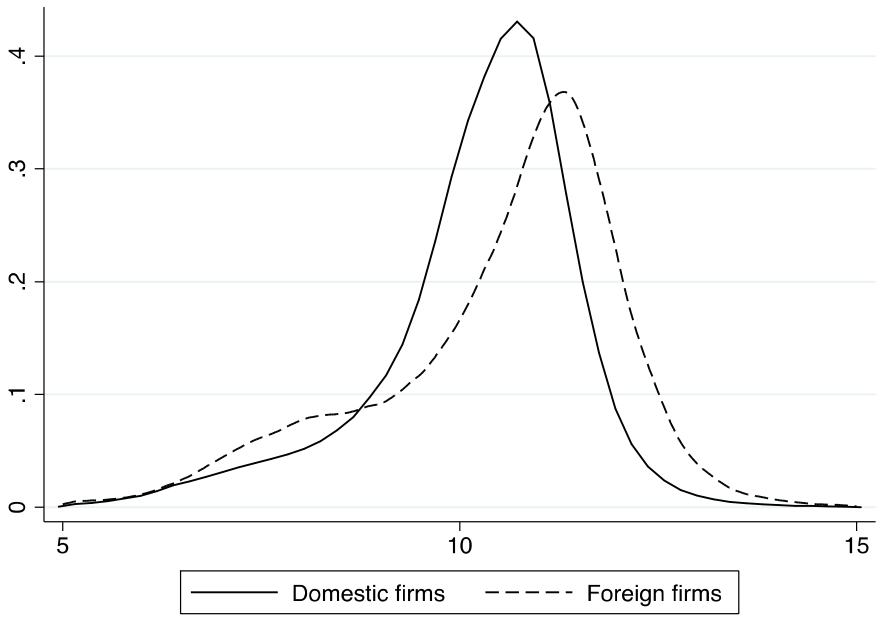


Figure 2.14: *WLP-TFP* distributions for domestic and foreign firms in Europe in the manufacturing industries.

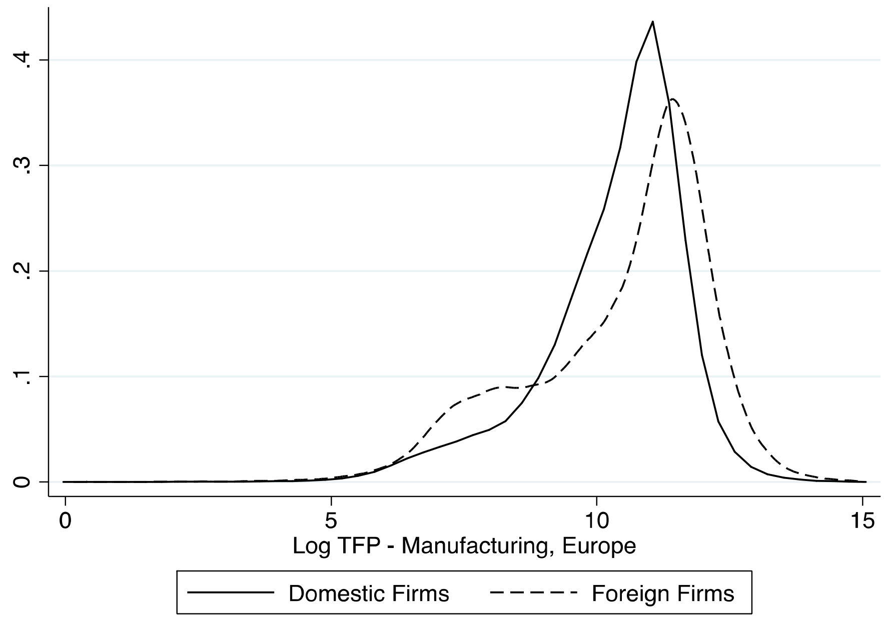
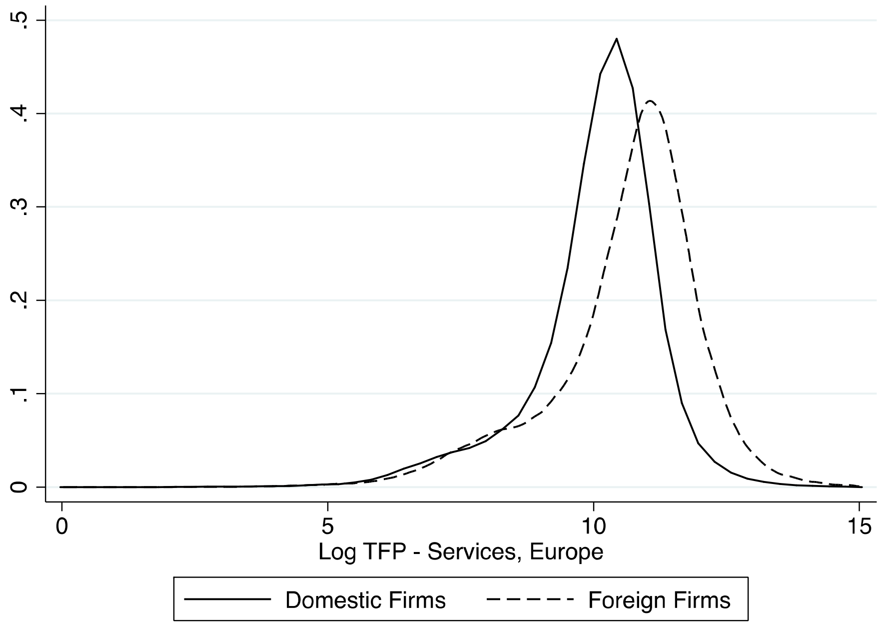


Figure 2.15: *WLP-TFP* distributions for domestic and foreign firms in Europe in the services industries.



When we split Europe in three macro-regions (East, North and South[34]), we can draw a similar conclusion from the first three panels of Figure 2.16: foreign firms appear to be more productive than domestic firms in all three regions. The last two panels in Figure 2.16 respectively illustrate the productivity distributions for the domestic firms in the three macro-regions and for the foreign firms in these three regions. For domestic firms we find a clear ranking with firms in the North outperforming firms in the South and firms in the South outperforming firms in the East. With respect to foreign firms, however, the distribution for South is closer to the distribution for North. Foreign firms in the East do seem to be considerably less productive on average.

In order to get further insight into the magnitude of foreign firms' premium in terms of TFP (and several other performance indicators), we perform an empirical exercise along the lines of Bernard and Jensen (1999). The analysis is fairly straightforward and consists of retrieving the foreign premium from estimating a regression of the following form:

$$lnX_{ijrt} = \alpha + \beta \text{Foreign}_{ijrt} + \delta L_{ijrt-1} + \gamma_t + \gamma_j + \gamma_r + \epsilon_{ijrt} \quad (2.2)$$

In equation (2.2), we regress the level of the performance indicator (X) on a dummy for foreign ownership (*Foreign*), the lagged size of the firm (L, measured as the natural log of the number of employees), and a set of time t, industry j, and region r dummies.[35]

In Table 2.10 we consider the full sample of foreign and domestic firms for which we are able to obtain *WLP-TFP*, but trim the data for extreme values of *WLP-TFP* by removing values below (above) the first (99th) percentile in each country-industry-size-year cell (to preserve the sample distribution in these dimensions).[36] Table 2.10 contains the estimated values for β in equation (2.2). The first column presents premiums for *WLP-TFP* for different subsamples as indicated by the row headings, the third column presents premiums for value added per worker as a comparison check. Columns (2) and (4) contain the number of observations used in the estimation. Premiums are

[34]East is BG, CZ, EE, HR, HU, (LT,) LV, PL, RO, SI, SK; North is AT, BE, DE, (DK,) FI, FR, (GB,) NL, SE; and South is ES, (GR, IE,) IT, PT. For countries between brackets TFP is not available for the period considered. North contains 425,539 observations; South 516,432; and East 373,783.

[35]Industries are defined as the '*broad*' NACE aggregates. We use NUTS2 region dummies (*cf. supra*).

[36]Size classes are defined as micro firms with less than 10 employees; small firms with between 10 and 50 employees; medium-sized firms with between 50 and 250 employees; and large firms with more than 250 employees.

Figure 2.16: *WLP-TFP* distributions for domestic and foreign firms in three Macro-regions (North, South and East) and *WLP-TFP* distributions for domestic and foreign firms separately in these regions compared.

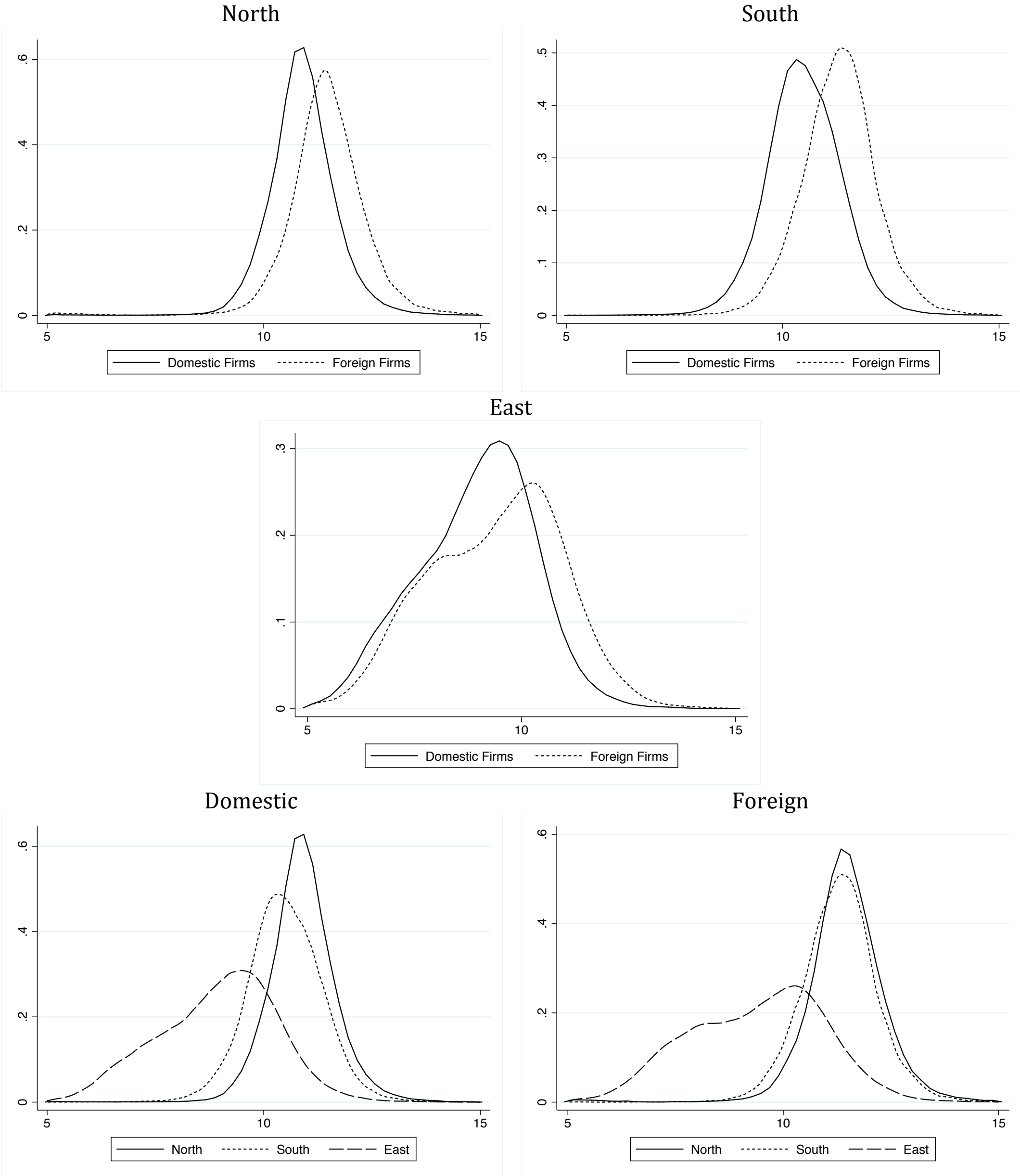


Table 2.10: *TFP* premium of foreign over domestic firms based on the EU-wide sample and different subsamples.

	ln WLP-TFP	# obs	ln VA pw	# obs
	(1)	(2)	(3)	(4)
trimmed	0.480	13,023,107	0.503	13,026,194
non-trimmed	0.555	13,238,694	0.573	13,238,694
manufacturing	0.324	3,020,906	0.345	3,021,943
services	0.557	7,919,363	0.584	7,920,568
before 2003	0.490	3,507,385	0.518	3,507,991
2003–2007	0.483	4,714,522	0.511	4,714,983
after 2007	0.469	4,801,200	0.484	4,803,220
majority foreign-owned firms	0.510	13,064,783	0.538	13,069,327
minority foreign-owned firms	0.335	13,064,783	0.365	13,069,327
micro firms ($L \leq 10$)	0.532	8,505,507	0.566	8,488,588
small firms ($10 < L \leq 50$)	0.441	3,462,642	0.458	3,467,654
medium firms ($50 < L \leq 250$)	0.335	871,972	0.367	879,794
large firms ($L > 250$)	0.329	182,986	0.368	190,158

Premiums for *WLP-TFP* and value added per worker (VA pw). Subsamples are obtained by considering manufacturing and services industries separately, by splitting up the sample period in three shorter periods, by distinguishing between majority and minority foreign-owned firms and by considering four size classes of firms. The foreign premium is statistically significant at the 1% level in all cases in columns (1) and (3). In the trimmed sample, values of *WLP-TFP* below (above) the 1st (99th) percentile in each country-industry-size-year cell are removed.

always significant at the 1% level. Controlling for size, time, industry, and region, we find that foreign firms' level of *WLP-TFP* is on average 48% higher in Europe. This number is confirmed for value added per worker in column (3) where we find a 50% premium. For the non-trimmed sample these premiums are about 7 percentage points higher. When we consider manufacturing and services industries separately we find, in line with Figures 2.14 and 2.15 above, that the premium is considerably larger for services industries. The premium seems fairly stable over time with potentially a slight tendency to decrease, but given changes in sample constellation (*cf.* Tables 2.A.5 and 2.A.6 in the Appendix), one should not read too much in this decrease. As indicated above the criterion to classify a firm as foreign is a single foreign owner controlling at least 10% of shares. When we split foreign firms in a group which is majority foreign-owned (more than 50%) and a group which is minority foreign-owned (more than 10% of the shares, but less than 50%)[37], we find that both groups outperform domestic firms, but that majority foreign-owned firms also outperform minority foreign-owned firms. Majority foreign-owned firms are 51% more productive than domestic firms, whereas minority foreign-owned firms are 33% more productive. Finally, we consider four size categories inspired by the EU's definition of micro (employing less than 10 employees), small (between 10 and 50 employees), medium (between 50 and 250 employees), and large (more than 250 employees) firms. The productivity premium decreases by size class. It is well over 50% for micro firms, about 45% for small firms, and about 33% for medium and large firms.

Tables 2.11 and 2.12 report foreign *WLP-TFP* premiums by country and by industry respectively.[38] For each country in Table 2.11 we consider four different samples of firms: *i*) all firms; *ii*) firms with more than 20 employees (which improves the representativeness of our data); *iii*) firms with more than 20 employees for the period 2003-2007 (resulting in a stable number of firms, and excluding crisis effects); and *iv*) firms in manufacturing with more than 20 employees for the period 2003–2007. We examine the fourth sample of firms because the *WLP-TFP* estimation algorithm is probably more tailored towards manufacturing firms.[39] In Table 2.12, foreign *WLP-TFP*

[37]Of the 313,677 foreign firms (after trimming) 51,523 firms are minority foreign-owned, while 262,154 firms are majority foreign-owned.

[38]Yasar et al. (2007) and Castellani and Giovannetti (2010) examine TFP and exporter premia for Turkey and Italy. These studies find that foreign firms are more productive than domestic counterparts.

[39]Services do account for large parts of value added in all countries.

Table 2.11: *TFP* premium of foreign over domestic firms. Cross-country differences for different subsamples.

	All firms	≥ 20 employees				All firms	≥ 20 employees		
	1995–2011	1995–2011	2003–2007	2003–2007 manuf.		1995–2011	1995–2011	2003–2007	2003–2007 manuf.
country	(1)	(2)	(3)	(4)	country	(1)	(2)	(3)	(4)
AT	0.140***	0.140***	0.130**	0.131	**BG**	0.439***	0.502***	0.555***	0.419***
	11,860	10,837	5,048	1,882		173,065	78,099	19,612	6,304
BE	0.259***	0.309***	0.313***	0.266***	**CZ**	0.541***	0.522***	0.536***	0.411***
	127,543	84,281	27,157	10,430		280,183	103,678	49,411	18,995
DE	0.334***	0.283***	0.299***	0.168***	**EE**	0.585***	0.435***	0.395***	0.258***
	111,977	85,763	34,642	12,652		170,610	26,134	9,880	3,392
ES	0.602***	0.443***	0.447***	0.260***	**HR**	0.489***	0.376***	0.395***	0.247***
	3,291,491	516,590	196,458	64,397		322,076	38,612	18,675	5,903
FI	0.461***	0.316***	0.318***	0.214***	**HU**	0.566***	0.517***	0.548***	0.471***
	314,284	47,284	17,990	6,711		56,296	31,345	5,212	2,465
FR	0.382***	0.242***	0.261***	0.093***	**LV**	0.727***	0.313***	0.328***	0.02
	2,668,045	486,644	163,455	55,449		1,743	444	196	103
IT	0.396***	0.354***	0.364***	0.234***	**PL**	0.659***	0.651***	0.674***	0.469***
	1,611,475	461,499	154,451	85,339		137,341	95,899	39,437	14,471
NL	0.323***	0.191***	0.190***	-0.055	**RO**	0.479***	0.358***	0.396***	0.279***
	12,343	7,565	3,367	1,200		2,018,655	255,552	89,272	37,437
PT	0.624***	0.415***	0.388***	0.322***	**SI**	0.381***	0.304***	0.303***	0.273***
	650,243	68,845	14,607	6,357		99,798	20,373	9,566	4,849
SE	0.334***	0.202***	0.168***	0.114***	**SK**	0.632***	0.594***	0.627***	0.518***
	515,775	48,124	19,665	9,756		58,930	21,751	9,625	4,146
NO	0.422***	0.224***	0.268***	0.124***					
	380,418	57,258	21,204	6,225					

Numbers in rows below estimated coefficients refer to the number of firm-year observations used in estimation. ***/**/* denotes significance at 1/5/10 percent. Columns (1) and (2) present results for the full sample period (1995-2011), columns (3) and (4) for the period 2003-2007. Column (1) shows results for all firms, in columns (2)-(4) only firms with at least 20 employees are considered. In column (4) only manufacturing firms with at least 20 employees are considered.

Table 2.12: TFP premium of foreign over domestic firms. Cross-industry differences for different subsamples.

	All firms	≥ 20 employees	
	1995–2001	1995–2011	2003–2007
	(1)	(2)	(3)
industry		**Manufacturing**	
15t16	0.353***	0.332***	0.341***
	430,871	129,738	44,564
17t19	0.436***	0.331***	0.347***
	322,683	130,767	43,618
20	0.158***	0.097***	0.127***
	187,631	43,839	15,907
21t22	0.336***	0.271***	0.279***
	295,332	77,745	27,193
23t25	0.348***	0.276***	0.273***
	251,737	112,947	39,221
26	0.324***	0.290***	0.325***
	161,834	58,257	20,345
27t28	0.262***	0.207***	0.213***
	579,383	185,578	66,012
29	0.277***	0.209***	0.192***
	274,235	109,596	38,412
30t33	0.277***	0.189***	0.204***
	215,967	79,367	28,219
34t35	0.331***	0.278***	0.306***
	86,758	39,677	13,934
36t37	0.278***	0.198***	0.206***
	232,107	64,456	22,807

	All firms	≥ 20 employees	
	1995–2001	1995–2011	2003–2007
	(1)	(2)	(3)
industry		**Mining and quarrying**	
10t14	0.267***	0.265***	0.298***
	54,111	17,311	5,707
industry		**Electricity, gas and water supply**	
40t41	0.596***	0.518***	0.541***
	48,194	26,159	9,573
industry		**Construction**	
45	0.457***	0.321***	0.368***
	1,962,901	333,339	121,122
industry		**Services**	
50t52	0.555***	0.475***	0.509***
	4,614,026	597,650	215,078
55	0.174***	0.174***	0.195***
	914,574	110,682	40,097
60t63	0.597***	0.536***	0.501***
	661,077	145,195	51,872
64	0.608***	0.454***	0.464***
	40,382	9,769	3,714
70t74	0.553***	0.530***	0.528***
	1,689,304	278,834	101,535

Numbers in rows below estimated coefficients refer to the number of firm-year observations used in estimation. ***/**/* denotes significance at 1/5/10 percent. Columns (1) and (2) report results based on the full sample period (1995-2011), results in column (3) refer to the period 2003-2007. Column (1) show results for all firms, in columns (2) and (3) only firms that employ at least 20 employees are considered.

Table 2.13: Foreign ownership and *WLP-TFP* growth 2005-2011. Evidence from a matched sample.

	All (1)	All (2)	All (3)	Manuf. (4)	Services (5)	All (6)
Foreign	0.030***	0.012***		0.028***	0.033***	0.034***
	[0.002]	[0.003]		[0.003]	[0.003]	[0.002]
Foreign (50%)		0.030***	0.039***			
		[0.003]	[0.002]			
Foreign*crisis						-0.009**
						[0.004]
Crisis						-0.161***
						[0.004]
log empl.$_{t-1}$	0.022***	0.022***	0.022***	0.020***	0.024***	0.022***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Age	-0.000***	-0.001***	-0.001***	-0.000***	-0.001***	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	375,122	375,122	375,122	128,385	230,715	375,122
R^2	0.063	0.063	0.063	0.046	0.077	0.063

Estimations contain year, industry, and country dummies. Standard errors in brackets; ***/**/* denotes significance at 1/5/10 percent.

premiums are evaluated by industry (for 21 broad industries) for samples *i*) to *iii*). In both Tables, numbers in rows below the coefficients refer to the number of firm-year observations used in the estimation. In Table 2.11 we detect the largest premiums for countries in Eastern Europe (on the right-hand side in the Table). For the old EU15 Members States, the largest premiums are found for Portugal, Spain, and Italy. For all countries *WLP-TFP* premiums decrease as the sample becomes more restrictive. The larger premiums are still typically found in Eastern Europe, though for a country such as Slovenia the premium is close to that of the Western European countries. In Table 2.12, we find that premiums in manufacturing industries range from 12.7% to 34.7% for the most restrictive sample. Except for mining (industries 10-14) and hotels and restaurants (industry 55), the premiums in all other industries are considerably higher than those in manufacturing.

We further analyse differences in productivity dynamics between domestic and foreign firms over the period 2005–2011, i.e. three years before and three years after the crisis. To mitigate potential endogeneity of foreign ownership, we employ a matching technique to create a missing counterfactual for each foreign firm's performance. This resolves to some extent the bias of foreign firms owning the most productive firms

Table 2.14: Productivity of *parents* with and without foreign affiliates. Evidence from *EUMULNET*.

	All aff.	$\leq$ 20 aff	$\leq$ 5 aff	Single	$\leq$ 5 aff	$\leq$ 20 aff, manuf.
	(1)	(2)	(3)	(4)	(5)	(8)
foreign aff.	0.281***	0.264***	0.238***	0.221***	0.124***	0.165***
	[0.006]	[0.006]	[0.006]	[0.009]	[0.012]	[0.008]
Lagged log empl.	0.061***	0.059***	0.065***	0.083***	0.064***	0.145***
	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]	[0.002]
# foreign aff.					0.082***	
					[0.007]	
Observations	213,154	210,158	189,865	119,247	189,865	54,123
R^2	0.46	0.46	0.46	0.50	0.47	0.64

Estimations contain year, industry, and country dummies. Standard errors in brackets; ***/**/* denotes significance at 1/5/10 percent.

in the first place and allows for a fairer comparison between domestic and foreign performance.[40] We therefore match each foreign firm with a domestic counterpart in the year 2005. We use Stata's *psmatch2* procedure and slightly modify it to make sure that firms are matched within the same 'broad' industry (and country). Observations are only retained if they belong to the common support. We focus on the year 2005 since this allows us to track a sample of decent size through pre- and post crisis years.[41] After obtaining the set of BVD ID numbers of the matched foreign and domestic firms for the year 2005, we add in the financial and other information for all available years (2005–2011). We use this sample to evaluate the effect of foreign ownership by comparing the growth performance of matched firms (pre- and during the crisis).

Table 2.13 presents the results of the estimation of the following equation.

$$\Delta \ln TFP^{WLP}_{ijct} = \alpha + \beta \text{Foreign}_{ijct} + \delta \ln L_{ijct-1} + \kappa \text{Age}_{ijct} + \gamma_t + \gamma_j + \gamma_c + \epsilon_{ijct} \quad (2.3)$$

We regress *WLP-TFP* growth of firm i in industry j in country c at time t on a dummy

[40] Ideally we would match on both the level and the growth of TFP of domestic and foreign firms prior to 2005, but this would severely limit the number of observations available.

[41] We retain all firms from *AUGAMA* that are present in the data set at least from 2005 onwards (could be earlier) and at least until 2009 (could be later) and that report sufficient information to obtain *WLP-TFP*. On this subset of *AUGAMA* we run a probit for the year 2005 to explain foreign ownership. The explanatory variables are lagged *WLP-TFP*, the lagged number of employees, and age (see Alfaro and Chen (2012)). The balancing hypothesis is satisfied for all three variables.

Table 2.15: Premium of foreign over domestic firms for other performance indicators.

	1995–2011		2003–2007	
performance indicator	premium (1)	# obs (2)	premium (3)	# obs (4)
ln WLP-TFP	0.480	13,023,107	0.266	358,463
ln Value Added per worker	0.503	13,026,194	0.300	361,217
ln Number of Employees*	0.691	11,426,678	0.401	352,576
ln Value Added	0.598	13,003,206	0.369	353,798
ln Capital per worker	0.231	12,992,328	0.298	360,920
ln Operating Revenue	0.565	13,003,535	0.425	353,858
ln Wage	0.326	12,961,377	0.192	361,195
ln Intangible Fixed Assets	0.262	6,247,780	0.298	265,298
ln Total Assets	0.757	12,993,792	0.475	354,003
ln Cash Flow	0.639	10,445,839	0.517	312,699
ln Profit/Loss before Tax	0.771	9,466,307	0.593	278,959
ln Profit/Loss after Tax	0.794	9,129,196	0.616	264,893
Implicit Tax Rate**	-2.457	12,838,221	-2.666	356,248

* specification does not contain lagged number of employees; ** denoted in percentage points. Samples are trimmed for extreme values of the respective indicators by removing values below (above) the 1st (99th) percentile in each country-industry-size-year cell. Columns (1) and (2) show results obtained from a sample covering all firms in all industries in the full sample period (1995-2011). Columns (3) and (4) show results for manufacturing firms with at least 20 employees in the period 2003-2007. The foreign premium is statistically significant at the 1% level in all cases in columns (1) and (3).

indicating foreign ownership, the lagged size of the firm, the age of the firm, and a set of industry, country, and year dummies. In the first column we find that the growth rate of total factor productivity of foreign-owned firms is about three percentage points higher than that of matched domestic firms. This is mainly driven by majority foreign-owned firms that record growth rates that are about four percentage points higher than that of domestic firms, whereas minority foreign-owned firms that record growth rates that are about 1.2 percentage points higher. The difference between foreign and domestic firms is larger for services industries according to point estimates, but the difference is not statistically significant. The last column shows that the difference in growth rates between foreign and domestic firm narrows during the crisis period but the combination of the foreign dummy and its interaction with the crisis variable (which equals one from 2008 onwards) suggests that foreign firms still outperform domestic counterparts by a little more than two percentage points in terms of *TFP* growth.

In Table 2.14 we compare the *WLP-TFP* level of parents with only domestic affiliates with that of parents with at least one foreign affiliate, i.e. an affiliate located in another European country than that of the parent. From *EUMULNET* we retain all parents and information on whether or not they own a foreign affiliate (*ForAff*). We then run the following regression for parents p in industry j in country c at time t:

$$\ln TFP^{WLP}_{pjct} = \alpha + \beta \text{ForAff}_{pjct} + \delta \ln L_{pjct-1} + \gamma_t + \gamma_j + \gamma_c + \epsilon_{pjct} \tag{2.4}$$

In the first column of Table 2.14 we consider all parents when estimating (2.4). We find parents that have at least one foreign affiliate to be 28% more productive than parents with only domestic affiliates. In columns two to four we restrict the sample by considering only firms with less than 21, 6, or exactly 1 affiliate. The *TFP* premium decreases, but when we focus on parents with only a single affiliate (foreign or domestic), we still find a productivity premium of 22% for parents with a single foreign affiliate. This evidence is in line with Antràs and Helpman (2004) who show that only the most productive domestic firms will set up a foreign affiliate. When we introduce the number of foreign affiliates as additional variable in the specification in column five (the sample is restricted to parents with at most five affiliates), we find that an additional foreign affiliate is associated with an eight percentage points larger *TFP* premium over parents with only domestic affiliates. Column six finally shows that the productivity premium decreases to 16.5% when we only consider parents in manufacturing industries.

Table 2.15, finally, considers other performance indicators than productivity. For ease of comparison the first two lines repeat the results of Table 2.10 for *WLP-TFP* and value added per worker. All performance indicators have been trimmed in a similar way as indicated for *TFP* above. The results reported in column (1) cover a sample of all firms and industries in period 1995–2011; the results in column (3) are obtained from an analysis using only manufacturing firms that employ at least 20 employees in period 2003–2007. We first focus on the results for the full sample of firms. Aside from being about 50% more productive on average, foreign firms create 60% more value added, have 56% more operating revenues, and generate a 64% larger cash flow. They do so by employing about 70% more employees than domestic firms, using 23% more capital per employee, 26% more intangibles (based on the subsample of firms reporting strictly positive intangibles). Foreign firms' total assets are on average 75%

larger. All this results in profits that are slightly less than 80% larger (based on the subsample of firms reporting strictly positive profits). Their implicit tax rate, calculated as profits before tax minus profits after tax divided by profits before tax, is on average 2.5 percentage points smaller than domestic firms' average implicit tax rate. When we restrict the sample to manufacturing firms with at least 20 employees in 2003-2007, we confirm the premiums of foreign over domestic firms for all performance indicators. As can be seen from column (3), the premiums are generally smaller. From Table 2.10 we know that this is more likely to be driven by the exclusion of services firms than by the narrower period considered. Only for capital per worker and intangible fixed assets the premium slightly increases. The difference in implicit tax rates is also slightly larger for this subsample.

2.7.1 Total Factor Productivity of Parents and Affiliates

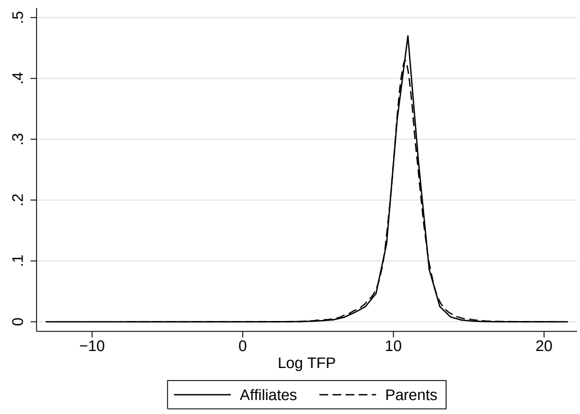
This section looks further at the TFP differences and correlations between the productivity of affiliates and their parents. For this purpose we use majority owned affiliates only for which the Wooldridge-Levinsohn-Petrin TFP is available. This results in close to 400,000 observations. We first plot the densities of these two groups (parents and affiliates), where parents of multiple affiliates are counted just once.

Figure 2.17 presents plots for all affiliates, as well as making a distinction between domestic and foreign ones. Panel (a) shows the results for all parents and affiliates, whilst in panels (b) and (c) we split the sample by domestic and foreign affiliates. While the distributions of the entire dataset look very similar, a K-S tests rejects their equality. A similar picture emerges when we only consider the domestic links, representing about 80% of our data. In case of foreign affiliates, the distributions are visibly different, with parents owning affiliates abroad being more productive than these affiliates. This is no surprising, given that multinationals are generally much more efficient than firms operating in one country only (see Helpman et al. (2004)).

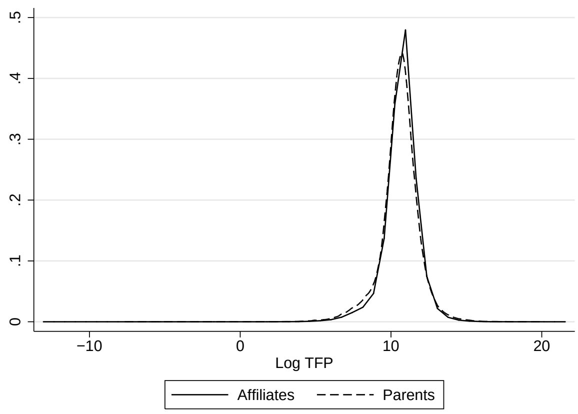
Next we look at how much of the variation in TFP of affiliates can be explained by the productivity of their parents. We run several simple regression with parent's TFP as explanatory variable and sequentially include year, industry and country fixed effects, as shown in (2.5):

Figure 2.17: TFP densities of parents and their affiliates

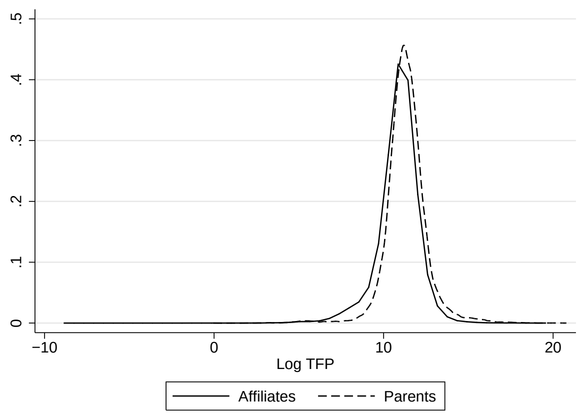
(a) All affiliates



(b) Domestic affiliates



(c) Foreign affiliates



Note: The graph shows WLP TFP distributions of parents and affiliates. The first panel contains the entire sample, while the last two panels splits the data by whether parents/affiliates are domestic or foreign.

$$\ln TFP_{ajct}^{WLP} = \alpha + \beta \ln TFP_{pjct}^{WLP} + \gamma_t + \gamma_j + \gamma_c + \epsilon_{ajct} \tag{2.5}$$

Table 2.16 presents the results for the entire sample, domestic and foreign affiliates. The last four columns shows the coefficients and R-squared for foreign affiliates when including country and industry fixed effects of parents only and of parents and affiliates respectively.

The results suggest that for the entire sample as well as for domestic affiliates, controlling for year, affiliate country and industry fixed effects, there is a positive and significant correlation between the TFP of parents and their affiliates. This correlation drops slightly if we only consider foreign affiliates, confirming the fact that multinationals and their affiliates are less alike than their domestic only counterparts. Moreover, affiliates' countries and industries are important in determining productivity, since we observed marked improvements in the R-squared as more controls are included.

The last two regressions should tell us more about the impact of parent's industry and country on this relationship. As expected, the parent's controls have a much lower explanatory power, TFP being influenced primarily by the affiliate's controls. Nonetheless, we observe a slight improvement in the R-squared when including both sets of fixed effects, suggesting that the country and industry of the parent will also have in impact on how productive the affiliate is.

Finally, we also check whether being a single affiliate in a parent's portfolio impacts the correlation between their productivities. We include a dummy taking value one if the parent owns other affiliates as well. We expect that firms belonging to larger networks to be more productive due to a more efficient allocation of tasks within these groups or learning synergies between affiliates belonging to the same parent (see Egger et al. (2014)). Table 2.17 presents the new results. We indeed see that having a parent with more than one affiliate is positively and significantly correlated with affiliate productivity, confirming our hypothesis.

Table 2.16: Explaining TFP performance of affiliates with that of their parent

	Affiliate controls						Parent controls		Parent and affiliate controls	
	All		Domestic		Foreign		Foreign		Foreign	
Controls	Parent TFP	Obs/R^2	Parent TFP	Obs/R^2	Parent TFP	Obs/R^2	Parent TFP	Obs/R^2	Parent TFP	Obs/R^2
	0.423***	395,600	0.460***	339,768	0.222***	55,832	0.222***	55,832	0.222***	55,832
	[0.088]	0.216	[0.103]	0.248	[0.041]	0.058	[0.041]	0.058	[0.041]	0.058
Year	0.420***	395,600	0.459***	339,768	0.218***	55,832	0.218***	55,832	0.218***	55,832
	[0.088]	0.219	[0.104]	0.250	[0.040]	0.067	[0.040]	0.067	[0.040]	0.067
Industry	0.368***	395,600	0.407***	339,768	0.164***	55,832	0.202***	55,832	0.200***	55,832
	[0.091]	0.319	[0.108]	0.351	[0.028]	0.191	[0.042]	0.094	[0.035]	0.212
Country	0.259***	395,600	0.262***	339,768	0.162***	55,832	0.205***	55,832	0.167***	55,832
	[0.025]	0.403	[0.031]	0.408	[0.013]	0.411	[0.032]	0.089	[0.013]	0.417
Year, Industry,	0.193***	395,600	0.191***	339,768	0.113***	55,832	0.174***	55,832	0.136***	55,832
Country	[0.020]	0.513	[0.023]	0.520	[0.011]	0.501	[0.032]	0.127	[0.014]	0.513

The table shows results of regression of affiliate TFP on the TFP of their parent. Controls are added in form of Year, Industry and Country fixed effects of affiliates and parents respectively.

Table 2.17: Explaining TFP performance of affiliates with that of their parent

	Affiliate controls						Parent controls		Parent and affiliate controls	
	All		Domestic		Foreign		Foreign		Foreign	
Controls	Parent TFP/ Multiple affiliates dummy	Obs/ R^2	Parent TFP/ Multiple affiliates dummy	Obs/ R^2	Parent TFP/ Multiple affiliates dummy	Obs/ R^2	Parent TFP/ Multiple affiliates dummy	Obs/ R^2	Parent TFP/ Multiple affiliates dummy	Obs/ R^2
	0.407***	395,600	0.446***	339,768	0.207***	55,832	0.207***	55,832	0.207***	55,832
	[0.085]	0.222	[0.102]	0.252	[0.031]	0.075	[0.031]	0.075	[0.031]	0.075
	0.221***		0.190***		0.411*		0.411*		0.411*	
	[0.043]		[0.036]		[0.212]		[0.212]		[0.212]	
Year	0.404***	395,600	0.444***	339,768	0.202***	55,832	0.202***	55,832	0.202***	55,832
	[0.085]	0.226	[0.103]	0.255	[0.030]	0.085	[0.030]	0.085	[0.030]	0.085
	0.225***		0.192***		0.420*		0.420*		0.420*	
	[0.043]		[0.035]		[0.210]		[0.210]		[0.210]	
Industry	0.351***	395,600	0.391***	339,768	0.151***	55,832	0.185***	55,832	0.185***	55,832
	[0.089]	0.326	[0.108]	0.357	[0.022]	0.204	[0.033]	0.106	[0.029]	0.221
	0.233***		0.209***		0.361**		0.351**		0.317**	
	[0.035]		[0.030]		[0.143]		[0.166]		[0.129]	
Country	0.246***	395,600	0.250***	339,768	0.156***	55,832	0.195***	55,832	0.163***	55,832
	[0.024]	0.408	[0.030]	0.412	[0.011]	0.415	[0.027]	0.100	[0.011]	0.420
	0.185***		0.166***		0.181***		0.343**		0.159***	
	[0.018]		[0.020]		[0.044]		[0.162]		[0.040]	
Year, Industry, Country	0.180***	395,600	0.178***	339,768	0.108***	55,832	0.163***	55,832	0.132***	55,832
	[0.019]	0.518	[0.022]	0.525	[0.009]	0.504	[0.027]	0.136	[0.013]	0.515
	0.198***		0.183***		0.173***		0.308**		0.134***	
	[0.014]		[0.018]		[0.028]		[0.129]		[0.026]	

The table shows results of regression of affiliate TFP on the TFP of their parent. Controls are added in form of Year, Industry and Country fixed effects of affiliates and parents respectively.

2.8 Conclusions

This chapter documents in detail the build of two datasets on the basis of raw data taken from the Amadeus database issued by Bureau Van Dijk Electronic Publishing. The first dataset, *AUGAMA*, is a large panel of firms in 26 European countries. More than 18 million observations with all information needed to obtain a measure of *TFP* are available for more than 3.6 million firms. The data stretch the period 1995-2012 and for the average firm 6.1 years of data is available. The coverage for the years before 2002 is generally lower (not for all countries though) and not for all firms information for 2012 is already available. We show that *AUGAMA* adequately approximates the structure of the European economy across countries, regions, and industries as portrayed by data from Eurostat (Structural Business Statistics) and Cambridge Econometrics. The second dataset, *EUMULNET*, is a dataset of European multinational networks with 'full' information, i.e. *TFP*, for both parents and their European affiliates. We have more than 600,000 parent-affiliate-year observations, 16% of these are links between a parent and affiliate in different countries. The period is 1997-2012, but as with *AUGAMA* the coverage is lower for earlier years and information for 2012 is not available for all parent-affiliate combinations. The advantage of *AUGAMA* and *EUMULNET* is that both datasets cover cross-country comparable firm-level data for a large number of European countries in a single dataset. This allows for cross-country research at the firm level while maintaining representativeness that is e.g. comparable to the 'distributed micro-data analysis' by CompNet ((CompNet, 2014)).

We use *AUGAMA* to estimate foreign firms' productivity premium over domestic firms. We follow Bernard and Jensen (1999) but estimate productivity premiums for foreign firms rather than for exporters using simple regression analysis. We find that across Europe on average foreign firms are 48% more productive than domestic firms. This is mainly driven by services sectors where foreign firms are on average about 56% more productive, whereas the difference amounts to 32% in manufacturing industries. Majority foreign-owned firms (>50% of the shares are foreign-owned) outperform minority foreign-owned firms, who in turn still outperform domestic firms. We also find the premium to be smaller in larger firm-size categories. Productivity premiums are typically larger in Eastern European countries than in Western European countries. For a number of other firm characteristics (e.g. value added, profitability, intangible assets) we find similar premiums for foreign firms. In a framework similar to Alfaro and Chen (2012) we find that foreign firms grow about 3%-points faster over the period 2005–2011. During the crisis period the growth gap between foreign and domestic firms becomes smaller but remains positive. Finally, in line with Antràs and Helpman (2004) we find parents with foreign affiliates to be 27% more productive than parents with only domestic affiliates. The number decreases to 14.4% when we only consider parents in manufacturing industries. We also find that there is a positive correlation between affiliates' productivities and those of their parents, as productive parents will share know-how with the affiliate. Moreover, companies with parents owning multiple affiliates are also more productive, potentially due to the fact that they learn from their 'siblings' and that larger networks benefit from economies of scale or scope.

Appendix 2.A Additional graphs and tables

Table 2.A.1: List of the NACE 2-digit industries included in the data.

Broad category	NACE 2-digit	Description
C		**Mining and quarrying**
C	10	Mining of coal and lignite; extraction of peat
C	11	Extraction of crude petroleum and natural gas
C	12	Mining of uranium and thorium ores
C	13	Mining of metal ores
C	14	Other mining and quarrying
D		**Manufacturing**
DA	15	Manufacture of food products and beverages
DA	16	Manufacture of tobacco products
DB	17	Manufacture of textiles
DB	18	Manufacture of wearing apparel; dressing and dyeing of fur
DC	19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
DD	20	Manufacture of wood and products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
DE	21	Manufacture of pulp, paper and paper products
DE	22	Publishing, printing and reproduction of recorded media
DF	23	Manufacture of coke, refined petroleum products and nuclear fuel
DG	24	Manufacture of chemicals and chemical products
DH	25	Manufacture of rubber and plastic products
DI	26	Manufacture of other non-metallic mineral products
DJ	27	Manufacture of basic metals
DJ	28	Manufacture of fabricated metal products, exc. machinery/equipment
DK	29	Manufacture of machinery and equipment n.e.c.
DL	30	Manufacture of office machinery and computers
DL	31	Manufacture of electrical machinery and apparatus n.e.c.
DL	32	Manufacture of radio/television/communication equipment/apparatus
DL	33	Manufacture of medical/precision/optical instruments, watches/clocks
DM	34	Manufacture of motor vehicles, trailers and semi-trailers
DM	35	Manufacture of other transport equipment
DN	36	Manufacture of furniture; manufacturing n.e.c.
DN	37	Recycling

Table continued on the next page

Table 2.A.2: List of the NACE 2-digit industries included in the data (Continued).

Broad category	NACE 2-digit	Description
E		**Electricity, gas and water supply**
E	40	Electricity, gas, steam and hot water supply
E	41	Collection, purification and distribution of water
F		**Construction**
F	45	Construction
G		**Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods**
G	50	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel
G	51	Wholesale trade and commission trade, except of motor vehicles and motorcycles
G	52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods.
H		**Hotels and restaurants**
H	55	Hotels and restaurants
I		**Transport, storage and communication**
I	60	Land transport; transport via pipelines
I	61	Water transport
I	62	Air transport
I	63	Supporting and auxiliary transport activities; activities of travel agencies
I	64	Post and telecommunications
K		**Real estate, renting and business activities**
K	70	Real estate activities
K	71	Renting of machinery and equipment without operator and of personal and household goods
K	72	Computer and related activities
K	73	Research and development
K	74	Other business activities

Table 2.A.3: Definition of Nomenclature of territorial units for statistics (NUTS) of the European Union. Minimum and maximum population thresholds indicated.

NUTS level	Description	Min.	Max.
NUTS 1	Major socio-economic regions	3 million	7 million
NUTS 2	Basic regions for application of regional policies	800,000	3 million
NUTS 3	Small regions for specific diagnoses	150,000	800,000

See also "Regions in the European Union; Nomenclature of territorial units for statistics, NUTS 2006/EU-27". NUTS favours administrative divisions. `http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-020/EN/KS-RA-07-020-EN.PDF`.

Table 2.A.4: NUTS regions: Example for Belgium

Level	Code	Description
NUTS 0	BE	Belgique / België
NUTS 1	BE1	Region de Bruxelles-Capitale / Brussels Hoofdstedelijk Gewest
NUTS 2	BE10	Region de Bruxelles-Capitale / Brussels Hoofdstedelijk Gewest
NUTS 3	BE100	Arrondissement Bruxelles-Capitale / Arrondissement Brussel-Hoofdstad
NUTS 1	BE2	Vlaams gewest
NUTS 2	BE21	Provincie Antwerpen
NUTS 3	BE211	Arrondissement Antwerpen
NUTS 3	BE212	Arrondissement Mechelen
NUTS 3	BE213	Arrondissement Turnhout
...		

Table 2.A.5: Number of firms with at least 20 employees and *WLP-TFP* available in *AUGAMA* (period 1996–2011).

	avg. # years	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
AT	4.3	118	243	331	403	331	266	952	1,349	1,793	1,699	2,074	2,006	1,875	1,768	1,982	1,703
BE	10.0	5,322	5,356	5,478	5,450	5,640	5,691	5,740	5,884	6,037	6,132	6,311	6,470	6,547	6,587	6,681	6,630
BG	6.5	3,921	3,260	3,745	6,158	6,240	6,194	3,412	3,693	3,767	4,238	7,248	9,540	11,466	11,793	11,534	6,599
CZ	7.1	153	147	119	425	4,152	5,608	7,710	9,228	11,456	12,522	13,584	14,574	12,148	15,401	14,604	10,455
DE	3.7	1,073	1,304	1,834	2,593	2,862	3,176	5,794	8,445	9,954	13,426	15,682	15,495	15,359	15,541	13,384	12,561
EE	8.7	220	1,620	1,939	2,217	2,405	2,218	2,375	2,398	2,085	2,138	2,162	2,215	2,116	1,789	1,748	1,689
ES	9.7	23,376	25,347	28,042	29,847	33,563	36,217	39,636	41,510	43,317	45,206	47,160	46,816	46,790	48,029	45,200	34,593
FI	7.6	2,275	3,212	3,698	3,779	3,984	3,969	4,203	4,363	4,349	4,385	4,533	4,553	3,961	4,017	3,635	3,826
FR	8.1	31,019	35,325	47,357	52,495	53,110	44,176	43,341	43,606	43,778	43,343	38,054	36,518	31,714	30,353	34,671	38,268
GR	2.2			4,566	2,126	3,734	1,890										
HR	8.4	18	27	2,833	3,583	392	647	3,646	3,964	3,990	4,105	4,341	4,372	4,380	4,637	4,492	4,326
HU	4.2	407	2,028	2,754	2,571	2,034	1,093	4,423	1,066	1,119	2,558	3,046	9,831	5,555	9,905	5,759	4,940
IT	8.1	28,824	31,175	32,775	32,409	34,513	34,414	38,071	36,453	36,872	29,778	39,469	40,557	37,030	37,194	34,074	49,012
LV	4.3					14	33	43	56	68	75	82	90	73	64	134	134
NL	3.5	236	244	246	501	575	661	932	943	1,003	1,079	1,111	1,212	1,266	1,246	984	474
NO	7.5	3,709	1,932	1,999	2,295	2,492	2,553	2,530	2,722	5,877	6,018	6,092	6,461	6,704	6,858	6,748	6,736
PL	5.0	921	2,063	2,251	4,538	5,651	6,652	6,718	8,161	8,611	9,649	13,394	11,948	13,789	22,246	10,821	6,477
PT	5.5	1,472	1,442	1,804	474	669	618	660	798	793	810	14,119	14,407	14,756	14,761	13,934	12,780
RO	8.9	12,682	13,510	14,353	15,736	17,726	17,639	18,480	20,262	21,465	22,832	20,004	23,985	22,835	21,280	21,500	21,472
SE	6.8		2,229	5,370	5,690	6,085	5,861	6,390	5,435	4,134	4,445	4,453	4,133	1,612	1,651	1,630	1,673
SI	7.6		963	458	401	709	825	1,704	1,877	2,041	2,131	2,304	2,304	2,273	2,324	2,782	2,688
SK	5.9	10	49	107	167	364	744	1,152	1,686	1,689	3,007	3,553	3,606	3,021	4,178	3,660	2,293

Table 2.A.6: Number of *foreign* firms with at least 20 employees and *WLP-TFP* available in *AUGAMA* (period 1996–2011).

	avg. # years	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
AT	4.9	36	52	60	87	71	52	206	345	491	461	552	519	484	434	507	429
BE	12.1	993	1,012	1,039	1,026	1,050	1,047	1,100	1,170	1,232	1,270	1,290	1,369	1,366	1,370	1,418	1,380
BG	8.1	249	223	312	557	639	683	487	541	574	638	869	1,041	1,142	1,145	1,089	705
CZ	7.7	32	33	14	97	845	1,226	1,571	1,816	2,225	2,380	2,589	2,738	2,367	2,911	2,839	2,136
DE	4.8	104	114	178	256	255	311	568	826	960	1,654	2,193	2,349	2,430	2,422	2,130	1,403
EE	8.8	65	379	431	501	555	535	570	569	487	503	516	524	500	435	429	416
ES	10.4	1,445	1,649	1,674	1,738	1,922	2,001	2,118	2,177	2,261	2,262	2,352	2,363	2,366	2,462	2,380	1,703
FI	9.1	241	299	331	317	339	372	422	452	456	472	497	480	408	427	396	384
FR	9.0	2,297	2,725	3,359	3,528	3,622	3,120	3,175	3,197	3,173	3,172	2,917	2,864	2,290	2,180	2,588	2,451
GR	2.2			348	135	236	152										
HR	8.9	2	3	226	273	20	41	304	346	378	395	405	425	436	454	450	445
HU	6.5	89	355	436	400	300	187	719	212	206	390	299	613	552	656	659	580
IT	10.4	1,961	2,075	2,188	2,161	2,249	2,229	2,294	2,272	2,393	2,223	2,451	2,504	2,438	2,398	2,335	2,341
LV	5.5					2	10	13	17	25	28	28	29	32	25	33	32
NL	4.4	55	61	69	133	145	161	226	222	217	237	248	268	311	332	247	121
NO	8.1	279	132	141	183	213	211	213	228	522	544	544	534	565	566	593	608
PL	6.1	177	335	437	893	988	1,175	1,285	1,519	1,608	1,918	2,260	2,148	2,527	3,454	2,229	1,285
PT	6.0	81	64	95	40	51	48	55	66	66	58	563	613	730	756	708	615
RO	9.0	2,778	3,031	3,452	3,933	4,251	3,872	4,039	4,397	4,728	5,119	4,605	5,231	5,013	4,781	4,767	4,783
SE	7.3		151	330	354	380	380	430	304	199	215	223	206	164	169	165	175
SI	9.3		121	61	63	100	128	235	255	280	296	321	325	329	333	352	346
SK	6.5	2	10	20	33	80	161	248	339	323	648	748	726	631	765	678	477

3

Distance, Time since Foreign Entry, and Productivity Spillovers from Foreign Direct Investment*

3.1 Introduction

Nowadays, countries actively and fiercely compete to attract foreign direct investment (see Harding and Smarzynska Javorcik (2011)). Policymakers are eager to do so for several reasons. First of all, they expect to benefit in terms of faster economic growth in their country through additional foreign capital and higher employment. Second, foreign firms are expected to bring more advanced technology (see Markusen (1995)) which policymakers believe to 'spill over' to domestic firms, with increased domestic productivity as a result. Spillovers from foreign to domestic firms have been investigated at least since Caves (1974). Initially, it proved difficult to detect clear evidence of aggregate positive FDI spillovers (see Görg and Greenaway (2004); Crespo and Fontoura (2007)). Following Javorcik (2004), the literature now distinguishes

*This chapter is based on a paper forthcoming in Papers in Regional Science. Co-authored with Bruno Merlevede.

between spillovers within the same industry (horizontal spillover effects) as well as those resulting from vertical links along the supply chain (backward and forward spillover effects). The recent literature seems to have established fairly robust evidence of positive backward spillover effects from foreign firms to their domestic suppliers. By means of a meta-analysis Havránek and Irŝová (2011) confirm that the average backward spillover effect is both statistically and economically significant.

Following new theoretical insights that stress the importance of firm level heterogeneity (see Melitz (2003) and Helpman et al. (2004)), the literature has moved away from the idea that spillovers are unconditional and uniform. The focus has instead turned to the identification of characteristics that facilitate positive spillover effects, concerning domestic firms' characteristics such as absorptive capacity (e.g. Merlevede and Schoors (2007)) or foreign firms' characteristics such as ownership structure (e.g. Javorcik and Spatareanu (2008) and Kamal (2014)).

We contribute to this literature by analysing the pattern of technology diffusion as a function of distance between domestic and foreign firms and foreign firms' maturity in the domestic economy. Although both dimensions have been analysed before in firm-level spillover studies, they have not yet been combined. The combination of both is, however, important and can yield new insights. This is illustrated by Comin et al. (2012) who study cross-country technology adaption and find that technology diffuses slower to locations far away, but the effect of distance vanishes over time. Firm-level spillovers, especially backward spillovers are likely to be sensitive to the combination of distance and multinational enterprises' (MNEs) maturity. MNEs have a keen interest in technological upgrading by their suppliers and therefore an incentive to provide them with explicit assistance (see Javorcik (2004)). As successful assistance and upgrading requires human interaction, communication, and monitoring, nearby suppliers will be able to learn and upgrade faster and potentially more (cf. Giroud (2012) and Keller and Yeaple (2013)).

As *AUGAMA* was not yet available, in this chapter we use a pilot dataset of Romanian firms with variation across NUTS 3 regions to identify patterns of productivity spillovers.[1] We find negative horizontal spillover effects for medium maturities of for-

[1]The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU. A NUTS level 3 region is a "small region for specific diagnoses". The minimum and maximum population thresholds for a NUTS 3 region are defined as

eign firms, but larger positive spillover effects for foreign firms that have been present for at least four full years in the domestic economy. On aggregate, the effect of distance on horizontal spillovers is limited. Point estimates do suggest that distance mitigates initial negative effects, while it increases the positive spillovers from foreign firms with longer presence. Point estimates cannot, however, be rejected to be equal over different distances at conventional levels. This suggests that the competition and labour market channels driving horizontal spillovers are largely independent of distance. Backward spillovers, on the other hand, are affected by distance. Domestic firms located in the same region as a foreign client experience a bonus effect quickly following foreign entry. 'Relocating' a firm from its own region to Bucharest-Ilfov, the capital region and top FDI location, is associated with an 11 per cent larger backward spillover. For medium maturities of foreign firms, we find a significant positive backward spillover which is unrelated to distance. However, these spillovers disappear after several years of foreign presence. Taking into account regional heterogeneity, the pattern of backward spillovers is confirmed for above median productivity regions, while for below median productivity regions, we no longer observe a significant within-region backward bonus upon entry of foreign clients. Consistent with the view that it might take time for spillovers to be absorbed over larger distances, domestic firms in below median productivity regions do experience positive backward spillovers from further away, but more mature foreign firms.

The remainder of the chapter is structured as follows. In section 2 we discuss the related literature and derive expectations regarding the interaction of distance and maturity. Section 3 introduces our regional time-since-foreign-entry approach to spillovers, while in section 4 we give a short overview of our data. Section 5 presents the results and we conclude in section 6.

3.2 Literature

The literature on spillover effects distinguishes between two types of spillovers. Horizontal spillovers run from a foreign firm to a host country firm competitor, whereas vertical spillovers originate in customer-supplier relationships between foreign and

150,000 and 800,000. With "distance" we refer to proximity and regional patterns rather than pure distance.

domestic firms, and can be of a backward or forward type. Teece (1977) suggests two main channels for horizontal spillovers: technology imitation and mobility of workers trained by foreign firms (see also Fosfuri et al. (2001) and Marin and Bell (2006)). Foreign entry may also fuel competition in the domestic market. Fiercer competition urges host country firms to either use existing technologies and resources more efficiently or adopt new technologies and organizational practices, which provides another important channel for horizontal effects of foreign presence (see Aitken and Harrison (1999), and Glass and Saggi (2002)). The latter channel does not represent a transfer of knowledge, but is simply a gain or loss from increased competition (see Havránek and Irŝová (2013)). The standard empirical framework (*cf. infra*) captures the net effect of these different channels by the productivity semi-elasticity of the spillover variable. We follow the existing literature in labelling the within-industry impact of foreign presence on domestic firms' productivity horizontal 'spillover effects' (see Crespo and Fontoura (2007) and Havránek and Irŝová (2013)), bearing in mind that the effect is not limited to externalities, such as technology imitation and the mobility of workers, but also captures competition effects.

None of the above effects is necessarily positive. Labor market dynamics may entail negative spillovers such as a brain drain of local talent to foreign firms to the detriment of local firm productivity (see Blalock and Gertler (2008)) or an overall increase in wages irrespective of productivity improvements caused by foreign firms paying higher wages (see Aitken et al. (1996)). Where foreign technology is easily copied, the foreign investor may choose to avoid leakage costs on state-of-the-art technology by transferring technology that is only marginally superior to technology found in the host country (see Glass and Saggi (1998)). This limits the scope for horizontal spillovers via demonstration effects. Higher productivity of foreign affiliates may also lead to lower prices or less demand for the products of domestic competitors. If domestic firms fail to raise productivity in response to the increased competition, they will be pushed up their average cost curves (see Aitken and Harrison (1999) on this market-stealing effect). These partial effects are hard to disentangle empirically and a general measure for horizontal spillover potential is typically used to identify the net effect of all these channels.

Backward spillovers run from the foreign firm to its upstream local suppliers. Thus, even if foreign firms attempt to minimise technology leakage to direct competitors (i.e. a horizontal effect), they may still want to assist their local suppliers in providing inputs of sufficient quality in order to realise the full benefits of their investment. In other words, they want the inputs from the host country to be of lower cost yet similar in quality to inputs in the home country. The foreign firm may even transfer technology to more than one domestic supplier and encourage upstream technology diffusion to circumvent a hold-up problem. Rodríguez-Clare (1996) shows that the backward linkage effect is more likely to be favourable when the good produced by the foreign firm uses intermediate goods intensively and when the home and host countries are similar in terms of the variety of intermediate goods produced. Under reversed conditions, the backward linkage effect could damage the host country's economy. The literature defines backward spillovers both as externalities to suppliers that are unaccounted for and not appropriated by the MNE as well as direct support offered to suppliers (see Moran et al. (2005)). The latter includes technology transfer and assistance that is intentional, possibly even formalized in an agreement. Therefore, what we label backward 'spillovers' in the remainder of the chapter should be understood as the combined impact of both types of effects of foreign presence in downstream industries.

Forward spillovers run from a foreign firm to its local client. In most of the literature forward spillover effects are small or insignificant (see Havránek and Irŝová (2011)). This is likely due to the fact that the literature is largely focused on developing and transition countries where foreign involvement is typically in manufacturing and not in services. Further, Damijan et al. (2013) indicate that foreign affiliates in Eastern Europe are mainly engaged in end-user consumer goods which limits the scope forward spillover effects. However, for Italy, for example, Mariotti et al. (2014) do find important forward spillover effects. Therefore, they should not be ruled out a priori.

The current literature has evolved from the search for an average (positive) effect towards the identification of factors that promote or obstruct spillover effects. We combine two of these factors: time-since-foreign entry (the maturity of foreign firms in the host country market) and spatial proximity.

The maturity of foreign firms has received limited attention. Nevertheless worker mobility is inherently linked to maturity as workers first need to be trained at the foreign firm before taking their skills to domestic firms. Labour poaching by new foreign entrants, on the other hand may entail negative spillover effects. Likewise, vertical spillovers driven by sourcing inputs produced from or supplying inputs to MNEs might not necessarily be instantaneous nor permanent. For the Irish electronics sector Görg and Ruane (2001) find that foreign firms start off with a relatively low extent of local linkages, but as they get accustomed, they proceed to develop more local input linkages. Based on their Volvo case study Ivarsson and Alvstam (2005) conclude that technology transfer to suppliers is more efficient for older MNE plants. Within MNEs technology is also not necessarily easily or rapidly transferred (see Urata and Kawai (2000)). Zhang et al. (2013) find spillover effects to be larger in industries where foreign firms are more mature on average. Merlevede et al. (2014) model the time-since-entry pattern of foreign firms and show that adequately accounting for it reveals new insights in the case of country-wide spillovers. Positive horizontal spillover effects require the presence of mature foreign firms, while positive backward spillover effects arise rapidly following foreign entry, but are transient.[2]

As far as the proximity aspect is concerned, many studies have tried to identify a regional element in spillover effects, but results are mixed (see Crespo and Fontoura (2007)). For Venezuela Aitken and Harrison (1999) found no evidence of local, nor country-wide horizontal spillovers. Mariotti et al. (2014) find that spillovers are strong in knowledge intensive sectors, but proximity is not relevant. Both Keller (2002) and Halpern and Muraközy (2007), on the other hand, do find that spillover effects decline or disappear with distance. Using data for Portugal, Crespo et al. (2009) confirm the importance of proximity between MNEs and domestic firms for FDI spillovers. Finally, Altomonte and Colantone (2008) and Wen (2014) report mixed results, with only some regions recording positive spillovers, suggesting that, aside from distance, other regional characteristics, such as differences in foreign firms' entry and maturity patterns, might be relevant drivers of spillover effects.

It is important to keep in mind that effects may vary among developed and developing countries, as MNEs locate in the former in order to gain technological competences acquired in the host country (see Ivarsson (2002)), while location in the latter is consistent with supplying the local market with superior products (see Cantwell and

[2]In related work, Liu (2008) does distinguish between the level and growth effects of foreign presence, but it is not linked through to maturity.

Piscitello (2007)). For Romania, we do not expect strong distance patterns in horizontal spillovers for the following reasons. First, MNEs will show no particular tendency to agglomerate near domestic firms as local companies do not posses a comparative advantage in their industry (see e.g. Mariotti et al. (2010)).[3] Second, even if located near domestic firms, horizontal spillover effects will be limited because MNEs will prevent technology leakage to local competitors. Likewise, the impact of imitation or reverse engineering will be limited due to low absorptive capacity of domestic Romanian firms (see e.g. Girma (2005)). Therefore we do not expect much of an advantage of being close to an MNE. Because of lower absorptive capacity, spillovers are more likely to manifest themselves through increased competition on the domestic market, potentially leading to negative spillover effects (see Aitken and Harrison (1999)). Since MNEs typically operate nationwide, the impact of distance will be limited. Finally, potential regional horizontal spillover effects might arise from workers moving between MNEs and domestic firms. Labour turnover could be detrimental to domestic firms when their best employees are cherry picked by MNEs (see Sinani and Meyer (2004)), or positive if skills acquired at MNEs move to domestic firms as workers move. As these effects mainly concern mobile, high-skilled workers, the potential impact of distance is again likely to be limited (see Davis and Dingel (2012)).

Backward spillover effects originate from a linkage between a domestic supplier and a foreign client. In this case, we do expect proximity to be more important. Since it is in the interest of the MNE to cooperate with its supplier, backward linkages -in contrast to horizontal relations- imply incentives to share technical knowledge (see Glaeser et al. (1992)). This type of relation will be more sensitive to distance. Giroud (2012) finds that plants located closer to headquarters show higher investment rates because it is easier to monitor nearby plants and acquire information on them. Likewise nearby suppliers should be easier to assist and domestic firms located in the vicinity of MNEs will have access to more face-to-face interactions and will be able to learn faster (see Comin et al. (2012); Keller and Yeaple (2013)). Cristea (2011) confirms the importance of face-to-face communication by showing that increased exports are associated with increased demand for business travel. Keller and Yeaple (2013) also show that

[3] A centrally planned economy was introduced in Romania after World War II, seeking to create a large self-sufficient industrial base, rather than integrating in the global economy by focusing on comparative advantage industries.

inputs highly dependent on non-codifiable knowledge call for more communication favouring nearby locations. Bernard et al. (2014) point to the importance of distance for the formation of linkages. Finally, for a cross-country setting Comin et al. (2012) show that the impact of distance on technology diffusion dies out over time. Depending on how specialized inputs are, forward spillover effects could have a similar transmission mechanism as backward effects, favouring close-by buyers and clusters of cooperating MNEs and domestic firms (see Ivarsson (2002)). Alternatively, if MNEs provide cheaper or better quality inputs to the entire downstream sector, a geographic aspect might be absent. Because in transition economies multinationals have targeted mainly manufactured end-user consumer goods (see Damijan et al. (2013)), we also expect less forward linkages to be established in the first place. For Romania, we therefore expect forward spillover effects to be of lesser importance.

Summarising, from the above we expect that horizontal spillover effects are likely to be sensitive to the maturity of foreign firms, but less to spatial proximity. The key reason for the latter is that horizontal spillovers will manifest themselves mainly via competition and labour channels. These channels will only be affected by distance in a limited way. However, as discussed above, there are good reasons to expect a link between the maturity of foreign firms and horizontal spillover effects. With respect to backward spillover effects, especially the combination of maturity and spatial proximity yields the expectation of faster (in terms of foreign maturity) effects at shorter distances because of MNEs' incentive to assist their suppliers. Since technical upgrading at local suppliers is unlikely to be a smooth process, learning will be faster at closer distances because of easier (face-to-face) contact, visits of engineers, managerial assistance, ... Further, if the probability of buyer-supplier relations tends to decrease with distance, we should also expect smaller average spillover effects over larger distances. It is also unclear whether spillover effects are long-lived. Once its domestic supplier has reached a sufficient level of TFP, more mature foreign firms may scale down technical assistance. Finally, in line with other recent work on Eastern Europe such as Damijan et al. (2013), we do not expect much evidence of forward spillover effects.

3.3 Empirical framework and measurement

3.3.1 Standard measurement

The empirical framework to analyse spillover effects can be seen as an 'augmented' production function, where spillover variables are added to other explanatory variables such as labour, capital, and material inputs. The typical measure employed to identify horizontal or within-industry spillover effects is given by Equation (3.1), where the foreign presence is captured by the share of the foreign firms' output in that industry. For a (domestic) firm i in industry j at time t it is of the following form:

$$\mathrm{HR}_{jt} = \frac{\sum_{i \in j} F_{it} Y_{it}}{\sum_{i \in j} Y_{it}} \tag{3.1}$$

where Y is output and F is the percentage of shares owned by foreign investors. In line with the definition commonly applied by the OECD or the IMF, at least 10% of shares should be owned by a single foreign investor for a firm to be considered as foreign. HR_{jt} in (3.1) measures the share of output that is produced by foreign firms in industry j. Since this spillover variable is built up to industry level from firm-level data, HR_{jt} has the same value for all firms i in industry j at time t. The choice of this measure is not without drawbacks, since we assume a homogeneous reaction of domestic firms within the same industry to the increase in foreign presence in that industry. However, the absorptive capacity of the firm as well as its location with respect to its industry's technological frontier will determine how well it will react to increased foreign presence/competition (see Aghion et al. (2005)). Given the data at hand, we can't easily make these distinctions, and therefore follow the literature in using the aggregated measure.

The definition of the backward spillover variable, BK_{jt}, starts from the horizontal measure and combines it with information from input-output tables as in:

$$\mathrm{BK}_{jt} = \sum_{k \neq j} \gamma_{jkt} * HR_{kt} \tag{3.2}$$

where γ_{jkt} is the proportion of industry j's output supplied to sourcing industry k at time t. The γ's are calculated from (time-varying) IO-tables for intermediate consumption. Inputs sold within the firm's industry are excluded ($k \neq j$) because this is captured by HR_{jt}. Since firms cannot easily, nor quickly switch industries to buy inputs, this approach avoids the problem of endogeneity by using the share of industry output sold to downstream domestic markets k with some level of foreign presence HR_{kt}. Employing the share of firm output sold to foreign firms in different industries would cause endogeneity problems if the latter prefer to buy inputs from more productive domestic firms.

In line with BK_{jt}, we can define FW_{jt} as $\sum_{l \neq j} \delta_{jlt} * HR_{lt}$ where δ_{jlt} is the proportion of industry j's inputs sourced from industry l at time t. As in the case of horizontal spillovers, we proxy the production linkage of domestic and foreign firms belonging to different industries by using an aggregate measure, mainly the average percentage of output supplied by industry j to industry k or from industry l to industry j. However, a firm-specific linkage would be a much more precise way of measuring this relationship, since it is likely that only a small number of firms actually engage in supplier-buyer relationships with foreign firms. Given the lack of such data, we follow the literature by using the aggregated measure. Moreover, since using I-O tables would introduce a bias preventing us from seeing significant results, the presence of average spillovers should confirm a very strong effect for firms active in the supply chain of multinationals.

Unlike the direct effect that FDI has on productivity of firms via M&A (see Arnold and Javorcik (2009)), we are interested in the indirect effect of foreign presence on local firms. However, since we lack the direct supply-chain links between these firms nor can we measure the impact of MNEs on domestic firms via labour or competition/imitation channels, we use the variation in the share of FDI in the relevant market as a proxy for these channels. Therefore, an increase in the foreign share of output would, on average, increase the number of supplier/buyer links as well as local competition and labour mobility, potentially leading to productivity improvements for domestic firms. The lagged growth of spillover variables HR_{jt}, BK_{jt}, and FW_{jt} are thus regressed on the productivity growth of (domestic) firm i in industry j. The size, sign, and significance of the resulting coefficients are then taken as evidence of spillover effects. The estimation is performed in first differences in order to avoid unobserved firm level characteristics such as managerial performance and other factors from driving the effect. We do however include firm age and firm size as controls, since older/larger firms tend to grow slower than younger/smaller firms. Finally, country/region, industry and year fixed effects are used in order to take into account unobservable factors which could impact TFP growth.

As indicated above, FDI spillovers are commonly analysed in a production function framework. Firm level total factor productivity is obtained in a first step estimation and in a second step the FDI spillover variables together with some further controls are treated as additional 'input' explaining domestic firms' productivity. The careful estimation of the production function is thus an important building block in the analysis. The basic problem in estimating productivity is that firms react to firm-specific productivity shocks that are not observed by the researcher.

Griliches and Mairesse (1995) provide a detailed account of this problem and make the case that inputs should be treated as endogenous variables since they are chosen on the basis of the firm's unobservable assessment of its productivity. The semi-parametric approaches by Olley and Pakes (1996) (OP) and a more recent modification of it by Levinsohn and Petrin (2003) (LP), and the dynamic panel data approach by Blundell and Bond (1998) (DPD) are alternative methodologies to overcome the endogeneity bias in estimating production functions. Both types of methodologies have been widely used in the recent literature on firm level heterogeneity for derivation of total factor productivity measures. More recently, Ackerberg et al. (2008) (ACF) argue that, while there are some solid and intuitive identification ideas in the papers by Olley and Pakes (1996) and Levinsohn and Petrin (2003), their semi-parametric techniques suffer from collinearity problems casting doubt on the methodology. They suggest an alternative methodology that make use of the ideas in these papers, but do not suffer from these collinearity problems. We therefore use the ACF estimator to obtain our indicator of total factor productivity (TFP). A measure of TFP for firm i in industry j at time t is obtained as the difference between output and capital, labour, and material inputs, multiplied by their estimated coefficients:

$$tfp_{ijt} = Y_{ijt} - \widehat{\beta}_{lj} l_{ijt} - \widehat{\beta}_{kj} k_{ijt} - \widehat{\beta}_{mj} m_{ijt} \tag{3.3}$$

Following the literature (e.g. Javorcik (2004)), in the second step tfp_{ijt} is related to a firm specific effect, a vector of spillover variables, $\mathbf{FDI}_{jt}$, and firm and industry level controls, $\mathbf{Z}_{i(j)t}$. Note that (3.4) now pools firms from all industries together in one large panel, whereas (3.3) is estimated by industry. This approach is what Havránek and Irŝová (2011) define as best practice.

$$\Delta tfp_{ijt} = \alpha_i + \Psi_1 f\left(\Delta \mathbf{FDI}_{jt-1}\right) + \Psi_2 \Delta \mathbf{Z}_{i(j)t} + \xi_{ijt} \tag{3.4}$$

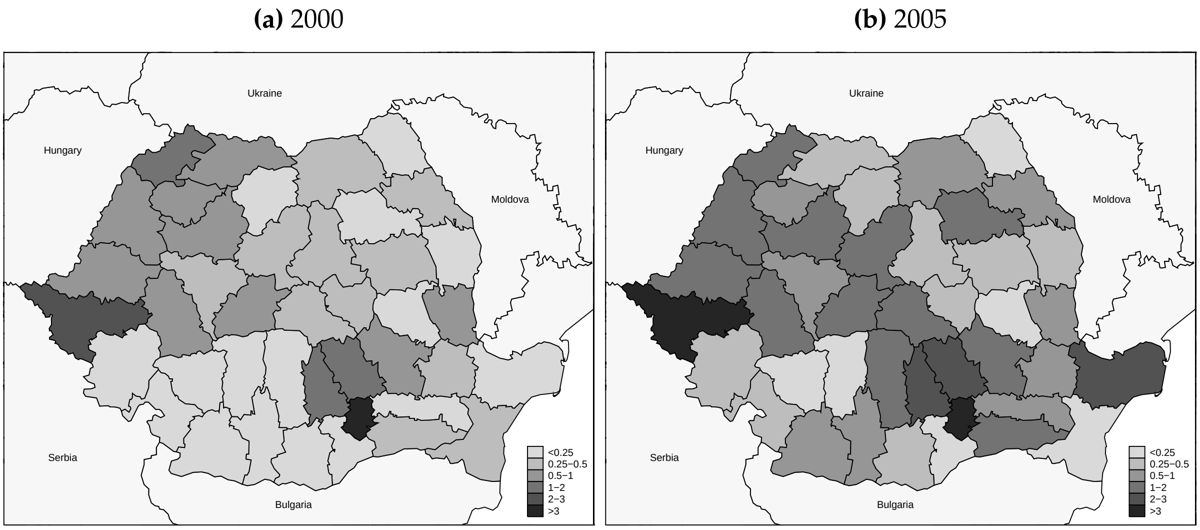
Ψ_1 in equation 3.4 allows us to identify the sign, size, and significance of the impact of foreign presence on the productivity of local firms. In the next section we define our vectors of spillover variables, $\mathbf{FDI}_{jt}$, and control variables, $\mathbf{Z}_{i(j)t}$.

3.3.2 A regional dynamic approach to spillovers

In this section we introduce our regional dynamic approach to the identification of spillover effects in the above framework. Whereas we do not have information on the exact location of foreign firms in our dataset (cf. infra), we do know in which NUTS 3 region a firm is located. We use the NUTS 3 classification as our regional dimension. At this level Romania is divided in 42 territorial units, i.e. 41 counties and the capital Bucharest. Our data, however, do not allow us to discriminate between Bucharest and the surrounding county Ilfov. Therefore, we have 41 territorial units in our analysis. The NUTS 3 level is appropriate because we find quite some heterogeneity between regions in terms of foreign presence and larger regional aggregates (e.g. the NUTS 2 division) would hide this heterogeneity. The NUTS 3 division also follows an original administrative structure for which we are able control by means of region fixed effects.

Figure 3.1 plots the share of industry output produced by foreign firms in a region from the total country-wide output of that industry (the average over manufacturing industries in a given region is plotted). The figure clearly shows that foreign presence is not uniformly spread across the country, but is mainly concentrated along border regions and in the capital, as consistent with predictions from the literature (see e.g. Mariotti and Piscitello (1995); McCann and Acs (2011); Spies (2010)). Over the sample period the dispersion of FDI intensity at the regional level (measured by the standard deviation) has increased from 1.7 in 2000 to 1.9 in 2005. Given the potential contribution of spillover effects to economic growth, it is important to test whether these regional differences in FDI intensity have an impact on where spillover effects are generated, whether spillover effects differ in size across regions and whether and how spillover effects spread from one region to another. We now introduce our methodological approach. We first introduce the distance dimension, and then interact it with the time-since-foreign-entry dimension.

Figure 3.1: Industry output produced by foreign firms in a NUTS 3 region as a share of total country-wide output of the industry (the average over manufacturing industries in a given region is plotted)



3.3.2.1 Spillovers and distance

An often used measure to capture within-industry spillover effects HR_{jt} in (3.1) is the share of output that is produced by foreign firms in industry j. Alternatively one could use the number of foreign firms in the respective industry (see Altomonte and Pennings (2009)). However, given that we expect bigger firms to have a greater impact on domestic productivity changes and potentially have a larger geographical reach as well, we require a measure capturing the intensive margin of foreign presence and thus use output instead of the number of foreign firms.

For a given firm i in industry j in region r at time t we can break HR_{jt} down into different 'geographical' subcomponents as follows:

$$\mathrm{HR}_{jt} = \frac{\sum\limits_{i \in j} F_{it} Y_{it}}{\sum\limits_{i \in j} Y_{it}} \tag{3.5}$$

$$= \frac{\sum R_{it} F_{it} Y_{it}}{\sum Y_{it}} + \frac{\sum NB_{it} F_{it} Y_{it}}{\sum Y_{it}} + \frac{\sum \left(1 - R_{it} - NB_{it}\right) F_{it} Y_{it}}{\sum Y_{it}} \tag{3.6}$$

where R_{it} indicates whether firm i is located in region r, and NB_{it} indicates whether firm i is located in a contiguous region of r. Finally, $(1 - R_{it} - NB_{it})$ will equal 1 if firm i is located in a further-away non-neighbouring region, i.e. a rest-of-country category.[4] We refer to this regional decomposition as 'distance'.

From (3.6) it becomes clear that introducing HR_{jt} as a single variable in a regression involves the implicit assumption that the spillover intensity, as measured by the coefficient obtained on HR_{jt}, is the same within and across regions. In our empirical analysis we relax this assumption and allow the coefficients to differ between the different subcomponents in (3.6), obtaining estimates for region, neighbour and rest-of-country components respectively. A regional definition for BK follows from (3.2) above. Since we only have input-output tables at the country-level, we assume that technical coefficients are similar across regions and equal to those derived from country-level input output tables.[5]

We differ from earlier literature by explicitly structuring the regional dimension as a decomposition of the traditional nation-wide definition. By introducing all three subcomponents of (3.6) in our analysis we also differ from part of the regional FDI spillover literature that does not allow for cross-regional spillovers. Often only the first term of the decomposition is included among the regressors, thereby implicitly assuming that spillovers are confined to region boundaries and do not cross borders. This runs counter to Halpern and Muraközy (2007) who find that horizontal spillovers vary with distance, but do not disappear. It also runs counter to macro-spillover studies

[4] Clearly, (3.6) could be further decomposed in a straightforward manner to account for second- or even higher-order neighbours. However, since adding second-order neighbour effects does not affect our estimates with respect to region, neighbour, and rest-of-country, we focus on the three aforementioned dimensions. These results are available on request.

[5] The IO-tables are only available at the country-wide level. Therefore we need to assume that technical coefficients do not differ across regions. As far as technology is not too different across regions, the impact on the results should be limited. Since technological coefficients refer to the share of a broad input category used in the production process, we do not expect large differences across regions. Nevertheless one should bear this caveat in mind.

as Keller (2002) who finds that spillovers between countries are declining with distance. We further distinguish ourselves from the existing literature by modelling cross-region spillover effects. Consider the following reformulation of (3.6), where summation is over firms i in industry j:

$$\begin{aligned}\mathrm{HR}_{jt} &= \frac{\sum R_{it}F_{it}Y_{it}}{\sum R_{it}Y_{it}} \times \frac{\sum R_{it}Y_{it}}{\sum Y_{it}} \\ &+ \frac{\sum NB_{it}F_{it}Y_{it}}{\sum NB_{it}Y_{it}} \times \frac{\sum NB_{it}Y_{it}}{\sum Y_{it}} \\ &+ \frac{\sum \left(1-R_{it}-NB_{it}\right)F_{it}Y_{it}}{\sum \left(1-R_{it}-NB_{it}\right)Y_{it}} \times \frac{\sum \left(1-R_{it}-NB_{it}\right)Y_{it}}{\sum Y_{it}}\end{aligned} \tag{3.7}$$

Studies that focus on regional spillovers typically apply the traditional nation-wide definition to their regional spillover variable. This spillover variable is constructed as output produced by foreign firms in industry j in region r as a share of total *regional* industry j output, i.e. the first part of the first term in (3.7), rather than as a share of country-wide industry j output.

The definition of an appropriate measure relates to one's idea about spillover potential (this is what the variables are intended to capture). Consider two regions A and B. In region A 10 out of the total of 100 units are produced by foreign firms, while in region B 10,000 out of the total 100,000 units are produced by foreigners. $\sum R_{it}F_{it}Y_{it} / \sum R_{it}Y_{it}$ is appropriate if one believes that the spillover potential is the same in both regions. In the former case, spillovers should be thought of as limited to the region level since it is difficult to carry this definition through to the cross-region level. Suppose regions A and B are neighbours. Following a logic of relative within territorial unit presence, the spillover from neighbours could be measured as $\sum NB_{it}F_{it}Y_{it} / \sum NB_{it}Y_{it}$. This results in a value of 0.10 for both region A and B. However, it seems counter-intuitive that the spillover potential from region A to B would equal the spillover potential from B to A. This is not the case when using the second subcomponent of our decomposition in (3.6). In our example, this results in a spillover potential from A to B of $10/100,100$ and a spillover potential from B to A of $10,000/100,100$. These values seem better aligned with the cross-region spillover potential one would expect. Regions where a larger share of the foreign activity is located carry a larger spillover potential and therefore should be reflected in the measure employed in empirical work. Therefore, we apply the decomposition in (3.6) and allow for coefficient heterogeneity for the different subcomponents.

3.3.2.2 Spillovers and time-since-foreign-entry

Abstracting for the moment from the geographical dimension in HR_{jt} discussed above, (3.1) hides another important dimension that deserves attention. Indeed, upon closer inspection various spillover transmission channels imply an impact of foreign maturity as discussed above. Following Merlevede et al. (2014) we therefore introduce a time-since-foreign-entry pattern in our analysis of spillover effects. Consider the following alternative breakdown of (3.1):

$$\mathrm{HR}_{jt} = \frac{\sum \widetilde{F}_{it}^{1} Y_{it}}{\sum Y_{it}} + \frac{\sum \widetilde{F}_{it}^{2} Y_{it}}{\sum Y_{it}} + ... + \frac{\sum \widetilde{F}_{it}^{n} Y_{it}}{\sum Y_{it}} \tag{3.8}$$

where $\widetilde{F}^{x}$ is a variable indicating foreign ownership status *and* entry timing. $\widetilde{F}_{it}^{x}$ equals the percentage of shares owned by foreign investors in firm i if at least 10% of shares were owned by a single foreign investor in year $t-x+1$ and firm i was not foreign owned in year $t-x$, i.e. the investment took place between $t-x+1$ and $t-x$. So $\widetilde{F}_{it}^{x}$ is set to the percentage of shares owned by foreign investors if:

$$\left(\sum_{v=0}^{x-1} F_{i,t-v} = x\right) \wedge \left(\sum_{w=x}^{\infty} F_{i,t-w} = 0\right) \tag{3.9}$$

HR_{jt} is thus broken down into HR_{jt}^{1}, HR_{jt}^{2}, and so on, along the lines of foreign entry timing (note the difference with pure calendar time or taking lags of HR_{jt}). A time-since-foreign-entry definition for BK_{jt}^{x} follows from (3.2) and (3.8) above (a definition for forward spillover variables straight forwardly follows):

$$\mathrm{BK}_{jt}^{x} = \sum_{k \neq j} \gamma_{jkt} * HR_{kt}^{x} \tag{3.10}$$

We combine the regional and time-since-foreign-entry aspects into a single comprehensive approach, as summarized in Table 3.1. A failure to find cross-region spillovers on the basis of aggregate variables as in (3.6) could be due to the fact that time-since-foreign-entry has been neglected, rather than that these spillovers are truly non-existent. Indeed, as indicated above the time-since-entry pattern for within-region spillovers may well be different from the time-since-entry pattern for cross-region or rest-of-country spillovers, since it may take more time for domestic firms to absorb spillovers from foreign firms in further-away regions. Further note that some papers limit the scope of spillovers to the boundaries of a region by construction of the spillover variables (see Nicolini and Resmini (2010)). We model the potential regional pattern explicitly and combine it in a novel way with time-since-foreign-entry effects.

Table 3.1: Coefficient heterogeneity in a regional time-since-foreign-entry approach

Region/Time-since-foreign-entry	t	t-1	t-2	t-3	t-4+
same region	$c_{R,t}$	$c_{R,t-1}$	$c_{R,t-2}$	$c_{R,t-3}$	$c_{R,t-4+}$
neighbour region	$c_{NB,t}$	$c_{NB,t-1}$	$c_{NB,t-2}$	$c_{NB,t-3}$	$c_{NB1,t-4+}$
rest of country	$c_{RoC,t}$	$c_{RoC,t-1}$	$c_{RoC,t-2}$	$c_{RoC,t-3}$	$c_{RoC,t-4+}$

3.3.3 Empirical framework

Our empirical approach detailed in (3.4) above closely follows the existing literature described earlier. We estimate domestic industry production functions using the ACF estimator separately for each NACE[6] 2-digit manufacturing industry j in the period 1996-2005. Capital, labour, and material inputs elasticities are thus treated as industry-specific. Firms that are foreign at some point in time are excluded from the estimation.

The vector of spillover variables ($\mathbf{FDI}_{jt-1}$) covers different horizontal and vertical spillover variables described above. More specifically, HR, BK, and FW are decomposed in function of both the geography and time-since-foreign-entry dimensions (for clarity we drop industry and time subscripts in (3.11)). We consider three different regional dimensions: within-region spillovers, HR_reg^{t-x}, first-order neighbour spillovers, HR_nb^{t-x}, and spillovers from the regions that make up the rest-of-country, HR_roc^{t-x}. Considering the time span of our dataset (1996-2005, *cf. infra*) we opt to include HR_x^{t} to HR_x^{t-3} and create a variable HR_x^{t-4+} which aggregates all foreign firms that have been present for at least four full years on the domestic market, hence the summation from t to $t-4+$ in (3.11). Since we do not have information on exact dates of foreign entry prior to 1996, the time span of the dataset for the estimations is reduced to 2001-2005 because of missing values.

[6] NACE stands for *Nomenclature générale des Activités économiques dans les Communautés Européennes.*

$$\begin{aligned}\Psi_1 f\left(\mathbf{FDI}_{jt-1}\right) = &\sum_{x=0}^{4+}\left(\alpha_{h,reg}^{t-x} HR\_reg^{t-x} + \alpha_{h,nb}^{t-x} HR\_nb^{t-x} + \alpha_{h,roc}^{t-x} HR\_roc^{t-x}\right)\\ &+ \sum_{x=0}^{4+}\left(\alpha_{b,reg}^{t-x} BK\_reg^{t-x} + \alpha_{b,nb}^{t-x} BK\_nb^{t-x} + \alpha_{b,roc}^{t-x} BK\_roc^{t-x}\right)\\ &+ \sum_{x=}^{4+}\left(\alpha_{f,reg}^{t-x} FW\_reg^{t-x} + \alpha_{f,nb}^{t-x} FW\_nb^{t-x} + \alpha_{f,roc}^{t-x} FW\_roc^{t-x}\right) \quad (3.11)\end{aligned}$$

Through the vector $\mathbf{Z}_{i(j)t}$ we control for competition within the industry, measured by the Herfindahl index, import competition in the industry, the share of intermediates supplied in total industry output, and firm size and age. Further we use the region-industry share of national industry activity and the region's share of national manufacturing activity to control for region and region-industry agglomeration effects. Specification (3.4) is first-differenced and estimated by OLS. We also introduce industry (α_j), region (α_r), and time dummies (α_t) in the first-differenced specification to account for unobserved factors that could be driving *growth* performance at the region or industry level. This results in (3.12) as final specification to be estimated. Since $\mathbf{FDI}_{jt}$ and some control variables are defined at the industry level, and estimations are performed at the firm level, standard errors need to be adjusted (Moulton (1990)). Standard errors are therefore clustered for all observations in the same region, industry and year (see Javorcik (2004)).

$$\Delta tfp_{ijrt} = \Psi_1' \Delta f\left(\mathbf{FDI}_{jt-1}\right) + \Psi_2' \Delta \mathbf{Z}_{i(j)t} + \alpha_t + \alpha_j + \alpha_r + \epsilon_{ijrt} \quad (3.12)$$

3.4 Data

For this chapter we use a pilot firm-level data for a panel of Romanian manufacturing firms during 1996-2005. We do not employ the larger *AUGAMA* due the fact that we did not yet have it available and we wanted to focus only on one country for this analysis. Since most foreign investment entered the country after 1996, Romania makes a very good candidate to study the dynamic impact of recent foreign investment on domestic firm productivity (see Hilber and Voicu (2010)). Moreover, in a bid to bring all regions to a similar level of economic development by creating a homogeneous working class, the country has undergone a massive forced industrialisation for about two decades prior to 1990s (see Ronnås (1984)). This was at least partially successful due to the wide dispersion of natural resources across the country. Although the process did not level out all regional differences that developed over centuries, it did reduce some disparities and created a more standardised structure of counties with strong manufacturing bases and improved urban networks. The fall of communism in 1989 was accompanied by a reversal of at least some of these policies, with severe restructuring in the industrial set-up of the country as a consequence. As a result, regional inequalities have risen due to both market forces and a decrease in state intervention (see Antonescu (2012)). While we do not have data for the 1990-1995 period, our sample still covers the early stages of the transition period and therefore a relatively homogeneous regional set-up.

Our firm-level data are taken from Amadeus, a Bureau van Dijk Electronic Publishing database. Amadeus is a pan-European database of financial information on public and private companies. Every month Bureau Van Dijk issues a new DVD with updated information. A single issue of the DVD contains only the latest information on ownership and firms that go out of business are dropped from the database fairly rapidly. Furthermore, because Bureau Van Dijk updates individual ownership links between legal entities rather than the full ownership structure of a given firm, the ownership information on a specific DVD-issue often consists of a number of ownership links with different dates, referring to the last verification of a specific link. To construct our dataset with entry, exit, and time-specific foreign entry in local Romanian firms, we therefore employed a series of different issues of the database. However, since ownership information is gathered at irregular intervals, we do not have ownership information for all firm-owner-year combinations.[7] Given these specificities of Amadeus, we first created a dataset at the firm-owner-year-level with the available information from Amadeus. We then filled out missing firm-owner-year-entries under restriction that the full ownership structure cannot exceed 100%. In case of time gaps between entries for the same owner-firm combination but with a different share-size we assume that changes show up immediately in the database. Finally, we then fill out the gaps

[7] Identifying the same owner in different issues is not always straightforward since an ID is only listed in case the owner is a firm that is listed in Amadeus itself. For all other owners matching is done on the basis of the name. Differences in spacings, plurals, addition to the name of a company-type, the use of characters specific to Romanian versus standard Roman characters in different issues are corrected for.

with the older information.[8] We focus on a sample of firms that report unconsolidated data.

Data are deflated using industry price level data at NACE rev.1.1 2-digit level. These are taken from the Industrial Database for Eastern Europe from the Vienna Institute for International Economic Studies (2008) and from the Romanian National Statistical Office (RNSO) (2005). Real output Y is measured as operating revenues deflated by producer price indices of the appropriate NACE industry; real material inputs M, are deflated by a weighted intermediate input deflator where the industry-specific weighting scheme is drawn from the IO tables. Labor L is expressed as the number of employees. Real capital K is measured as tangible fixed assets, deflated by the average of the deflators for the following five NACE industries: machinery and equipment (29); office machinery and computing (30); electrical machinery and apparatus (31); motor vehicles, trailers, and semi-trailers (34); and other transport equipment (35) (see Javorcik (2004)). Detailed IO tables containing 105 (59 manufacturing) sectors for the period 1996–2005 were obtained from the RNSO.

We restrict the dataset to firms with on average at least 5 employees over the sample period. The dataset is further trimmed for outliers by removing the top and bottom percentiles of the annual growth rates of real operating revenues, real capital, labour, and real material inputs.[9] The share of foreign firms in the total number of sample firms steadily increased from 17% to 24% (10% to 15% if small firms are not excluded). Table 3.2 lists summary statistics for both domestic and foreign firms. The stylized facts commonly found in the literature are confirmed in our dataset. Foreign firms are larger in terms of employment and capital, produce more output and are more productive. The productivity bonus of foreign over domestic firms is 26% in case of the ACF methodology.

Based on 15164 industry-region-year observations Table 3.3 reveals that on average over industries, regions, and years about 25% of output is produced by foreign firms. On average 20.7 percentage-points of it is produced in the rest-of-country. Within-

[8] e.g.

	Amadeus	immediate
2000	40	40
2001	.	40
2002	50	50

[9] If the 'outlier' is the first or last observation for a specific firm and other data points appear 'normal', the other firm-year data are kept. If not all observations for this firm are dropped from the dataset.

Table 3.2: TFP summary statistics

	All firms n=133154		Domestic firms n=105854		Foreign firms n=27300	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
ln(real output)	13.74	1.90	13.53	1.84	14.52	1.94
ln(employment)	3.08	1.47	2.93	1.40	3.67	1.57
ln(capital)	12.08	2.32	11.82	2.26	13.06	2.29
ln(tfp) ACF	5.74	1.52	5.69	1.52	5.95	1.47

Summary statistics for foreign and domestic firms.

region and neighbouring-region aggregates account for 1 and 3.2 percentage-points on average. Taking into account the median value, the interquartile range, minima, and maxima suggests a reasonable variation in the data for within-region and neighbouring-region aggregates. Some industries are dominated by foreign firms concentrated in a single region as testified by the maximum value of 80.3%. Backward spillover variables show a more mitigated pattern as they are a weighted average of downstream horizontal variables. Correlations between region-neighbour, region-rest-of-country, and neighbour-rest-of-country are virtually zero at 0.03, -0.05, and -0.07. Table 3.4 shows that for each regional aggregate foreign firms which have been present for at least four years account on average for the largest share of output produced by foreign firms in the industry and region. One should bear these numbers in mind when interpreting the results below.

Table 3.3: Summary statistics for the distance decomposition

	Mean	Median	IQR	P25	Min	Max
horizontal						
region	1.0	0.0	0.2	0.0	0.0	80.3
neighbour	3.2	0.6	3.1	0.0	0.0	83.5
rest-of-country	20.7	17.5	21.5	8.5	0.0	87.2
backward						
region	0.6	0.1	0.3	0.0	0.0	38.7
neighbour	1.7	0.7	1.2	0.3	0.0	26.8
rest-of-country	14.4	14.5	9.8	9.3	0.1	62.9

Table entries refer to the share of total country-wide industry output produced by foreign firms in the region, neighbouring region, and rest-of-country regional aggregates. Numbers are based on 15164 industry-region-year observations.

Table 3.4: Summary statistics for the distance and time-since-foreign-entry decomposition

	t	t-1	t-2	t-3	t-4+
region	0.04	0.12	0.14	0.14	0.78
neighbour	0.13	0.37	0.45	0.47	2.59
rest-of-country	0.78	2.40	2.98	3.22	16.52

Table entries refer to the share of total country-wide industry output produced by foreign firms of a given maturity (indicated in column headings) in the region, neighbouring region, and rest-of-country regional aggregates. Numbers are averages over 15164 industry-region-year observations.

3.5 Results

This section presents results of different sets of estimations. For the sake of clarity and in order to keep the tables manageable we do not report the results on the control variables here. If not mentioned otherwise, we include firm size and age, industry competition, competition from imports in the industry, the share of intermediates supplied in total industry output, and time, industry and region dummies as control variables. We consider horizontal, backward and forward spillovers. The latter turn out to be insignificant and for reasons of clarity and space we only report forward spillover results in the first results table and omit them from further tables. We think of them as additional control variables. We first discuss results that only focus on the distance decomposition of the spillover variables. Then we combine the distance and time-since-foreign-entry decompositions of the spillover variables and present our main results. We refer the reader to Merlevede et al. (2014) for results on the time-since-entry decomposition by itself.

3.5.1 Distance decomposition

Table 3.5 presents results for the distance decomposition. The table contains both the estimated coefficients for spillover effects over different distances and an F-test for the equality of the estimated coefficients over distance. We observe that all horizontal spillover coefficients are significant. The estimated coefficients increase with distance which suggests that while positive spillover channels dominate, negative effects such as increased competition are somewhat stronger for nearby foreign firms. We cannot reject the different coefficients to be equal, however. Backward spillover effects are only statistically significant from firms located further away, specifically in the rest-of-country area. This result could reflect that backward spillover effects originate from firms concentrated in a small number of regions and from there spread to other regions of the country. For the average Romanian region these regions would pertain to the rest-of-country category. Nonetheless, the test for equality of coefficients again is unable to reject the null hypothesis that coefficients are equal. The third column shows that forward spillovers are insignificant, a finding which is not uncommon in the literature (see e.g. Damijan et al. (2013)).

Table 3.5: Results when applying the distance decomposition to spillover variables

	Horizontal	Backward	Forward
same region	0.834*	1.577	-1.924
	[0.461]	[1.058]	[1.261]
neighbouring region	1.117***	0.304	1.465
	[0.287]	[0.916]	[0.915]
rest of country	1.438***	1.355***	0.082
	[0.141]	[0.375]	[0.322]
Reg=NB=RoC	1.268	0.692	2.325*
Observations		49,074	
R-squared		0.05	

The table presents both the regression results for the geographical component alone and a test for the equality of coefficients at regional, neighbour and rest-of-country level with the corresponding F-test. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. F-test * rejected at 10%; ** rejected at 5%; *** rejected at 1%.

3.5.2 Distance and time-since-foreign-entry decomposition

Results in Table 3.5 do not account for the time-since-foreign-entry dimension. Table 3.6 therefore shows our central result that combines the regional dimension with the time-since-foreign-entry dimension. From specification (1) we infer that horizontal spillover effects (column a) are negative in the first years after entry, but turn positive when foreign firms 'mature'. The intensity of horizontal spillovers again does not seem to vary in terms of FDI location (nearby or far away), with very similar spillover coefficients for same region, neighbouring region, and rest-of-country aggregates. This is confirmed by the results in column 1 of Table 3.7, as we cannot reject the equality of the horizontal spillover coefficients across the regional dimension. Therefore, for a similar share of foreign sales at each geographic level, domestic firms experience similar productivity effects whether foreign firms are located in their own region, in a neighbouring region or in the rest-of-country.

The time-since-entry dimension suggest that domestic firms experience an initial

Table 3.6: Horizontal and Backward spillovers

	(1)		(2)		(3)	
	(a) Horizontal	(b) Backward	(a) Horizontal	(b) Backward	(a) Horizontal	(b) Backward
entry in t						
same region	-0.574	18.231***	-0.427	19.852***	-0.164	20.151**
	[0.953]	[6.921]	[0.964]	[7.073]	[1.030]	[9.875]
neighbouring region	0.889	-0.346				
	[0.546]	[5.192]				
rest of country	0.183	2.712	0.421	2.405		
	[0.332]	[1.740]	[0.307]	[1.682]		
entry in t-1						
same region	-2.483***	10.985***	-2.394***	11.735***	-2.621***	9.943***
	[0.814]	[2.929]	[0.820]	[2.937]	[0.861]	[3.327]
neighbouring region	-1.574***	5.201				
	[0.577]	[3.750]				
rest of country	-1.089***	8.464***	-1.104***	8.176***		
	[0.374]	[1.094]	[0.354]	[1.062]		
entry in t-2						
same region	-1.427***	3.343***	-1.394***	3.579***	-1.528**	2.570**
	[0.539]	[1.122]	[0.541]	[1.134]	[0.618]	[1.287]
neighbouring region	-2.461***	6.227**				
	[0.562]	[2.820]				
rest of country	-1.427***	4.224***	-1.549***	4.325***		
	[0.269]	[0.745]	[0.263]	[0.726]		
entry in t-3						
same region	0.263	3.845***	0.237	3.970***	-0.027	2.760*
	[0.437]	[1.415]	[0.439]	[1.464]	[0.490]	[1.453]
neighbouring region	0.084	2.784				
	[0.525]	[2.456]				
rest of country	0.583***	4.287***	0.555***	4.126***		
	[0.193]	[0.971]	[0.186]	[0.939]		
entry in t-4 or earlier						
same region	1.546***	0.525	1.539***	0.543	1.212***	-0.192
	[0.412]	[1.214]	[0.416]	[1.215]	[0.453]	[1.350]
neighbouring region	1.814***	-1.368				
	[0.315]	[1.025]				
rest of country	2.108***	-0.439	2.053***	-0.594		
	[0.155]	[0.435]	[0.152]	[0.418]		
Observations	49,074		49,074		49,074	
R-squared	0.071		0.070		0.045	

Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

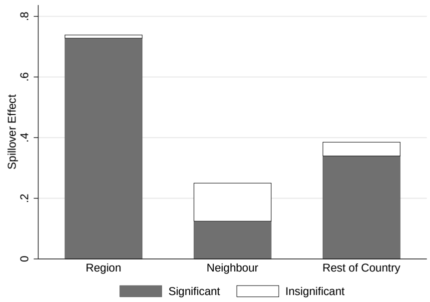
negative impact following foreign firms' entry, which could be due to considerable negative competition effects or labour cherry picking. However, once foreign firms have been present for a sufficiently long period in the domestic economy (entry in t-4 or earlier), positive spillover effects do arise and they are sufficiently large to compensate for earlier negative effects. In line with Table 3.5, point estimates again hint at a larger (or less negative) impact for foreign firms located in the rest-of-country versus the own region, confirming that distance might offer some protection against negative effects.

Backward spillover effects in specification (1) column (b), on the other hand, show a larger sensitivity to the distance between domestic and foreign firms. Both distance and time-since-foreign-entry are important determinants of the magnitude of the backward spillover effect accruing to domestic firms. Taking time-since-foreign-entry into account, we find that entry of foreign firms in the same region entails a faster positive contribution to the domestic firms' productivity. Positive spillover effects from further away foreign firms do manifest themselves, but take more time to do so. To better illustrate the impact of backwards spillovers, in Figure 3.2 we plot the total expected contribution to a domestic firm's TFP level of a foreign entrant producing 2 per cent of downstream output annually, i.e. the backward spillover effect. From panel (a) of Figure 3.2 we clearly infer that a domestic firm would prefer to see the foreign firm enter in its own region, as the expected backward spillover effect over the foreign firm's life time is at least twice as large there compared to those from the other two regional dimensions. The F-tests in Table 3.7 confirm a statistically significant bonus of being close to foreign clients. Moreover, specifications (2) and (3) in Table 3.6 show that the 'being close' bonus is not driven by a specific correlation structure between the different elements of the regional decomposition, since the exclusion of the *neighbour* variables or the *neighbour* and $rest-of-country$ variables does not affect the results of the *region* variable. Panel (b) of Figure 3.2 shows that positive and significant backward spillover effects are arising from foreign firms with limited maturity (this is in line with Merlevede et al. (2014)). The time-since-foreign-entry pattern shows a strong initial impact of supplying intermediate inputs to foreign firms located in the host region that decays over time and disappears for more mature foreign firms. The patterns for backward spillover effects from foreign firms in neighbouring regions or in rest-of-country regions are similar and not statistically different from one another (cf. F-tests in Table 3.7). These spillover effects take more time to manifest themselves and also disappear once a foreign firm has been present for a longer period in the host country. This explains the difference with our findings in Table 3.5 where we did not account for the time-since-foreign-entry pattern. Since only recent foreign entrants drive regional differences in backward spillover effects, lumping all firms together in terms of time-since-foreign-entry prevents us from observing this effect.

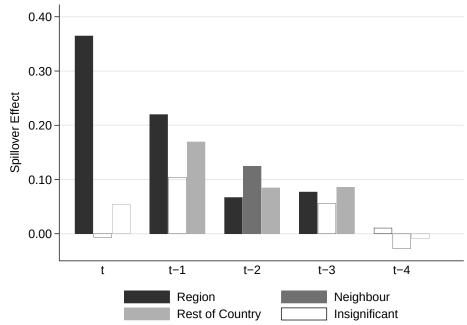
These findings are consistent with our expectations. Horizontal spillover effects

Figure 3.2: Backward spillover effect of a foreign firm producing 2 per cent of downstream output

(a) Distance impact



(b) Distance and time-since-foreign-entry impact



The figure shows the actual contribution to a domestic firm's TFP level of a foreign firm each year producing 2 per cent of downstream output.

Table 3.7: Test for equality of coefficients

F-test	Horizontal	Backward
Reg_t=NB_t=RoC_t	1.117	2.536*
Reg_t=NB_t	1.838	4.280**
NB_t=RoC_t	1.241	0.330
Reg_t=RoC_t	0.558	4.813**
Reg_{t-1}=NB_{t-1}=RoC_{t-1}	1.314	0.777
Reg_{t-2}=NB_{t-2}=RoC_{t-2}	1.656	0.513
Reg_{t-3}=NB_{t-3}=RoC_{t-3}	0.606	0.198
Reg_{t-4+}=NB_{t-4+}=RoC_{t-4+}	1.248	0.834

The table presents a test for the equality of coefficients at regional, neighbour and rest-of-country level with the corresponding F-test value. * rejected at 10%; ** rejected at 5%; *** rejected at 1%.

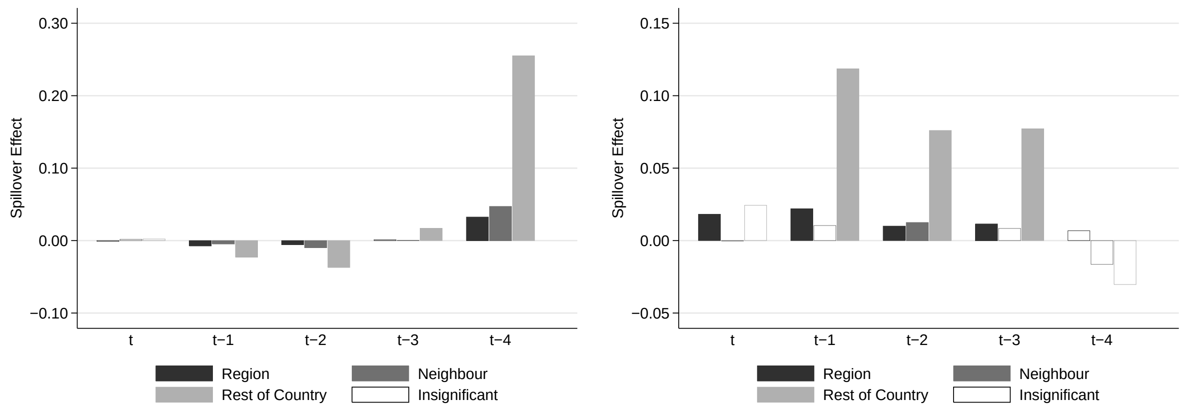
vary with distance on the basis of point estimates, but differences are not statistically significant. There is thus some indication that the closer to a foreign firm, the smaller the spillover effect will be because of a stronger initial negative competition and labour market effects. However, because MNEs typically compete nation-wide and high-skilled workers in demand by foreign firms are likely to be highly mobile, the distance effect is not significant. Backward spillover effects, on the other hand, are faster and larger for nearby domestic firms. This is because MNEs have an incentive to assist their local suppliers and learning is faster at closer distances due to easier (face-to-face) contact and visits of engineers. This does not rule out effects over longer distances, but we find them to be smaller and to take more time to manifest themselves. Backward spillover effects are limited in time-since-foreign-entry, as linkages with foreign clients no longer stimulate technological upgrading after considerable initial improvements of productivity.

Since Figure 3.2 does not convey much about the actual in-sample contribution to productivity for Romanian firms, we offer two additional views of our results. First, we calculate the period-average contribution to the TFP-level of the average domestic firm of the different spillover variables, as shown in Figure 3.3. This is achieved by multiplying the average amount of foreign presence at all three regional levels with their respective coefficients. Moreover, it gives a better indication of what foreign entry has brought for Romanian firms over 2001-2005. Figure 3.3 suggests that the average contribution of the rest-of-country horizontal spillover is larger than those from either the own or neighbouring regions. This is due to the combination of similar sized coefficients as well as the fact that on average the amount of foreign activity in the rest-of-country regional aggregate is much larger than in either the own or neighbouring regions. From panel (a) of Figure 3.3 it is also clear that the positive contribution after four years of foreign presence outweighs the negative effects foreign firms generate in the first three years. This confirms the fact that after an initial adjustment period, domestic firms do benefit from the presence of foreign firms in their own industry (cf. Merlevede et al. (2014)). Backward spillover effects are limited to the first years after foreign entry. For the average domestic firm the rest-of-country backward spillover effect is the largest because most of the foreign firms are located there, but the within-region contribution is non-negligible.

Second, since the previous result does not take into account regional heterogeneity in Romania, we recalculate the expected impact in terms of spillover effects on TFP from moving a domestic firm between regions, as seen in Figure 3.4. We first 'relocate' a firm from Vaslui (denoted as VS), a subpar performing region in the North-East at the border with Moldova, to Timiş (denoted as TM), a regional hub in the South-West of Romania. This move increases the total spillover effect on the log of the TFP-level of the firm with 0.036. The effect is for about two thirds driven by an increase in the horizontal spillover effect which in turn is driven mainly by an increase in the within-region effect (the negative neighbour and positive rest-of-country components are smaller and cancel out). A further move from Timiş to Bucharest-Ilfov (denoted

Figure 3.3: Average effect

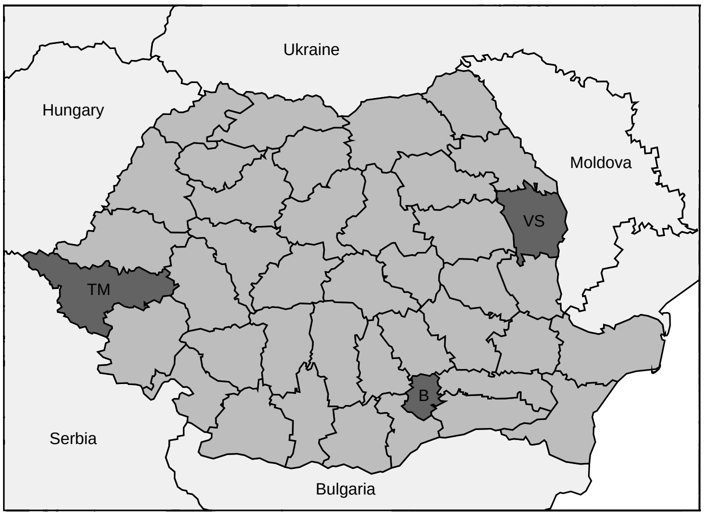
(a) Horizontal Spillovers **(b)** Backward Spillovers



The figure shows the actual contribution of the FDI on the productivity of domestic firms, where each coefficient is augmented by the amount of foreign presence at the respective regional dimension.

as B) increases the total spillover effect by 0.058. In contrast to the previous move, the increase is now due to the combination of a decrease in the horizontal effect (-0.032) that is more than compensated by a larger increase in the backward spillover effect (0.091). The latter is due to the fact that being in the same region as foreign entrants carries a statistically significant bonus in terms of backward spillover effects, as well as that Bucharest-Ilfov is the main TFP-hub in Romania, dwarfing the other regions in terms of foreign presence and foreign entry over the sample period. This also explains a within-region negative horizontal effect in Bucharest-Ilfov, compared to Timiş. In the latter the horizontal effects from the many foreign firms located in Bucharest-Ilfov have a more benign effect as they are part of the rest-of-country component for firms located there. The regions recording the smallest total spillover effects over the sample period are neighbours of Bucharest-Ilfov. This occurs due to the fact that they are less protected from negative horizontal effects from foreign firms in Bucharest-Ilfov as well as that they are not close enough in order to benefit from the immediate positive backward spillover effects generated by the large number of foreign entrants in Bucharest-Ilfov. Note that all this is derived on the basis of point estimates and should be considered as indicating the direction of the effects. F-tests revealed that only the 'immediate' within-region backward spillover effect (the largest source of spillover effects) is found to be statistically different from the other geographical components.

Figure 3.4: The NUTS3 regions in Romania



In Table 3.8 we further explore regional heterogeneity in the estimation by testing whether the identified patterns are stable across regions that perform above and below median region-productivity. We use the approach of Foster et al. (2001) to calculate initial regional TFP from our firm-level data.[10] Regions with above median region-TFP levels could be interpreted as more dynamic regions with larger absorptive capability, yielding a rationale to expect different patterns. As Table 3.8 indicates, more productive regions show slightly higher horizontal spillover coefficients compared to the entire sample. This might suggest that firms located in these regions are on average better at adapting to foreign presence in their industry. Coefficients are similar at all regional dimensions, indicating that location is not relevant for horizontal spillover absorption. With respect to backward spillover effects, we confirm the 'being-close bonus', i.e. the significantly larger within-region backward spillover effect from recent foreign entrants. Overall, these patterns are fairly similar to those obtained using the full sample. Differences emerge when we consider spillover patterns in below median region-productivity regions. Within-region positive horizontal spillover effects from foreign firms with

[10]We compute this value by a weighted sum of their individual productivities: $P_r = \sum s_{ir} * p_i$ and where s_{ir} is the regional output share of firm i in region r and p_i is its productivity.

sufficient maturity are no longer detected (whereas they still are positive and significant at the neighbour and rest-of-country levels). Similarly, for backward spillover effects, the within-region 'being-close bonus' for new foreign entrants disappears, as do the positive within-region backward spillover effects from foreign firms entering two and three years earlier. It also takes more time for other positive backward spillover effects to manifest themselves in these regions (3-4 years).

Table 3.8: Horizontal and Backward spillovers in Above and Below median productivity regions

	(1) Above		(2) Below	
	(a) Horizontal	(b) Backwards	(a) Horizontal	(b) Backwards
entry in t				
same region	-0.504	17.622**	0.869	49.737
	[1.130]	[7.697]	[0.994]	[47.079]
neighbouring region	-0.937	-3.246	1.799***	17.826
	[0.849]	[5.905]	[0.520]	[11.645]
rest of country	0.571	2.256	-0.584	4.253
	[0.404]	[2.204]	[0.596]	[2.708]
entry in t-1				
same region	-2.266**	11.672***	-4.238	46.073*
	[0.899]	[3.213]	[5.869]	[26.034]
neighbouring region	-1.659**	3.812	-3.358***	8.437
	[0.705]	[4.240]	[1.090]	[8.256]
rest of country	-0.852*	8.711***	-1.247***	8.071***
	[0.498]	[1.319]	[0.446]	[1.883]
entry in t-2				
same region	-1.316**	3.580***	-10.836**	14.002
	[0.582]	[1.189]	[4.726]	[15.817]
neighbouring region	-2.505***	2.036	-3.253***	19.130***
	[0.705]	[3.139]	[0.889]	[5.051]
rest of country	-1.475***	3.762***	-1.157***	4.963***
	[0.342]	[0.961]	[0.389]	[1.039]
entry in t-3				
same region	0.408	3.810***	-4.773	21.948
	[0.450]	[1.383]	[3.559]	[24.716]
neighbouring region	0.161	1.019	0.089	8.415*
	[0.643]	[2.815]	[0.845]	[4.590]
rest of country	0.766***	3.682***	0.404	5.775***
	[0.253]	[1.283]	[0.277]	[1.240]
entry in t-4 or earlier				
same region	1.713***	0.264	1.651	10.422
	[0.413]	[1.246]	[1.473]	[9.899]
neighbouring region	2.135***	-1.01	1.585***	-2.831
	[0.398]	[1.208]	[0.482]	[2.175]
rest of country	2.301***	-0.384	1.999***	-0.4
	[0.204]	[0.571]	[0.234]	[0.656]
Observations	33,693		15,381	
R-squared	0.075		0.072	

Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3.9: Test for equality of coefficients - Above/Below

F-test	(1) Above		(2) Below	
	(a) Horizontal	(b) Backward	(a) Horizontal	(b) Backward
Reg_t=NB_t=RoC_t	1.507	2.304	4.854***	1.094
Reg_t=NB_t	0.102	4.402**	0.867	0.441
NB_t=RoC_t	2.687	0.814	9.637***	1.340
Reg_t=RoC_t	0.797	3.745*	1.678	0.933
Reg_{t-1}=NB_{t-1}=RoC_{t-1}	1.174	1.160	1.878	1.062
Reg_{t-2}=NB_{t-2}=RoC_{t-2}	1.162	0.146	4.330**	4.168**
Reg_{t-2}=NB_{t-2}	1.911	0.210	2.487	0.091
NB_{t-2}=RoC_{t-2}	2.059	0.289	5.023**	7.752***
Reg_{t-2}=RoC_{t-2}	0.067	0.014	4.118**	0.325
Reg_{t-3}=NB_{t-3}=RoC_{t-3}	0.622	0.452	1.107	0.342
Reg_{t-4+}=NB_{t-4+}=RoC_{t-4+}	1.004	0.318	0.380	1.215

The table presents a test for the equality of coefficients at regional, neighbour and rest-of-country level with the corresponding F-test value. * rejected at 10%; ** rejected at 5%; *** rejected at 1%.

3.5.3 Foreign firms' location choice

In Table 3.8 we found that regions above and below the median region-productivity show different spillover patterns. One could argue that foreign firms would tend to locate in the regions where they expect domestic firms with higher productivity (growth) to be located. In order to make sure that our results are not driven by such factors, we analyze foreign firms' location choice within Romania. From panel (a) in Figure 3.5 one can observe that the majority of foreign companies locates either near the Western border with Hungary or in Bucharest-Ilfov, the capital region. This indicates that location choice is potentially non-random.

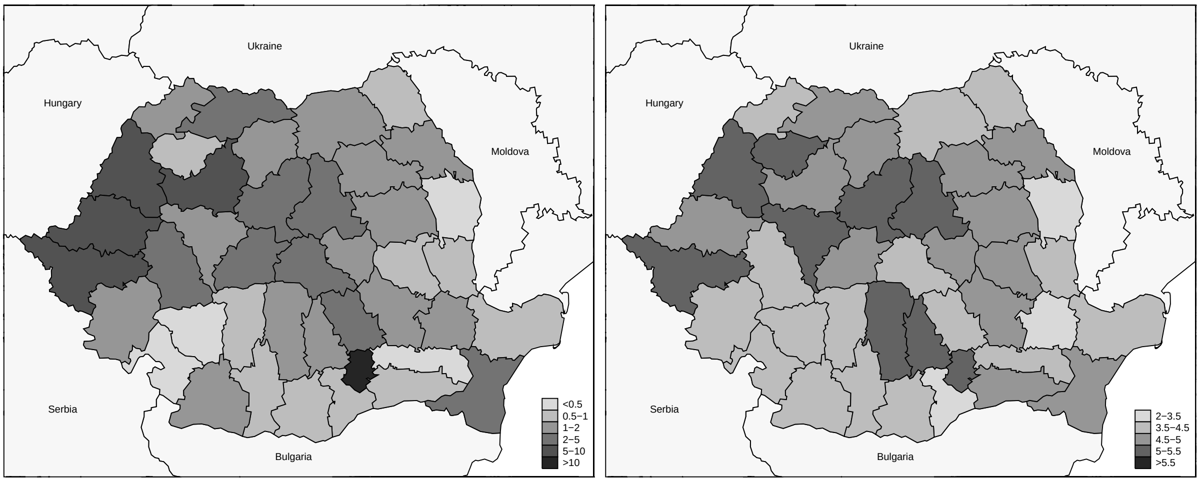
Location choice can be explained by several factors. First of all, it has been suggested that foreign companies investing in developing countries such as Romania face very specific obstacles like widespread bureaucracies, corruption, insufficiently developed financial markets and unpredictable legal systems (see Bitzenis (2006)). Therefore, instead of focusing solely on labour costs, foreign firms would locate in areas with high services agglomeration or, in other words, large cities which allow them to have access to lawyers, accountants, translators and the banking industry (Hilber and Voicu (2010)). Second, location of foreign subsidiaries might also be explained by the proximity to Western borders. Since a large share of foreign investment has European roots, choosing a location closer to home might constitute an advantage for their parents. Moreover, since Western border regions have for a large part of recent history been under the influence of the Habsburg empire, locating in this area might be more appealing from a cultural sense as well (see Becker et al. (2011)).

Nonetheless, there might still be an issue if the most productive (domestic) companies are also located in these regions. Comparing panels (a) and (b) of Figure 3.5 suggests no immediate problem. Nevertheless, we run two simple regressions to investigate how regional productivity growth is related to the location choice of new foreign firms. We perform the analysis at both the region and the region-industry level. We include a Western border dummy because we expect the border to have a significant impact on location choice due to the closeness to Western markets. Further we add a dummy that is set to one if the main national road connecting Bucharest with Hungary passes through the region.[11] Finally, we also include the regional rural rate obtained from the Romanian National Statistical Office (RNSO) (2014) as we expect foreign firms to be located in urban areas, giving them access to services and higher educated labour force. The results in Table 3.10 indicate that location is indeed heavily influenced by our control variables but is not related to the lagged first difference of regional productivity of domestic firms. We therefore conclude that foreign firm location is not influenced by the presence of fast growing domestic firms in the region.

Figure 3.5: Share of foreign firms and regional productivity

(a) 2005 share of foreign firms as percentage of total country number of foreign firms

(b) 2005 regional productivity of domestic firms computed as a weighted sum of their individual productivities



[11]Other roads were underdeveloped and of poor quality during our sample period.

Table 3.10: Location of foreign firms

	New foreign firms (region)	New foreign firms (region industry)
Regional productivity growth	0.021	-0.006
	[0.576]	[0.006]
DN1 road	11.146***	0.220***
	[2.550]	[0.048]
HU border	10.564***	0.313***
	[2.424]	[0.057]
Rural rate	-52.326***	-1.740***
	[11.965]	[0.194]
Observations	369	6,293
R-squared	0.356	0.061

The table shows the regression results of local firm productivity change on foreign firm location. Standard errors are reported in brackets. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent. The dependent variables are the lag of the first difference in regional/region-industry firm productivity of domestic firms, a dummy indicating whether the main national road is crossing the region, a dummy for bordering regions with Hungary and the regional rural rate.

3.6 Conclusion

This chapter analyses horizontal and vertical productivity spillover effects of foreign direct investment on domestic Romanian manufacturing companies from 1996 to 2005. We add to the literature by analysing the pattern of technology diffusion as a function of both distance (proximity) and time-since-foreign-entry as mediating factors of spillover effects.

Horizontal spillovers are found to be fairly homogeneous over distance. Recent foreign entrants have a negative impact on local competitors' TFP which is more than compensated by positive effects once foreign firms have been present for a longer period in the domestic economy. This indicates that it takes time for domestic firms to adjust to foreign presence, with productivity improvements being realised provided these companies withstand the initial pressures from foreign entrants and are able to absorb the new foreign technology. This finding does not differ between above and below median productivity regions.

In terms of backward spillovers, our results indicate that these manifest themselves relatively fast after foreign entry, but fade away when foreign firms have been present for a longer period. Being located in the same region as foreign firms carries a rapidly arriving productivity bonus compared to foreign firms located further away. This suggests that over larger distances spillovers are absorbed, but at a slower pace. The effect is particularly strong for domestic firms located in regions with above median productivity. For below median productivity regions, positive backward spillover effects mainly originate from further away regions and the being close bonus disappears.

These patterns are in line with what one would expect for a developing economy. Horizontal spillovers mainly manifest themselves via competition and labour channels which are only affected by distance in a limited way. In the time-since-foreign-entry dimension these effects are non-linear, with a negative competition and labour poaching effects at first, followed by positive spillovers once domestic firms have adjusted to the increased competition and workers trained by foreign firms start to switch back to domestic firms. Because MNEs have an incentive to assist their local suppliers and learning is faster at closer distances due to easier (face-to-face) contact and visits of engineers, we find a bonus of being located close to a foreign client. The latter does not preclude effects over longer distances, but they are smaller and take more time to manifest themselves. Irrespective of distance, backward spillover effects are limited in time-since-foreign-entry. Once a domestic supplier has reached a sufficient level of TFP, a linkage with a foreign client no longer stimulates technological upgrading. We do not find evidence of forward spillover effects which is in line with other recent work on Eastern Europe.

Overall spillover effects from foreign direct investment are likely to be positive, but both horizontal and backward spillover effects vary considerably with foreign firms' maturity. Backward spillover effects are also faster and larger for nearby domestic firms. Horizontal spillover effects vary with distance on the basis of point estimates (the closer, the smaller the spillover effect because of a negative competition aspect), but these differences are not statistically significant.

4

FDI and Borders: Evidence of Knowledge Spillovers from Neighbouring Countries *

4.1 Introduction

Together with enterprise creation, encouraging inward foreign direct investment (FDI) is one of the cornerstones of most industrial policy in both developing and developed countries. Policymakers do so because FDI is considered to be an important channel via which growth can be boosted. Multinational firms (MNEs) are not only expected to bring new capital, resources and jobs, but they are also expected to generate positive productivity spillover effects towards domestic firms. It is generally agreed that MNEs are more productive than their domestic counterparts and posses superior technologies and knowledge (see Helpman et al., 2004). Given that it is impossible for MNEs to fully internalise these advantages, these technologies and know-how are expected to spill over to domestic firms, resulting in productivity gains for the latter. This paper brings together three strands of literature to investigate the impact of

*This chapter is based on a paper co-authored with Bruno Merlevede.

national borders in productivity spillover effects from FDI. Our paper is most strongly related to the micro-econometric literature on productivity spillover effects from foreign to domestic firms. Further it draws on the macro-oriented literature on technology transfer across countries and the literature on border effects in trade.

Research into the existence and economic significance of these spillover effects has provided mixed results, however (see e.g. Görg and Greenaway, 2004; Crespo and Fontoura, 2007). This was especially the case for the earlier literature that focused on spillover effects from MNEs on domestic firms that operate in the same industry (horizontal spillover effects). A more consistent picture emerges when one looks at vertical spillover effects, i.e. those arising between a foreign firm and its domestic suppliers (and to a lesser extent buyers). First established by Javorcik (2004), these findings suggest that MNEs have a greater interest in sharing their advanced knowledge, be it either in terms of operational or management techniques, with domestic firms that are related to the foreign firm through the supply chain. Following Javorcik (2004), a large set of papers have replicated this finding for other countries, with consistently positive and economically significant backward spillover effects and small, but often statistically significant forward spillovers (see Havránek and Irŝová, 2011).

Despite the great breath of papers in this field, the literature almost exclusively consists of single country studies. This is due to the fact that large, representative multi-country firm-level datasets are scarce. The literature has therefore turned attention more towards characteristics that potentially facilitate the occurrence of (positive) spillover effects. Relevant characteristics include the absorptive capability of the domestic firms (see e.g. Narula and Marin, 2005), the characteristics of foreign affiliates such as technological capability, embeddedness, and autonomy (e.g. Giroud et al., 2012; Marin and Bell, 2006; Lenaerts and Merlevede, 2015), as well as the distance between foreign and domestic firms (Merlevede and Purice, 2015).

In this paper we exactly draw upon the large, representative multi-country firm-level dataset for European countries constructed in Chapter 2. This allows us to analyse the potential for cross-border spillover effects. Using country-level data, Keller (2002) finds that technology transfer declines and in the end disappears with distance. Comin et al. (2012) confirm this finding and show that distance has a negative impact on technology diffusion, although they also find that the impact of distance dies out over time. These results clearly point to a role for distance. In our multi-country firm-level setting we analyse the impact of distance in a discrete way and investigate whether national borders constitute a barrier for spillover effects from MNEs to local firms. This further relates to the literature that investigates border effects in trade. McCallum (1995) finds that regions within the US trade about 20 times more than US and Canadian regions over a similar distance. By means of a meta-analysis, Havranek and Irsova (2015) confirm that within-country regions trade about 20% more than regions with similar characteristics, but located in different countries. Clearly, border effects in trade are very relevant to vertical spillover effects since they are explicitly related to the supply chain.

Our data allows us to analyse the potential for cross-border spillover effects for a large number of domestic firms in Eastern European countries that have joined the EU in 2004 and 2007. For an economic union such as the EU, where borders should constitute less of a barrier for the movement of capital and goods, it is not unlikely that domestic Eastern European firms near the border with Western Europe are interacting with firms in neighbouring countries. Especially client-supplier relations could be established across borders. We further investigate whether some heterogeneity in borders exists by using bilateral Schengen membership as a further test of the strictness of borders. This question also parallels part of the trade literature that has been concerned with the impact of (heterogeneity in) borders on the trade between and within countries. In line with this border effect literature, we analyse whether (cross-border) FDI spillover effects can be detected in Eastern Europe and whether there are significant differences between within country and cross-border spillover effects for similar distances between foreign and domestic firms.

We believe this to be an important question for several reasons. First of all, testing for such spillover effects can be interpreted as a test of the integration in the internal EU market. Are borders still a too strong barrier for domestic firms located near the border to be able to benefit from the MNE activity in the neighbouring region across the border? Second, the existence of spillover effects from MNEs located over the border could suggest potential benefits of cooperation between neighbouring countries when trying to attract FDI in their regions.

We perform the analysis using firm-level data on Central and Eastern Europe during the 2000-2010 period. We use information on the region where firms are located, and combine this with the distance between regions to analyse whether domestic firms within a certain distance from the border experience productivity gains as a result of MNE activity nearby, but across the border. Our findings suggest that national borders constitute an almost insurmountable barrier for horizontal and forward spillover effects. In the case of backward spillover, national borders significantly dampen cross-border spillover effects, but the size of the impact of the border seems related to the '*depth*' of the border as evidenced by the difference between Schengen and non-Schengen borders.

The remainder of the paper is structured as follows. In section 2 we give a brief overview of our methodology. Section 3 introduces our data and explains how we constructed the dataset. Section 4 presents the results and we conclude in section 5.

4.2 Spillovers and border effects

In order to identify the presence of cross-border spillover effects, we first introduce the traditional country-wide approach to the analysis of spillover effects. We then proceed by introducing cross-border spillover effects in the standard framework for the analysis of productivity spillover effects from FDI.

4.2.1 Country wide spillovers in the standard framework

The existing literature on country-wide spillover effects from foreign direct investment makes and important distinction between within industry, or horizontal spillover effects, and between industry, or vertical spillover effects. The main channels for horizontal spillovers are technology imitation (the demonstration effect, (see Teece, 1977) and mobility of workers trained by foreign firms (see Görg and Strobl, 2005). Foreign entry may also lead to increased competition in the domestic market. This could form an incentive for host-country firms to use existing technologies and resources more efficiently or adopt new technologies, which provides another important channel of horizontal spillover effects (see e.g. Aitken and Harrison, 1999). These effects are not however necessarily positive. Backward spillovers run from a foreign firm to its upstream local suppliers. In this case foreign firms are more likely to assist their local suppliers in providing inputs of sufficient quality to realise the full benefits of their investment. A forward spillover goes from a foreign firm to its downstream local buyer of inputs. These inputs may be of better quality and enhance the productivity of local firms that use them. However, these inputs may also be more expensive and less adapted to local firm requirements, potentially denting local firm productivity. The existing literature on country-wide spillover effects from foreign direct investment makes and important distinction between within industry, or horizontal spillover effects, and between industry, or vertical spillover effects. The main channels for horizontal spillovers are technology imitation (the demonstration effect, (see Teece, 1977) and mobility of workers trained by foreign firms (see Görg and Strobl, 2005). Foreign entry may also lead to increased competition in the domestic market. This could form an incentive for host-country firms to use existing technologies and resources more efficiently or adopt new technologies, which provides another important channel of horizontal spillover effects (see e.g. Aitken and Harrison, 1999). These effects are however necessarily positive. Backward spillovers run from a foreign firm to its upstream local suppliers. In this case foreign firms are more likely to assist their local suppliers in providing inputs of sufficient quality to realise the full benefits of their investment. A forward spillover goes from a foreign firm to its downstream local buyer of inputs. These inputs may be of better quality and enhance the productivity of local firms that use these inputs, but these inputs may also be more expensive and less adapted to local firm requirements, so they dent local firm productivity.

The standard methodology for identifying these spillovers effects can be viewed as an 'augmented' production function approach. In this approach variables capturing foreign presence are added to more standard explanatory variables such as labour, capital, and material inputs in explaining total factor productivity. The standard measure for the (country-wide) horizontal spillover variable is given in Equation (4.1). For a (domestic) firm i in industry j at time t it is of the following form:

$$HR_{jt}^{CW} = \frac{\sum\limits_{i \in j} F_{it} Y_{it}}{\sum\limits_{i \in j} Y_{it}} \tag{4.1}$$

where Y is output and F is the percentage of shares owned by foreign investors. (4.1) is the standard used throughout the literature. However, there are other ways to measure the FDI intensity within a country. Altomonte and Pennings (2009) for example use a simple count of foreign firms. However, we prefer to use the share of output produced by foreign firms, as this takes into account the relative size of foreign activity. We denote as foreign any firm with a non-domestic direct ownership of more than 50%. HR_{jt}^{CW} in (4.1) thus measures the share of output that is produced by foreign firms in industry j in the country at time t. The spillover variable is then built up to industry level from firm-level data, such that HR_{jt}^{CW} has the same value for all firms i in industry j at time t.

The backward spillover variable, BK_{jt}^{CW}, is computed in (4.2) following an approach that is by now standard in the literature:

$$BK_{jt}^{CW} = \sum_{k \neq j} \gamma_{jk} * HR_{kt}^{CW} \tag{4.2}$$

here γ_{jk} is the proportion of industry j's output supplied to industry k. Typically the γ's are calculated using country-level input-output (I-O) tables for intermediate consumption. However, since we are looking at multiple countries and will also account for cross border spillovers, we make use of EU-27 I-O tables.[1] These tables are not available on an annual basis. We use the 2007-table provided by Eurostat. The backward spillover variable is thus an industry-level measure that captures the potential for a domestic firm in industry j to supply a foreign firm in downstream industry k.

Similarly to the backward spillover variable, we define the forward spillover variable as:

$$FW_{jt}^{CW} = \sum_{l \neq j} \delta_{jl} * HR_{lt}^{CW} \tag{4.3}$$

where δ_{jl} is the proportion of industry j's inputs sourced from industry l. The forward spillover variable then captures the potential for a domestic firm in industry j to buy its inputs from a foreign firm in upstream industry l.

Once computed, the spillover variables HR_{jt}^{CW}, BK_{jt}^{CW}, and FW_{jt}^{CW} are then added as explanatory variables in an equation estimating the productivity of (domestic) firm i

[1] An alternative to using one I-O table is applying country-specific tables for spillovers within a country and country pairwise tables for spillovers across borders, as allowed by the WIOD database. However, WIOD is computed at a more aggregate level, which could significantly underestimate the impact of backward spillovers (see Lenaerts and Merlevede, 2012).

in industry j ($cf.\ infra$). As common in the literature, the size, sign and significance of the resulting coefficients are taken as evidence of spillover effects. The country wide spillovers are then:

$$\Psi_1 f\mathbf{FDI}^{CW}_{jt-1} = \alpha_{hr,fc} HR^{CW}_{jt-1} + \alpha_{bk,fc} BK^{CW}_{jt-1} + \alpha_{fw,fc} FW^{CW}_{jt-1} \tag{4.4}$$

4.2.2 Accounting for borders

Starting with the seminal paper by McCallum (1995), the trade literature has been concerned with the impact of borders. McCallum (1995) was the first to compare how the movement of goods between two regions within a country (the US) compare to that between regions in different countries (the US and Canada). He found that the US-Canada border constitutes very a strong barrier to trade. Ever since a multitude of papers have tried to measure this border effect, with results suggesting that within-country regions trade around 20% more than regions from different countries with similar characteristics (see Havranek and Irsova, 2015). Results do however differ depending on the methodology used to determine internal distance within countries (see Head and Mayer, 2002). Moreover, estimates differ depending on the industries as well as the countries considered, with stronger border effects in more regional markets of perishable goods (see Chen, 2004) as well as in developing world compared to the OECD countries (see Havranek and Irsova, 2015). Taking inspiration from this rich border effect literature, we analyse whether cross-border FDI spillover effects are present in the EU or whether borders are too strong a barrier for firms to be able to benefit from MNE activity in the neighbouring countries and regions.

In addition to the country-wide spillover effects in (4.4), we therefore introduce cross-border spillover effects (horizontal and vertical) stemming from FDI in neighbouring countries. We start by determining the location of firms in our dataset up to the relative detailed NUTS 3-digit regional level.[2] For firms located in a region close to the border we define an 'area of interest' (AI) and consider those with a national border present within this area. Figure 4.1 illustrates this procedure for a domestic firm located in the

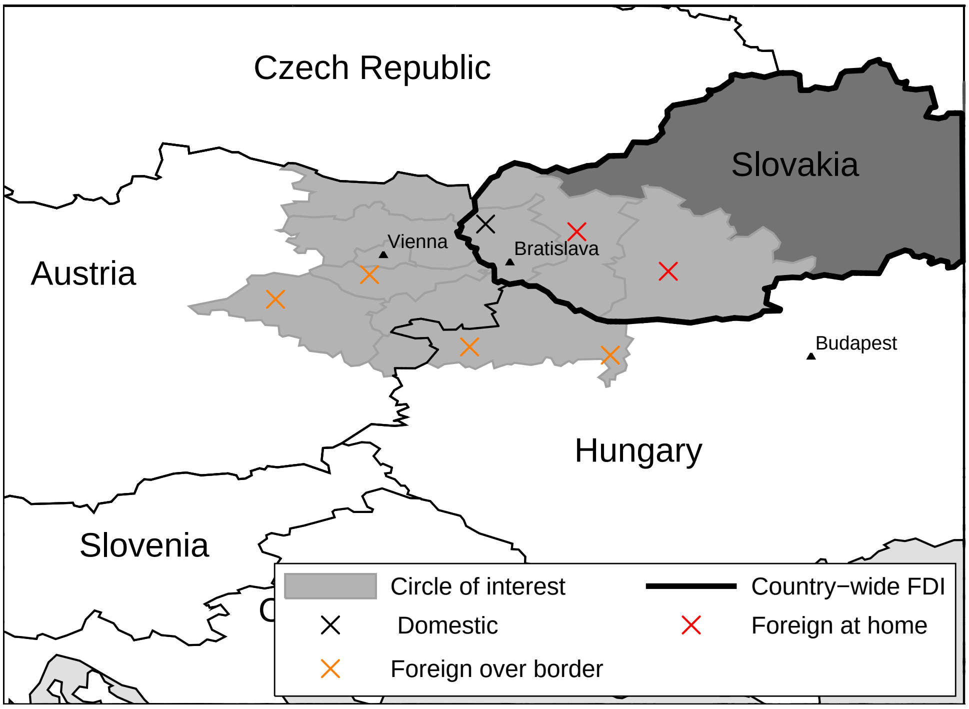
[2]Since we do not posses all the exact addresses where firms are located, this is the best measure we have to determine how far from each other foreign and domestic firms are located. For a reasonable amount of firms, the data only allow to assign our firms to the centre of their region because of missing zip-codes.

Bratislava NUTS 3-digit region in Slovakia. The light grey area is the area of interest for a firm in the Bratislava region. It encompasses all NUTS 3-digit regions whose centre is within 75 kilometres of the centre of the Bratislava region.[3] This light grey area forms the basis for defining variables to analyse cross-border spillover effects. We prefer to consider home country regions that have a neighbouring cross-border regions within 75 kilometres region centre distance over pure neighbour regions because NUTS 3-digit regions differ in size across countries. Figure 4.1 shows that for a firm in Bratislava this area of interest includes both national NUTS 3-digit regions as well as regions in two neighbour countries, Austria and Hungary. A Bratislavan firm thus would potentially face both spillovers from within Slovakia itself as well as spillovers from its two neighbours. Note that while the Czech Republic is a neighbour of Slovakia as well, the distance between Bratislava and the nearest NUTS 3-digit Czech region centre is above 75 km, therefore falling outside our 'area of interest'. Bearing in mind that there is some consensus on the backward spillover channel as the main source of substantial positive spillover effect, the choice for an area of interest of limited distance can be motivated by recent research that points towards the role of proximity in the formation of linkages. Using a Japanese transaction dataset, Bernard et al. (2014) show the importance of distance for the formation of linkages. Key to the backward spillover channel is that MNEs have an interest in technological upgrading by their suppliers and therefore an incentive to provide them with explicit assistance (see Javorcik, 2004). As successful assistance and upgrading requires human interaction, communication, and monitoring (see Giroud, 2012; Keller and Yeaple, 2013), nearby suppliers are the most likely place to detect positive spillover effects. Keller and Yeaple (2013) further show that more knowledge-intensive inputs call for more communication, thereby favouring nearby locations. Therefore, a distance-limited area of interest including national borders is the best setting to detect border effects in productivity spillover effects from MNEs to local firms.

Our empirical strategy amounts to defining spillover variables for firms in 'border' regions that allow to identify whether spillover effects from foreign firms within the area of interest -both within the home country and across the border- are different from the traditional country-wide spillovers defined above. We build on (4.1) and define

[3]We consider other distances as robustness check.

Figure 4.1: The spillovers region for a firm in Bratislava



The figure shows how we calculate the FDI spillover potential for a domestic firm located in Bratislava, Slovakia.

spillover variables within the area of interest (AI) as follows:

$$\begin{aligned} HR_{jt}^{AI} &= \frac{\sum\limits_{i \in j, i \in AI} F_{it} Y_{it}}{\sum\limits_{i \in j, i \in AI} Y_{it}} \\ &= \frac{\sum\limits_{i \in j, i \in AI-H} F_{it} Y_{it} + \sum\limits_{i \in j, i \in AI-CB} F_{it} Y_{it}}{\sum\limits_{i \in j, i \in AI} Y_{it}} \\ &= HR_{jt}^{AI-H} + HR_{jt}^{AI-B} \end{aligned} \tag{4.5}$$

where Y is output and F is the percentage of shares owned by foreign investors as before. The difference with (4.1) is that HR^{AI} is limited to firms located in the area of interest and includes regions in the neighbouring country. Further HR^{AI} can be straightforwardly split in foreign activity in the home country (HR^{AI-H}) and foreign activity across the border (HR^{B}). This allows us to separate the impact of both types of foreign activity in the estimation. Differences in the estimated impact then make inference about the impact of the border possible. In the calculation of foreign presence in (4.5) we exclude home country firms with foreign activity across the border. We do so due to the fact that although technically foreign in the neighbouring country, there might be endogeneity concerns. Domestic firms over the border might be there due to higher productivity at home to begin with, thus crossing the border due to this advantage in the first place.

Backward spillover variables can then be calculated as before, again making us of the I-O tables as in (4.2) (forward spillover variables are obtained in a similar way):

$$\mathrm{BK}_{jt}^{AI-H} = \sum_{k \neq j} \gamma_{jk} * HR_{kt}^{AI-H} \tag{4.6}$$

$$\mathrm{BK}_{jt}^{AI-CB} = \sum_{k \neq j} \gamma_{jk} * HR_{kt}^{AI-CB} \tag{4.7}$$

The additional spillover variables for firms in regions close to the border can then be grouped as:

$$\begin{aligned} \Psi_2 f \mathbf{FDI}_{jt-1}^{AI} = {} & \alpha_{hr,fb} HR_{jt-1}^{AI-H} + \alpha_{hr,fh} HR_{jt-1}^{AI-CB} \\ & + \alpha_{bk,fb} BK_{jt-1}^{AI-H} + \alpha_{bk,fh} BK_{jt-1}^{AI-CB} \\ & + \alpha_{fw,fb} FW_{jt-1}^{AI-H} + \alpha_{fw,fh} FW_{jt-1}^{AI-CB} \end{aligned} \tag{4.8}$$

4.2.3 Empirical framework

As indicated above, FDI spillovers are typically introduced as additional inputs explaining total factor productivity (TFP) in a production function framework. In this paper, we rely on the standard framework and follow what Havránek and Irŝová (2011) describe as 'best practice'. In particular, we consider a two-step procedure where we use firm-level data to estimate a production function in order to obtain a TFP-measure in the first step, and then relate TFP-growth to foreign presence and additional control variables in a second step.

In the estimation of total factor productivity (TFP) one is confronted with an endogeneity problem due to the fact that firms and observe their productivity (shocks) and adjust their input choices accordingly. To account for this problem, several semi-parametric techniques have been suggested, among which those of Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP). A recent modification of the LP methodology was suggested by Wooldridge (2009) (WLP), which combines the benefits of OP and LP, whilst applying a joint GMM estimation which both enhances efficiency and accounts for serial correlation and heteroskedasticity. We therefore use the WLP estimator to obtain our indicator of total factor productivity (TFP). A measure of TFP for firm i in industry j at time t is obtained as the difference between output and capital, labour, and material inputs, multiplied by their estimated coefficients (lower cases refer to logs):

$$tfp_{ijt} = y_{ijt} - \widehat{\beta}_{lj} l_{ijt} - \widehat{\beta}_{kj} k_{ijt} - \widehat{\beta}_{mj} m_{ijt} \tag{4.9}$$

In a second step we regress the change in productivity of domestic firms on the spillovers variables defined above and a set of control variables. We then estimate the following specification using the sample of all domestic firms:

$$\Delta tfp_{ijrt} = \Psi_1' \Delta f\left(\mathbf{FDI}_{jt-1}^{CW}\right) + \Psi_2' \Delta f\left(\mathbf{FDI}_{jt-1}^{AI}\right) + \Psi_3' \mathbf{Z}_{ijt} + \alpha_t + \alpha_j + \alpha_r + \epsilon_{ijrt} \tag{4.10}$$

where Ψ_1 (see (4.4)) refers to the impact of traditional country-wide foreign presence on the productivity of local firms and Ψ_2 (see (4.8)) identifies the additional impact of foreign firms from within the area of interest both at home and across the border. $\mathbf{Z}_{ijt}$ stands for firm specific control variables such as age, size. We further introduce time (α_t), industry (α_j) and region (α_r) dummies in the first-differenced specification (4.10) to account for unobserved factors that could be driving the change in performance at the region or industry level. This is relevant because we consider a sample of Eastern European countries where specific features of their transition from plan to market may still have some impact in certain regions or industries. We estimate specification (4.10) by OLS. Finally, since $\mathbf{FDI}_{jt}$ and some control variables are defined at the industry level and estimations are performed at the firm level, standard errors are clustered for all observations in the same region, industry and year (see Javorcik, 2004; Moulton, 1990).

4.2.4 Alternative estimation techniques

Since we allow distance to determine whether a domestic firm will be impacted by its foreign counterparts, this paper has a clear spacial element. Therefore, an alternative estimations for our analysis might be the spacial econometric technique. In essence, we already use a specification in the spirit of the spatial lag of X model as suggested by LeSage (2014), where:

$$y = X\beta_1 + WX\beta_2 + \epsilon \tag{4.11}$$

with y being the change in domestic TFP productivity, X representing the share of FDI in the region or across the border and W is an a adjacency matrix with values of zero and one, depending on whether a foreign firm is located within our distance of interest or not.[4]

However, we do not use the more famous Spatial Durbin Model (SDM), which is a global spillover specification and includes a lag vector WY, where Y is the dependent variable from the neighbouring regions. In our case this would imply that local firms across the border have an impact on the productivity of domestic firms as well. We believe that this is implausible, as it is the multinationals that are more mobile and more productive in the first place. Moreover, recent papers have expressed concern about the use of spacial econometrics without proper theoretical justification and suggested some remedies for researchers. For example, Partridge et al. (2012) recommend using a Moran-I test which would indicate whether we are dealing with spacial dependence and thus whether including a spacial lag WY is warranted in the first place.

Following this advise, we perform a Moran-I test checking for spacial autocorrelation between the growth of domestic firms at home and over the border. Given that we are dealing with panel firm level data for several industries, we first compute the regional unweighted productivity growth of domestic firms in manufacturing industries. The choice of aggregating firm productivities is due to the fact that very few industries are present in all regions. We then test for spacial autocorrelation by year. The table with yearly Moran-I statistics and respective p-values is given below.

As Table 4.1 suggests, except for 2009, we can not reject the null hypothesis that

[4] We have also tried a specification where the weights are determined purely by physical adjacency, with results similar the ones in the main specification. Results are available on request.

Table 4.1: Moran I test for spacial autocorrelation

Variable	Year	I	E(I)	sd(I)	z	p-value*
Average productivity growth	2001	0.046	-0.001	0.069	0.687	0.246
	2002	-0.067	-0.001	0.068	-0.972	0.165
	2003	0.043	-0.001	0.057	0.778	0.218
	2004	-0.042	-0.001	0.051	-0.798	0.213
	2005	-0.056	-0.001	0.069	-0.805	0.210
	2006	0.004	-0.001	0.068	0.077	0.469
	2007	0.029	-0.001	0.064	0.470	0.319
	2008	-0.023	-0.001	0.069	-0.316	0.376
	2009	0.106	-0.001	0.069	1.559	0.060
	2010	0.060	-0.001	0.069	0.879	0.190

The table shows the results of a 1-tailed Moran test for spacial autocorrelation. The variable tested is the average productivity growth of domestic manufacturing firms in bordering regions.

there is zero spatial autocorrelation present in the productivity growth of domestic firms at home compared to those across the border. We therefore conclude that spacial correlation is not an issue and that not accounting for the spatial lag WY is appropriate for our analysis.

4.3 Data

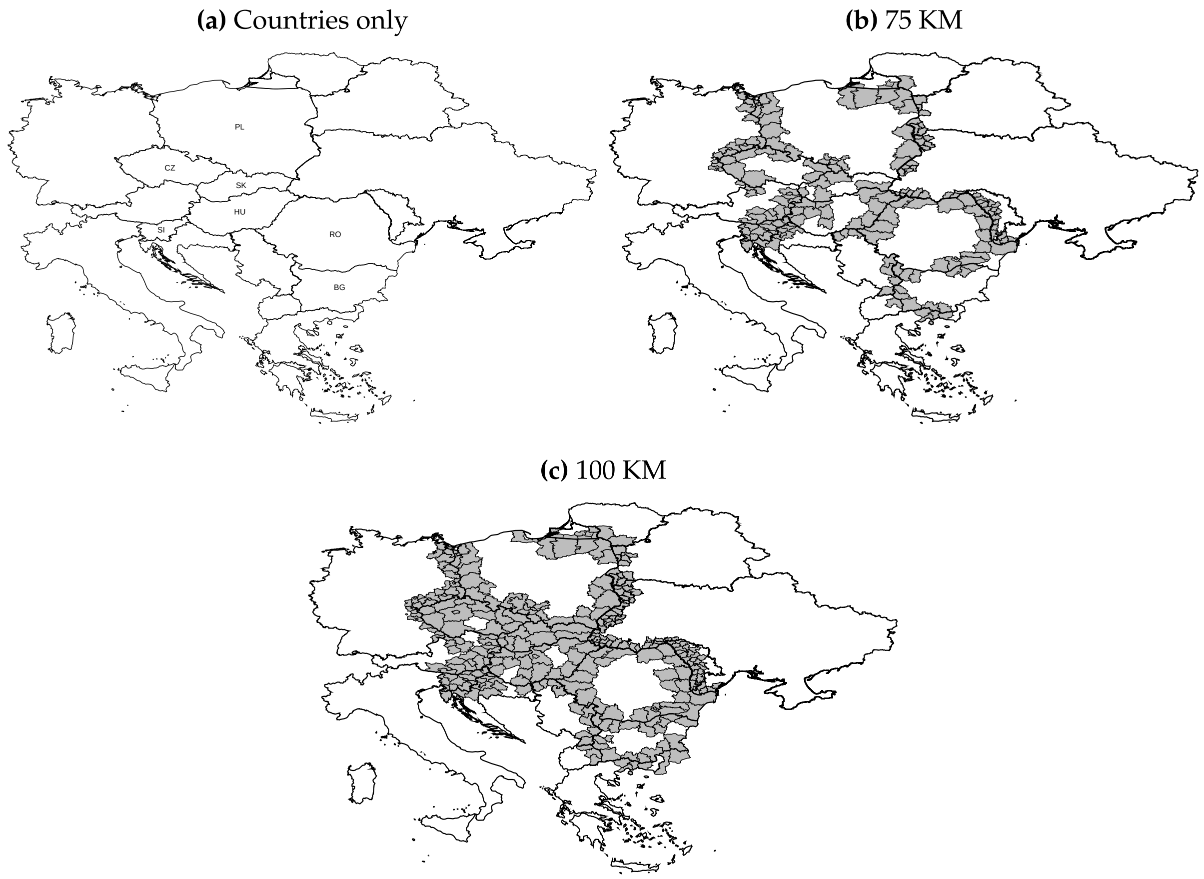
Our basic data source is the *AUGAMA*, a database consisting of financial and ownership information on public and private companies across Europe as described in Chapter 2. From this large database, we constructed a sample covering the period 2000-2010 that allows us to study FDI spillover effects on domestic manufacturing firms in seven Central and Eastern European countries (CEEC).[5] For a detailed account of how the data was constructed and cleaned, we refer to Chapter 2.[6] We choose to focus on these seven countries due to ample evidence that CEEC firms have benefited from the entry of foreign companies from Western Europe (see e.g. Damijan et al., 2013). During this period, our data records more than 200,000 domestic firms. We limit the sample to firms with at least 10 employees and an unconsolidated account. We consider a further eleven countries that border our focal seven CEEC countries.[7]

[5] Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovakia, and Slovenia.

[6] We use multiple issues (published on DVDs) of the database because a single issue is only a snapshot of the ownership information and firms that exit are dropped from the next issue released. In order to get a full overview of ownership and financials through time, multiple issues are required.

[7] Austria, Belarus, Germany, Greece, Croatia, Italy, Moldova, Macedonia, Serbia, Russia, and Ukraine.

Figure 4.2: Border regions in selected EU countries



The figure shows a map containing those countries used in the analysis in panel (a), as well as the 'border' regions considered in the analysis on the basis of th 75km (panel (b)) and 100km distance definitions (panel (c)).

Panel (a) in Figure 4.2 shows the countries used in the analysis. Figure 4.2 further provides an overview of bordering regions in our data, based on our preferred 75 km definition in panel (b) and based on a 100 km definition in panel (c). Comparing panels (b) and (c) illustrates our motivation for the choice of the 75 km definition as our base scenario. The 75 km definition is better in identifying border regions, whereas the 100 km definition includes nearly all domestic regions in the smaller countries in the analysis as border regions. We use the 100 km definition as a robustness check of our results. From panel (b) one can infer that the ratio between border and non-border regions varies across countries. As expected larger countries such as Romania and Poland have a smaller ratio; medium-sized countries, Hungary and Czech Republic, have a balanced mix between the two; and smaller countries such as Slovenia and Slovakia and have primarily border regions.

Clearly, ownership information is required to distinguish between foreign firms

Table 4.2: Number of foreign firms by country

	10% ownership		50% ownership	
Country	Number	Percentage	Number	Percentage
Austria	10,749	7.47	8,179	7.35
Bulgaria	1,046	0.73	706	0.63
Belarus	653	0.45	653	0.59
Czech Republic	4,418	3.07	3,936	3.54
Germany	60,706	42.18	43,170	38.82
Greece	1,077	0.75	847	0.76
Croatia	1,120	0.78	1,038	0.93
Hungary	3,870	2.69	2,449	2.2
Italy	6,882	4.78	5,460	4.91
Moldova	20	0.01	20	0.02
Macedonia	692	0.48	692	0.62
Poland	9,554	6.64	8,139	7.32
Romania	17,930	12.46	11,675	10.5
Serbia	21,321	14.81	21,321	19.17
Russia	1,971	1.37	1,442	1.3
Slovenia	756	0.53	569	0.51
Slovakia	983	0.68	780	0.7
Ukraine	172	0.12	135	0.12
Total	143,920	100	111,211	100

The table presents the total number of foreign firms by country, by ownership percentage for the entire period.

in the home country and domestic firms over the border. To properly identify the owner, we focus on FDI with foreign ownership of more than 50%. Whilst other papers typically use a lower boundary of 10% ($cf.$ the OECD definition), we employ the higher share in order to be able to properly account for the nationality of the direct owner. A simple count of firms in Table 4.2 suggests that while we do lose some firms by applying the higher cut-off, most of the foreign firms are majority foreign owned and the distribution across countries of foreign firms is very similar between the two definitions.

Table 4.3 presents an overview of domestic, foreign and neighbouring firms by country. Further, the last two columns provide an idea of which countries have larger exposure to foreign investment in neighbouring regions across the border. Bulgaria scores high in this ranking due to the proximity of several regions to the Romanian capital, Bucharest, that has attracted a significant amount of foreign investment. Slovakia, with its capital less than 75 km away from Vienna (Austria) also has a large exposure. Romania on the other hand, has very few foreign firms on its border with both Bulgaria and Hungary, potentially due to the location of their capital cities far away from the Romanian border.

Table 4.3: Number of firms by country

Country	At home						Over border	
	Domestic		Foreign (all)		Neighbours		Foreign	
	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage
Bulgaria	11,125	5.54	706	2.50	172	3.27	4,628	29.09
Czech Republic	38,286	19.08	3,936	13.93	1,451	27.61	1,956	12.29
Hungary	28,860	14.38	2,449	8.67	508	9.67	1,791	11.26
Poland	51,177	25.50	8,139	28.81	2,149	40.89	732	4.60
Romania	55,108	27.46	11,675	41.32	445	8.47	210	1.32
Slovenia	5,416	2.70	569	2.01	248	4.72	1,459	9.17
Slovakia	10,728	5.35	780	2.76	282	5.37	5,135	32.27
Total	200,700	100	28,254	100	5,255	100	15,911	100

The table presents a count of firms by country for the entire period. The foreign at home category contains both the third country foreign firms as well as foreign firms from neighbouring countries.

Table 4.4 shows average values of the different horizontal spillover variables for each country. Most countries show larger foreign shares of economic activity within their borders than outside these. This is again related to the size of the country and the location of major cities within these countries and around their borders. An exception from this are Slovenia and Slovakia, two rather small countries neighboured by Austria (and Italy). As their neighbours are two Western European states with significant foreign activity, they show very similar shares for these two categories. Although Poland shares a border with Germany, this is the former East-Germany where there is less activity.

Figure 4.3 shows the TFP distribution of two groups of domestic firms: those located close to the border and those located in the country centre. We do this for Romania and Czech Republic, two countries with a balanced number of firms over the two groups. While in Romania 'border' firms are slightly more productive than those located further away, in Czech Republic this picture is reversed. This might be driven by the location of their capital cities, and therefore that of the more productive firms: close to a border in the former and far away in the latter. We control for such potential effects by including region fixed effects in our estimations.

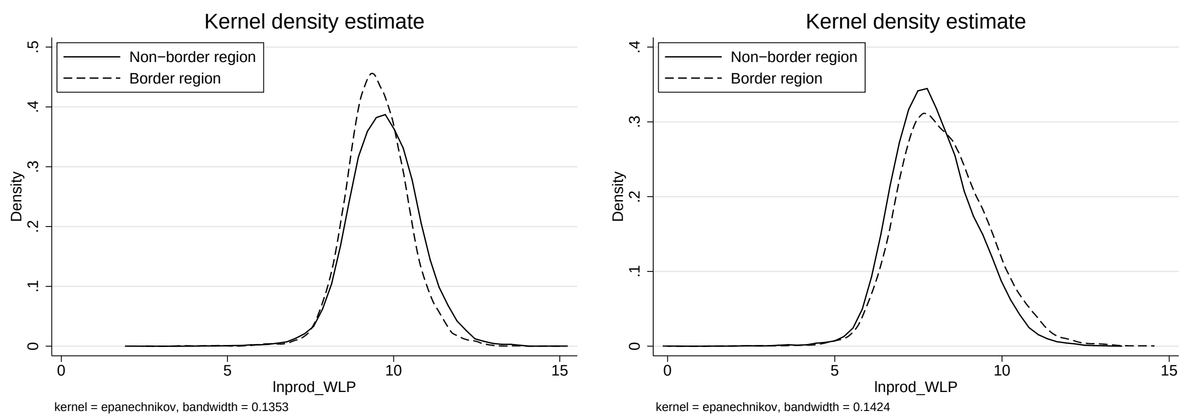
Table 4.4: Spillover intensities by country and type in 2007

Country	HR^{CW}	HR^{AI-H}	HR^{AI-CB}
Bulgaria	0.089	0.009	0.054
Czech Republic	0.082	0.023	0.038
Hungary	0.045	0.027	0.019
Poland	0.145	0.061	0.023
Romania	0.169	0.103	0.004
Slovenia	0.010	0.000	0.035
Slovakia	0.036	0.011	0.033

The table presents spillover intensities of border regions for the year 2007. First column shows the countrywide share of foreign activity, the second column shows the foreign activity as share of the 'circle of interest' and finally the last column contains the foreign share of activity from across the border.

Figure 4.3: Distribution of domestic firms' TFP

(a) Czech Republic **(b)** Romania



4.4 Results

This section presents our results. We first analyse the impact of a national border on productivity spillover effects from MNEs to domestic firms. In a second set of results we focus on border heterogeneity and distinguish between two types of borders, namely Schengen vs. non-Schengen borders. This analysis is motivated by the fact that while Schengen borders allow people and goods to travel freely across borders, non-Schengen borders are characterised by stronger border control, which may constitute a significant impediment for the formation of cross-border relationships. We then present a robustness analysis considering the choice of distance and ownership threshold.

4.4.1 The impact of borders on productivity spillover effects from FDI

Table 4.5 presents our first set of results. In column (1) we start with an analysis of traditional country-wide spillovers. The obtained coefficients represent the country-wide impact on domestic firms' productivity growth of all MNEs present in the home country, irrespective of their location. We detect positive horizontal spillovers in our sample of domestic firms in CEEC countries. On the basis of a large meta-analysis, Havránek and Irŝová (2013) find that the average horizontal spillover effect is zero, but that certain characteristics foster positive spillover effects. They cite a.o. a limited technology gap, higher levels of human capital, and non-fully foreign-owned projects as factors fostering positive spillover effects. These factors are likely to be present in our sample. CEEC countries are known for their high level of human capital and given the time period considered (2000-2010) a considerable catching up process in terms of technology has already been taking place prior to the start of our sample, and certainly near the end of our sample. Furthermore, Merlevede et al. (2014) find that MNEs of sufficient maturity (in terms of presence in the home country) generate positive horizontal spillover effects rather than the negative impact they have shortly after entry. Again, for our sample foreign presence in these CEEC countries is likely to be dominated by MNEs that have been present for a sufficiently long period to generate positive spillover effects overall. Our strong positive result for backward spillover effects does not come as a surprise. It is the most important channel found in the literature for positive spillover effects (see Havránek and Irŝová, 2011), deriving from the incentives of foreign firms to assist their local suppliers. Given that our sample is based on CEEC, where higher quality inputs may be too expensive or too advanced for local firms, the negative forward spillovers are not surprising and confirm previous studies (see e.g. Schoors and van der Tol, 2002). Furthermore MNEs in CEEC-countries tend to be mainly engaged in end-user consumer goods with limited interest in supplying local firms (Damijan et al., 2013).

In column (2) we add $Horizontal^{AI-H}$, $Backward^{AI-H}$, and $Forward^{AI-H}$, i.e. the spillover variables capturing the $differential$ impact from MNEs within 75 km from domestic firms, but located in the home country. This can be seen as an interaction effect that investigates whether being close to MNEs has an additional effect on top of the country-wide spillovers. From column (2) one can infer that results suggest no additional benefits of being close to MNEs. This is in line with Merlevede and Purice (2015), who show that regional patterns do exist but that they are heterogeneous in nature and require a more complex empirical strategy to be detected. This is beyond the scope of this paper as we are primarily interested in analysing border effects in the average spillover effect to domestic firms. This is done in columns (3) and (4) where the cross-border spillover variables $Horizontal^{AI-CB}$, $Backward^{AI-CB}$, and $Forward^{AI-CB}$ are introduced in the specification. We find only significant backward cross-border spillover effects, horizontal and forward cross-border spillover effects being not significantly different from zero. Comparing columns (3) and (4) suggests a negligible impact of the $AI - H$ spillover variables on this result. This implies that domestic firms near the border can benefit in terms of productivity by becoming a supplier to an MNE located nearby, but across the border. What does this imply in terms of border effects? As $Horizontal^{AI-CB}$, $Backward^{AI-CB}$, and $Forward^{AI-CB}$ capture the effect of MNEs in nearby regions across the border, their impact should be compared to the home country spillover effects to make inferences about the impact of national borders on the transmission of productivity spillovers. This is what we do in Table 4.6 where we perform F-tests for the equality of coefficients for home and cross-border effects. For all three spillover types we are able to reject the null hypothesis of equal coefficients. For the horizontal and forward spillover effects that are smaller, national borders seem to block any potential spillover for our sample of CEEC-countries. For the bigger backward productivity spillover effects, we find that national borders significantly decrease spillover effects, but these effects are not fully wiped out. Combined with the fact that the country-wide spillover potential is generally larger than the cross-border spillover potential (as measured by the value of the respective spillover variables, $cf.\ supra$), we establish an economic significant border effect.

Table 4.5: Home country and cross border FDI spillover effects

	(1)	(2)	(3)	(4)
Home country spillover effects				
Country-wide				
Horizontal^{CW}	0.270***	0.274***	0.270***	0.274***
	[0.045]	[0.043]	[0.045]	[0.043]
Backward^{CW}	1.686***	1.695***	1.695***	1.701***
	[0.204]	[0.210]	[0.205]	[0.210]
Forward^{CW}	-0.868***	-0.865***	-0.881***	-0.877***
	[0.154]	[0.148]	[0.154]	[0.148]
Area of interest vs. country-wide				
Horizontal^{AI-H}		-0.022		-0.022
		[0.071]		[0.071]
Backward^{AI-H}		-0.066		-0.047
		[0.197]		[0.196]
Forward^{AI-H}		0.011		-0.004
		[0.185]		[0.182]
Cross-border spillover effects				
Area of interest				
Horizontal^{AI-CB}			0.055	0.055
			[0.069]	[0.069]
Backward^{AI-CB}			1.019***	1.021***
			[0.263]	[0.265]
Forward^{AI-CB}			0.202	0.197
			[0.206]	[0.208]
Observations	314,150	314,150	314,150	314,150
R-squared	0.116	0.116	0.116	0.116

Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4.6: Test for equality of coefficients - country-wide vs. border

F-test	F-value	P-value
Country-wide vs. cross-border		
$HR^{home}=HR^{cross-border}$	7.54	0.0060
$BK^{home}=BK^{cross-border}$	5.31	0.0213
$FW^{home}=FW^{cross-border}$	17.54	0.0000

The table presents a test for the equality of coefficients for countrywide and over the border spillovers with the corresponding F-test value and P-value.

4.4.2 Border heterogeneity and productivity spillover effects from FDI - The impact of Schengen

From the previous subsection we retain that national borders were not 'strong' enough to reduce backward cross-border spillover effects to zero. However, the effect in the previous subsection reflects the average over all borders the CEEC countries have. The above analysis therefore neglects a potentially significant policy change in the form of some of these countries joining the Schengen Area in 2007. This is important as the cost for crossing the border between Schengen countries is almost zero, while crossing borders with non-Schengen countries - be it for people or goods- can be much more difficult due to border controls. Given our result of significant cross-border backward spillover effects and the fact that such relationships entail the transport of goods across the border, we might expect differences between Schengen and non-Schengen borders. In this subsection we therefore differentiate the impact of cross-border foreign presence between borders of two Schengen countries and borders involving at least one non-Schengen country. Naturally, we expect spillover effects to be stronger among Schengen states, since border controls at non-Schengen borders involve delays in merchandise shipments as well as the passage of people when entering or exiting these countries. We perform this test by introducing a Schengen dummy, SH, in the definition of HR^{AI-CB} in order to split the variable in a Schengen and a non-Schengen subcomponent. (4.12) describes the decomposition for the horizontal spillover variable. Backward and forward decompositions follow straightforwardly.

$$
\begin{aligned}
HR_{jt}^{AI-CB} &= \frac{\sum\limits_{i \in j, i \in AI-CB} F_{it} Y_{it}}{\sum\limits_{i \in j, i \in AI} Y_{it}} \\
&= \frac{\sum\limits_{i \in j, i \in AI-CB} SH_{it} F_{it} Y_{it} + \sum\limits_{i \in j, i \in AI-CB} (1 - SH_{it}) F_{it} Y_{it}}{\sum\limits_{i \in j, i \in AI} Y_{it}} \\
&= HR_{jt}^{AI-CB-SH} + HR_{jt}^{AI-CB-NSH} \qquad (4.12)
\end{aligned}
$$

Table 4.7 presents results that are analogous to columns (3) and (4) in Table 4.5 but now consider the further Schengen decomposition.

Table 4.7: Home country and cross border FDI spillover effects - Schengen vs. Non-Schengen borders

	(1)	(2)
Home country spillover effects		
Country-wide		
Horizontal^{CW}	0.274***	0.277***
	[0.045]	[0.043]
Backward^{CW}	1.713***	1.716***
	[0.207]	[0.212]
Forward^{CW}	-0.833***	-0.835***
	[0.152]	[0.146]
Area of interest vs. country-wide		
Horizontal^{AI-H}		-0.020
		[0.071]
Backward^{AI-H}		-0.026
		[0.194]
Forward^{AI-H}		0.025
		[0.180]
Cross-border spillover effects		
Area of interest - Schengen border		
Horizontal$^{AI-CB-SH}$	0.097	0.097
	[0.072]	[0.072]
Backward$^{AI-CB-SH}$	1.292***	1.288***
	[0.299]	[0.302]
Forward$^{AI-CB-SH}$	0.284	0.286
	[0.266]	[0.268]
Area of interest - Non-Schengen border		
Horizontal$^{AI-CB-NSH}$	0.031	0.031
	[0.080]	[0.080]
Backward$^{AI-CB-NSH}$	0.873***	0.871***
	[0.281]	[0.284]
Forward$^{AI-CB-NSH}$	0.131	0.133
	[0.255]	[0.256]
Observations	314,150	314,150
R-squared	0.116	0.116

Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4.8: Test for equality of coefficients -
Country-wide vs. Schengen vs. non-Schengen borders

F-test	F-value	P-value
Country-wide vs. Schengen border		
HR^{home}=$HR^{Schengen}$	4.75	0.0294
BK^{home}=$BK^{Schengen}$	1.70	0.1929
FW^{home}=$FW^{Schengen}$	13.56	0.0002
Country-wide vs. non-Schengen border		
HR^{home}=$HR^{non-Schengen}$	7.56	0.0060
BK^{home}=$BK^{non-Schengen}$	7.09	0.0078
FW^{home}=$FW^{non-Schengen}$	10.63	0.0011
Schengen vs. non-Schengen border		
$HR^{Schengen}$=$HR^{non-Schengen}$	0.76	0.3830
$BK^{Schengen}$=$BK^{non-Schengen}$	2.58	0.1082
$FW^{Schengen}$=$FW^{non-Schengen}$	0.22	0.6356

The table presents a test for the equality of coefficients for Schengen vs. non-Schengen borders with the corresponding F-test value and P-value.

With respect to the home country spillover effects, our previous results are confirmed. In terms of the Schengen-non-Schengen decomposition of cross-border spillover effects we find for cross-border horizontal and forward spillover effects no difference between Schengen and non-Schengen borders and our earlier results of non-significance is confirmed. For cross-border backward spillover effects we find a significant effect for both Schengen and non-Schengen borders. We do observe, however, a clear ranking in terms of the size of the point estimate. The within country backward spillover coefficient is larger than its cross-border Schengen counterpart, which in turn is larger than the non-Schengen coefficient. The F-tests in Table 4.8, however, indicate that we cannot reject the within country backward spillover coefficient to be equal to its cross-border Schengen counterpart, whereas it is significantly different from its cross-border non-Schengen counterpart at the 1% level. Comparing the Schengen and non-Schengen coefficients, we can almost reject equality at the 10% level.

4.4.3 Robustness

In this section we test the robustness of our results by varying two choices we made when calculating our spillover variables. The first relates to the definition of what constitutes a foreign firm, the second to the 75km as distance for defining the area of interest. For the definition of 'foreignness' we chose the 50% foreign ownership criterion for the reasons mentioned above. However, most spillover studies apply a 10% criterion for considering a firm as foreign. To define our area of interest, we opted for 75km distance between region centres because this results in a nicer 'corridor' around the borders. The 100km definition on the other hand does include all border regions, but leaves almost no pure home-country regions in the smaller countries of our sample, as can be seen from Figure 4.2 above.

Tables 4.9 and 4.10 repeat the last specifications from Tables 4.5and 4.7 respectively. Table 4.9 confirms our earlier findings with respect to the average cross-border backward spillover effect. Using the 100km definition, the point estimate goes down somewhat. Furthermore, the negative forward spillover effect in the home country seems to decrease with more foreign presence in upstream industries in the home country located in the area of interest. Table 4.10 suggests that the former result is mainly due to the fact that this alternative distance choice reduces the cross-border backward spillover effect to zero. Furthermore, using the the 100km-50% specification, we see some evidence of cross-border horizontal spillovers in Schengen countries. The F-tests for these specifications are reported in the Appendix.

All in all these results seem to strengthen our earlier conclusion that national borders constitute an almost insurmountable barrier for horizontal and forward spillover effects. In the case of backward spillover effects, national borders significantly dampen cross-border spillover. However, the size of the border impact seems related to the '*depth*' of the border, as evidenced by the difference between Schengen and non-Schengen countries. Clearly, this offers scope for the analysis of other border characteristics as can be found in the trade literature such as cultural similarities, euro adoption, etc. This is left for further research however.

Table 4.9: Home country and cross border FDI spillover effects - Robustness w.r.t. distance and ownership definitions

	(1) $75km - 10\%$	(2) $100km - 50\%$	(3) $100km - 10\%$
Home country spillover effects			
Country-wide			
Horizontal^{CW}	0.286***	0.264***	0.293***
	[0.043]	[0.045]	[0.045]
Backward^{CW}	1.745***	1.629***	1.721***
	[0.209]	[0.217]	[0.227]
Forward^{CW}	-0.882***	-1.083***	-1.068***
	[0.147]	[0.165]	[0.161]
Area of interest vs. country-wide			
Horizontal^{AI-H}	-0.026	0.015	-0.036
	[0.068]	[0.050]	[0.048]
Backward^{AI-H}	-0.261	0.142	-0.107
	[0.186]	[0.172]	[0.176]
Forward^{AI-H}	0.022	0.419***	0.424***
	[0.170]	[0.144]	[0.122]
Cross-border spillover effects			
Area of interest			
Horizontal^{AI-CB}	0.015	0.289	0.171
	[0.064]	[0.180]	[0.170]
Backward^{AI-CB}	0.811***	0.681**	0.575**
	[0.239]	[0.284]	[0.263]
Forward^{AI-CB}	-0.087	0.222	-0.107
	[0.226]	[0.238]	[0.237]
Observations	305,251	314,150	305,251
R-squared	0.115	0.117	0.115

Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4.10: Home country and cross border FDI spillover effects - Schengen vs. Non-Schengen borders - Robustness w.r.t. distance and ownership definitions

	(1) $75km - 10\%$	(2) $100km - 50\%$	(3) $100km - 10\%$
Home country spillover effects			
Country-wide			
Horizontal^{CW}	0.289***	0.267***	0.296***
	[0.043]	[0.045]	[0.045]
Backward^{CW}	1.759***	1.641***	1.731***
	[0.210]	[0.219]	[0.229]
Forward^{CW}	-0.845***	-1.033***	-1.021***
	[0.146]	[0.162]	[0.158]
Area of interest vs. country-wide			
Horizontal^{AI-H}	-0.024	0.017	-0.033
	[0.068]	[0.050]	[0.048]
Backward^{AI-H}	-0.245	0.161	-0.081
	[0.184]	[0.171]	[0.175]
Forward^{AI-H}	0.049	0.423***	0.428***
	[0.170]	[0.142]	[0.120]
Cross-border spillover effects			
Area of interest - Schengen border			
Horizontal$^{AI-CB-SH}$	0.042	0.184**	0.112
	[0.078]	[0.091]	[0.094]
Backward$^{AI-CB-SH}$	1.077***	0.944***	0.971***
	[0.277]	[0.311]	[0.286]
Forward$^{AI-CB-SH}$	-0.127	0.382	-0.153
	[0.278]	[0.250]	[0.262]
Area of interest - Non-Schengen border			
Horizontal$^{AI-CB-NSH}$	0.002	0.308	0.179
	[0.070]	[0.201]	[0.190]
Backward$^{AI-CB-NSH}$	0.676***	0.531*	0.398
	[0.254]	[0.304]	[0.280]
Forward$^{AI-CB-NSH}$	-0.074	0.095	-0.121
	[0.251]	[0.292]	[0.269]
Observations	305,251	314,150	305,251
R-squared	0.115	0.117	0.115

Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

4.5 Conclusion

The literature on productivity spillovers of FDI has primarily focused on within-country effects of MNE presence. In this paper we analyse the potential for cross-border spillover effects. We use a large multi-country firm-level dataset to measure indirect FDI spillovers in a number of Central and Eastern European countries during the 2000-2010 period. We employ information on the region where firms are located, and combine this with the distance between regions to analyse whether domestic firms within a certain distance from the border experience productivity gains as a result of nearby MNE activity from across the border. We compare this with the spillover effects generated within the local firm's home country and home region. Our findings suggest that national borders constitute an almost insurmountable barrier for horizontal and forward spillover effects. In the case of backward spillover effects, national borders significantly dampen cross-border spillovers, but the size of the border impact seems related to the '*depth*' of the border, as evidenced by differences between Schengen and non-Schengen countries.

Our finding has a broader implication for policy makers. On the one hand, it suggests a greater integration of the European internal market, where national borders play an ever smaller role in maintaining economic barriers between countries and allow for cross-country technology diffusion. On the other hand, since backward spillover effects potentially extend beyond the country of location, competition between policy makers for attracting FDI seems less grounded. While indeed the country-wide effect of FDI is significant and important, some spillover effects will inevitably reach neighbouring countries as well. Our results suggest that there could be benefits from cooperation among neighbouring countries in attracting FDI.

Appendix 4.A Additional tables

Table 4.A.1: Test for equality of coefficients for 75km - 10% specification

F-test	F-value	P-value
All borders		
$HR^{home}=HR^{cross-border}$	12.61	0.0004
$BK^{home}=BK^{cross-border}$	11.81	0.0006
$FW^{home}=FW^{cross-border}$	9.87	0.0017
Schengen borders		
$HR^{home}=HR^{Schengen}$	7.75	0.0054
$BK^{home}=BK^{Schengen}$	5.00	0.0254
$FW^{home}=FW^{Schengen}$	6.06	0.0138
non-Schengen borders		
$HR^{home}=HR^{non-Schengen}$	12.58	0.0004
$BK^{home}=BK^{non-Schengen}$	13.78	0.0002
$FW^{home}=FW^{non-Schengen}$	7.59	0.0059
Schengen - non-Schengen borders		
$HR^{Schengen}=HR^{non-Schengen}$	0.31	0.5782
$BK^{Schengen}=BK^{non-Schengen}$	3.12	0.0774
$FW^{Schengen}=FW^{non-Schengen}$	0.04	0.8343

The table presents a test for the equality of coefficients for Schengen vs. non-Schengen borders with the corresponding F-test value and P-value.

Table 4.A.2: Test for equality of coefficients for 100km - 50% specification

F-test	F-value	P-value
All borders		
HR^{home}=$HR^{cross-border}$	0.02	0.8913
BK^{home}=$BK^{cross-border}$	8.50	0.0036
FW^{home}=$FW^{cross-border}$	19.22	0.0000
Schengen borders		
HR^{home}=$HR^{Schengen}$	0.65	0.4184
BK^{home}=$BK^{Schengen}$	3.95	0.0468
FW^{home}=$FW^{Schengen}$	22.24	0.0000
non-Schengen borders		
HR^{home}=$HR^{non-Schengen}$	0.04	0.8459
BK^{home}=$BK^{non-Schengen}$	10.23	0.0014
FW^{home}=$FW^{non-Schengen}$	10.70	0.0011
Schengen - non-Schengen borders		
$HR^{Schengen}$=$HR^{non-Schengen}$	0.73	0.3928
$BK^{Schengen}$=$BK^{non-Schengen}$	2.09	0.1486
$FW^{Schengen}$=$FW^{non-Schengen}$	0.87	0.3517

The table presents a test for the equality of coefficients for Schengen vs. non-Schengen borders with the corresponding F-test value and P-value.

Table 4.A.3: Test for equality of coefficients for 100km - 10% specification

F-test	F-value	P-value
All borders		
HR^{home}=$HR^{cross-border}$	0.47	0.4931
BK^{home}=$BK^{cross-border}$	13.12	0.0003
FW^{home}=$FW^{cross-border}$	11.84	0.0006
Schengen borders		
HR^{home}=$HR^{Schengen}$	3.01	0.0828
BK^{home}=$BK^{Schengen}$	5.28	0.0216
FW^{home}=$FW^{Schengen}$	9.04	0.0027
non-Schengen borders		
HR^{home}=$HR^{non-Schengen}$	0.35	0.5534
BK^{home}=$BK^{non-Schengen}$	15.72	0.0001
FW^{home}=$FW^{non-Schengen}$	8.32	0.0039
Schengen - non-Schengen borders		
$HR^{Schengen}$=$HR^{non-Schengen}$	0.26	0.6107
$BK^{Schengen}$=$BK^{non-Schengen}$	4.81	0.0283
$FW^{Schengen}$=$FW^{non-Schengen}$	0.02	0.9000

The table presents a test for the equality of coefficients for Schengen vs. non-Schengen borders with the corresponding F-test value and P-value.

5

Do Parent-Affiliate Characteristics Affect Firm Performance: A View Through the Cycle*

5.1 Introduction

After several rounds of trade liberalisation, the past couple of decades have witnessed a significant increase in the fragmentation of production chains and the internationalisation of firm ownership. These developments however have given rise to concerns regarding the volatility of employment in companies that are either foreign-owned or operate in multiple locations. Moreover, for countries that are heavily relying on FDI, such as Ireland and Slovakia, this issue has become even more prominent during the latest financial crisis, when significant decreases in output, trade and employment have been recorded.

Extensive research has already looked at the performance of domestic companies compared to that of foreign firms, with mixed results. On the one hand we know that multinationals (MNEs) are usually more productive and pay higher wages (Bernard

*This chapter is based on a paper co-authored with Bruno Merlevede.

et al. (1995)). Moreover, older and larger firms, as many subsidiary owning firms are, will be more resistant to crises (see e.g. Dunne and Hughes (1994)). On the other hand, companies operating in several countries are accustomed to moving and could be more flexible in relocating to another country. The existing literature confirms this, as most studies either do not find a significantly better performance on the part of multinational affiliates compared to local firms, or record a higher level of exit among them. For example Görg and Strobl (2002) find that MNEs are more 'footloose' than domestic plants, ceteris paribus. For Belgian firms, Blanchard et al. (2012) conclude that MNEs are indeed 'footloose' and will exit much faster than similar national firms, with potential negative effects for the industry. Navaretti et al. (2003) find that MNEs adjust labour demand faster than national companies, however they are better at maintaining employment in the longer run. Bernard and Jensen (2007) find that even at home US multinationals are more 'footloose' compared to their domestic only counterparts, suggesting that the employment preferences are different not so much between domestic and foreign companies but between MNEs and *non*-MNEs. Comparing firm performance during a crisis in Chile, Alvarez and Görg (2007) do not find that MNEs react better than local firms. A similar study on Irish data by Godart et al. (2011) suggest that foreign firms exit the economy just as often as domestic firms during a period of economic downturn.

Nonetheless, a limited number studies which have a closer look at parent characteristics reveal that, under certain conditions, multinationals could represent a stabilizing force for the local economy. For example, MNEs will be less sensitive to changes in wages and thus less 'footloose' if they have a more extensive network of local linkages (Görg et al. (2009)). In case of local financial crisis, multinationals could be more stable than national counterparts due to their ability to relocate sales to other countries (see Dikova et al. (2013)). The distance from headquarters seems to matter as well, with papers finding that geographically dispersed companies will divest from further away locations compared to those in the same state/country as their headquarters (see e.g. Landier et al., 2009; Abraham et al., 2014). They suggest that multinationals will prefer maintaining plants that are closer by for two main reasons. First communication and information flows from larger distances are more difficult to maintain while monitoring is more costly. Second, these firms have closer ties with the community at home, speak the language and have a better knowledge of the local culture and will therefore try to avoid potential backlash from downsizing close-by (see e.g.Norback et al. (2013)).

Lastly, an emerging literature is looking at the impact of financial market development on multinationals and their investment activities. On the one hand, financing frictions are recognized as important factors for MNE's decision making and thus potentially as predictors of affiliate (re-)location. A recent paper by Poelhekke (2015) suggests that the rise of multinational banks has facilitated the expansion of foreign direct investment, especially coming from the same country as the respective bank. On the other hand, the internal capital market of multinationals is important in facilitating the activity of affiliates, especially during crisis. Hebous and Weichenrieder (2010) find that wholly owned affiliates are able to circumvent sharp currency depreciations by accessing capital from their parent companies. Using Chinese data, Manova et al. (2015) find that multinational subsidiaries are less liquidity constrained, having access to either the internal capital of its parent or broader financial markets. Finally, Alfaro and Chen (2012) compares the performance of domestic and foreign owned firms during the last financial crisis and conclude that foreign affiliates have performed better, especially if they had strong vertical production and financial ties with their parent.

Given how important the performance of affiliates could be in maintaining economic growth and employment, we investigate what factors impact their performance. Our contribution is twofold. First, unlike most of the papers so far, we try to asses affiliate performance based on the characteristics of parent. Moreover, since we have access to networks of multinationals in and from multiple EU counties, we can explore not only the traditional firm level characteristics of the parent, but also the financial institutions variation in the country of origin of parents and their affiliates. Second, we add to the literature by looking at these firm-specific factors through the cycle, by choosing a period that encompasses both the years prior to the latest financial crisis as well as those in which the crisis was recorded. Since firms might behave differently during good times compared to bad times, one of the aims of this paper is to pin down those characteristics that improve the resilience of affiliates to economic shocks through they cycle.

For the purpose of our analysis we use the *AUGAMA* firm-level database as well as the *EUMULNET* database, comprising ownership information of firms as well as financial and balance sheet information. Due to unique country identifiers assigned to each firm, we are able to reconstruct the network of affiliates a parent owns as well as important data regarding their activities. Besides looking at standard performance measures, we also take into account whether the parent is active in one national market or is a multinational, whether it owns one or multiple affiliates, is listed on the stock exchange or not and how well industrially diversified it is. On the institutional side, we account for the financial development of the country of parent/affiliate by including the country levels of credit to private sector as percentage of GDP.

Our results show that larger, more productive and listed parents are more likely to close down an affiliate, while belonging to a multinational parent improves affiliate survival. Moreover, industrial diversification of the parent improves affiliates survival, especially during the crisis years. Finally, we find parents that come from countries with looser credit environments positively affect affiliate survival, but this effect is mainly due to the pre-crisis years. In the aftermath of the crisis, even parents that on average had access to more loans, as proxied by county private credit ratios, are not able to improve the probability of survival of their affiliates.

The remainder of the paper is structured as follows. Section 2 introduces our data and explains how we constructed the dataset. Sections 3 and 4 present the results and robustness checks and we conclude in section 5.

5.2 Data

For this analysis we are using the *AUGAMA* firm-level database as well as the *EUMULNET* database. From the former database we obtain balance sheet information and profit and loss accounts for European parents and their affiliates, while from the latter we obtain the ownership information, or the EU network of affiliates belonging to one parent company. Since ownership information is gathered at irregular intervals, we do not have ownership information for all firm-owner-year combinations. Given this, we first create a dataset at the parent-affiliate-year-level with the available information and then fill out the missing firm-owner-year-entries under the assumption that the older information is valid until a change is recorded. After building this parent-affiliate panel, we then update it with the financial information from the country where the parent/subsidiary is located, as reported in *AUGAMA*.

The data are deflated using capital and output deflators as appropriate. Moreover, the data is cleaned by truncating at the 1st and 99th percentile in each country in terms of wage, employment, total assets and sales.[1] While most of the financial information, including several ratios, are directly available in the original database, we do construct some variables ourselves. Notably we compute the average wage per firm by taking the ratio of the cost of employees to the number of employees. To further get rid of outliers and reporting mistakes, we drop all firms where the ratio between the smallest and largest wage per firm during the sample period is larger than 3, which often implies that the costs of employees do not grow at a similar rate as the number of employees. This constraint is especially violated in firms operating in non-manufacturing sectors.

To this end our dataset consists of 20 European countries, which in turn consists of about 30,000 parent-subsidiary links where the ownership is above 50%. We only consider the majority owned subsidiaries since we assume that these parents would have a say in what happens with their affiliate. To account for the cycle, we focus on the period 2004-2010. Table 5.1 gives an overview of the data, by showing the frequency of firms across the various European countries. Due to the lack or limited availability of data for some countries, we are unable to consider all of the EU, however we do have a good mix of both Northern as well as Southern and Central and Eastern European countries (CEEC). We consider manufacturing and sales affiliates only, with robustness analysis performed separately on the manufacturing firms.

We use affiliate exit as determinant of its performance and assume that as long as the affiliate exists, it is profitable for the parent to maintain its operations. An exit is defined when the firm does not have any financial information available in the database. Moreover, we exclude from our analysis all firms that have been sold off or cease to be majority owned during the sample period. This leave us with an average exit rate of close to 7%. Figure 5.1 gives an overview of the average exit rate per country.

As the exit decision of an affiliate will be determined by both its own characteristics as well as those of its parent, we will next review some of the variables that we use in order to predict exit.

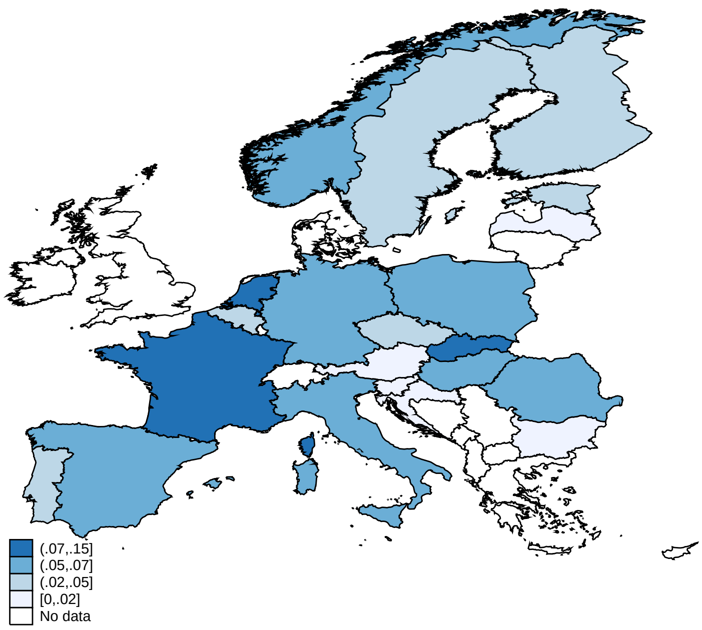
[1] For a detailed description of the data, please consult Chapter 2.

Table 5.1: Frequency of firms by country in 2007

	# parents	# affiliates owned	# abroad	# affiliates in country	# foreign owned
Austria	67	102	69	74	41
Belgium	420	634	162	531	59
Bulgaria	41	49	0	52	3
Czech Republic	34	42	3	70	31
Germany	274	633	176	567	110
Estonia	24	26	0	37	11
Spain	1854	2554	117	2678	241
Finland	44	46	5	42	1
France	2210	3383	268	3351	236
Croatia	72	91	0	91	0
Hungary	58	60	5	85	30
Italy	737	1027	172	975	120
Latvia	1	1	1	0	0
The Netherlands	48	86	68	35	17
Poland	41	46	3	121	78
Portugal	348	442	12	486	56
Romania	50	57	3	103	49
Sweden	927	1116	37	1104	25
Slovenia	4	7	7	0	0
Slovakia	2	2	1	2	1

The table shows a count of parents and affiliates for each country by domestic/multinational status in 2007. The first three rows contain information on the number of parents in one country, how many affiliates these parents own and how many of these owned affiliates are abroad. The last two rows sum the total number of affiliates in a country and how many of them have a foreign parent.

Figure 5.1: Affiliate exits as percentage of total number of firms



The figure shows the average exit rate of affiliates in own country as percentage of total firms in that country in years 2005-2010.

5.2.1 Affiliate characteristics

To control for the performance of the affiliate, we use a set of firm level variables as well as an industry level variable, as suggested by previous literature. Given that firm size and firm age are associated with higher firm longevity (see Dunne et al. (1988)), we employ these variables in our estimations. Our proxy for firm size is the log of number of employees while age is the number of years in operation (different from number of years owned by the parent). Moreover, as productivity is a good determinant of firm survival (see e.g. Melitz (2003)), we include a measure of affiliate labour productivity as well. We control for capital intensity of the firm, as it might have a positive impact on survival by acting as collateral in the face of a shock. We therefore add the log of capital-labour ratio as well as the log of the average wage paid by the firm, to proxy for employment costs.

As a measure of the competitiveness of the industry the firm is operating in, we borrow an industry entry cost measure from Bernard and Jensen (2007), which should give us an idea of whether an industry is expanding or contracting. This measure takes the yearly minimum of the entry and exit rates, such that:

$$\mathrm{EC}_{it} = 1 - \min[\mathrm{entryrate}_{it}, \mathrm{exitrate}_{it}]$$

where the rates are calculated for each industry i and year t irrespective of country.

5.2.2 Parent characteristics

Similarly to affiliate performance, we use size, age, wage, capital intensity and productivity of the parent to predict affiliate exit, since these characteristics could also be important factors in predicting whether affiliates survive or will shut down. As suggested by Bernard and Jensen (2007), larger and more productive parents might find it easier to relocate production, thus closing down an affiliate.

Table 5.2 shows summary statistics of some of the main parent and affiliate firm-level characteristics used in the estimations, as split by exiter/survivor status. We observe that at first look, the two groups are very similar both in terms of their own characteristics as well as those of their parents. Nonetheless, an F-test rejects their equality for most of the variables included. Survivors are generally larger, more productive and more capital intensive. Their parents are also very different with larger, more productive and less capital intensive parents closing their affiliates faster.

Table 5.2: Summary statistics by exiting/surviving firm

		Exiter	Survivor	F-test
Affiliate	Size	2.75	2.88	0.00
	Age	17.97	19.78	0.00
	Wage	10.31	10.32	0.26
	Productivity	12.04	12.14	0.00
	Capital Intensity	9.11	9.23	0.00
Parent	Size	3.46	3.14	0.00
	Age	25.57	26.07	0.03
	Wage	10.60	10.66	0.00
	Productivity	12.16	12.09	0.00
	Capital Intensity	9.78	9.88	0.00
	Multi-firm	0.45	0.47	0.00
	Multinational	0.11	0.15	0.00
	Listed parent	0.01	0.01	0.33
	Observations	9154	60264	

The table presents summary statistics of main variables by exiter status (i.e. the firms will exit sometime during the sample period). The F-test suggests that aside from affiliate wage and whether one's parent is listed, there are significant differences between affiliates that survive and those that do not.

Aside from these factors, we also consider additional variables that could impact affiliate exit.

Multinational Being a multinational, or owning affiliates in a different country than the one the parent is itself located, could impact affiliate survival due to the 'footloose' nature of the parent, as suggested by Görg and Strobl (2002) and Bernard and Sjöholm (2003). We therefore include a dummy for multinational status of the parent.

Multi-firm We also check whether owning several affiliates is detrimental to affiliate survival. We define these parents as "multi-firm". Given that most parents in our data own just one affiliate, we prefer a binary variable of whether the parent owns more than one affiliate instead of a continuous one with the actual number of affiliates owned. The literature is rather mixed when it comes to this factor, as some find that belonging to a group improves survival (see e.g. Disney et al. (2003)), while others on the contrary find that, controlling for plant characteristics, firms with multiple plants will shut down easier (see e.g. Bernard and Jensen (2007)). We expect parents with a larger number of affiliates to be more flexible with managing them, and thus to shut them down easier.

Table 5.3 provides an overview of the structure of the networks in the data. We can see that the majority of parents (66%) own a single domestic affiliate, while less than one percent of firms own more than 10 affiliates. Moreover, approximately 17% of affiliates are foreign and about 16% of parents are multinational.

Table 5.3: Distribution of the number of affiliates per parent for the year 2007

		Parents				Affiliates			
		Domestic		Multinational		Domestic		Foreign	
Affiliates	1	5321	0.733	594	0.082	5321	0.572	594	0.536
	2	875	0.121	109	0.015	1750	0.188	218	0.197
	3	243	0.033	33	0.005	729	0.078	99	0.089
	4	101	0.014	13	0.002	404	0.043	52	0.047
	5	42	0.006	4	0.001	210	0.023	20	0.018
	6	27	0.004	7	0.001	162	0.017	42	0.038
	7	22	0.003	2	0.000	154	0.017	14	0.013
	8	6	0.001	1	0.000	48	0.005	8	0.007
	9	5	0.001	2	0.000	45	0.005	20	0.018
	10 or more	19	0.003	2	0.000	472	0.051	42	0.038
	Total	7256		7256		9295		1109	

The table shows a count of parents by the number of affiliates they own and by domestic/multinational status.

Listed We include a dummy for whether the parent is a listed firm or not. This is motivated by the fact that firms whose shares are publicly owned are usually much larger and older than non-listed firms, characteristics which could be associated with higher exit rates for their affiliates. Moreover, these firms might be more sensitive to the stock market, closing their affiliates faster than non-listed firms, especially in period of economic downturn.

Table 5.4 breaks down the same characteristics by parent type, where parents divided in three major categories based on whether they own one or multiple affiliates, whether they operate in a single country or are a multinationals and whether they are listed or not. We observe that not only are "multi-firm", multinational and listed parents larger and older than their counterparts, but so are their affiliates. Moreover, while these different categories pay largely similar wages, there are visible TFP differences between them, with 'multi-firm", multinational and listed parents and their affiliates being more productive, as consistent with the literature.

Table 5.4: Parent and affiliate characteristics by parent type

		Single affiliate	Multi-firm	Domestic	Multinational	Non-Listed	Listed
Affiliate	Size	2.63	3.13	2.74	3.61	2.85	4.09
	Age	18.91	20.19	19.42	20.04	19.48	23.01
	Wage	10.25	10.39	10.29	10.45	10.32	10.08
	Productivity	12.02	12.25	12.08	12.39	12.13	11.93
	Capital intensity	9.19	9.23	9.19	9.33	9.21	9.30
Parent	Size	2.93	3.48	2.91	4.90	3.17	5.37
	Age	24.07	28.18	24.52	34.98	25.86	40.58
	Wage	10.49	10.83	10.61	10.92	10.65	10.43
	Productivity	12.03	12.18	12.01	12.62	12.10	11.70
	Capital intensity	9.85	9.88	9.82	10.13	9.86	10.07
	Observations	36871	32547	59682	9736	68843	575

The table shows summary statistics of main variables by single/multi-firm, domestic/multinational and listed/non-listed parent status.

Distance Consistent with the findings that multinationals are prone to closing down affiliates that are located further from headquarters, we add the distance between parent and affiliate as explanatory variable. We expect that, due to larger monitoring costs, distance will have a positive impact on affiliate exit (see e.g. Giroud (2012)).

Diversification level We test whether industrial diversification of parents has a positive impact on affiliate survival. This is the mirror image of multinationality and often is a complementary strategy of parents, as some will choose to diversify via going abroad while others will do so in terms of the number of industries they operate in (see Davies et al. (2001)). The literature suggests that diversification will increase the probability of exit (see Gibson and Harris (1996)). Nonetheless, more diversified parents could potentially protect their affiliates better if the industries their are active in do not experience shocks that are highly correlated. Firm diversification is measured using the entropy index as follows:

$$\text{Diversification} = \Sigma P_i ln(1/P_i)$$

where P_i is the share of the industry i in the total sales of the parent and $ln(1/P_i)$ is the weight for each industry in parent's portfolio. Higher values signify better diversification of the parent. To correct for the highly non-normal distribution of this variables due to many parents owning a single affiliate, we use the log of diversification

Table 5.5: Correlation coefficients of affiliate, parent and group characteristics

		Affiliate					Parent				
		Size	Age	Wage	TFP	K/L	Size	Age	Wage	TFP	K/L
Affiliate	Size	1.00									
	Age	0.24	1.00								
	Wage	0.03	0.13	1.00							
	Productivity	-0.04	0.04	0.54	1.00						
	Capital Intensity (K/L)	0.17	0.08	0.10	0.18	1.00					
Parent	Size	0.33	0.05	0.12	0.14	0.09	1.00				
	Age	0.18	0.15	0.10	0.08	0.05	0.38	1.00			
	Wage	0.18	0.09	0.44	0.23	0.06	-0.14	-0.01	1.00		
	Productivity	0.13	0.04	0.23	0.25	0.08	0.21	0.14	0.33	1.00	
	Capital Intensity (K/L)	0.08	-0.03	0.02	0.07	0.15	0.13	0.19	0.00	0.25	1.00

The table shows correlation coefficients of main variables.

as an independent variable.

Production linkages Given that production linkages have been found as important predictors of affiliate performance (see e.g. Alfaro and Chen (2012)), we include a dummy for whether affiliates and parents are active at the same Nace 4 digit level or not. We define an affiliate as vertical if it does not operate in the same industry as the parent. We expect vertical affiliates to have a lower exit rate due to their potentially important role in the production chain of the parent.

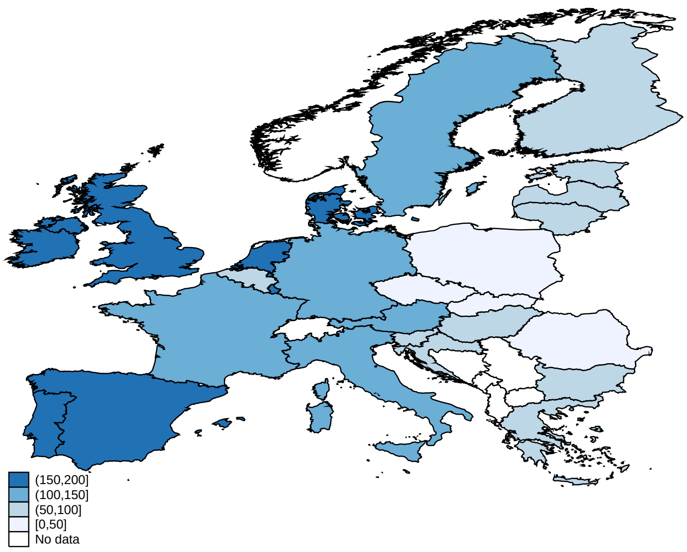
Financial linkages Consistent with the findings of Manova et al. (2015) that financial linkages can offer protection to affiliates via internal capital flows from parents, we check whether differences in the development of financial markets of parents and affiliates has an impact on affiliate exit. Since the true extent of parent's financial help is not available in the data, we assume that parents from countries with easier access to finance will pass that over to their affiliates if need be.

For data on country level financial development we use the private credit by deposit banks and other financial institutions as percentage of GDP provided by the World Bank (see Beck et al. (2000)). This measure has been previously used in the trade literature by Manova (2013) as well as Vandenbussche et al. (2013).[2] Since we are potentially dealing with two credit environments from the country of parent and that of the affiliate, for

[2]We replace the private credit values for Slovakia with those of the same variable taken from the World Bank's Financial structure database, since the original data contains missing values for this country. The two databases are very similar otherwise.

our estimations we include both the parent's country score as well as that of the affiliate. Figure 5.2 gives an overview of EU variation in terms of the average share of GDP of domestic credit, while Table 5.6 reports summary statistics for this variable.

Figure 5.2: Domestic credit to private sector as % of GDP



The figure shows the average domestic credit to the public sector as % of GDP for the years 2005-2010.

We observe that the countries with largest shares of consumer credit to GDP are primarily located in Western Europe, while CEEC levels of credit are relatively low. The discrepancy between the highest and lowest scoring countries is quite high, with Romania having a level as low as 20% in 2005 while in Spain this indicator is close to 150% for the same period. Naturally, this variable does not perfectly describe the credit constraints of companies, as part of this credit is owned by consumer an is therefore non-corporate.[3]

As robustness we employ industry level data on how dependent manufacturing industries are on external finance. This measure was first introduced by Rajan and Zingales (1998) in an attempt to control for endogeneity when explaining sectoral growth, and it is defined as the share of capital expenditure not financed by cash flow. While some sectors are inherently more dependent on finance, they will be the ones affected most by financial constraints. We borrow this measure form Bena and Ondko (2012), who calculate the external finance dependence based on European firm level data taken

[3]Notably in the Netherlands this ratio is high due to a large household debt.

Table 5.6: Domestic credit to private sector by country

Country	Mean	St. Deviation
Austria	115.73	4.00
Belgium	86.46	8.80
Bulgaria	60.15	14.53
Czech Republic	44.17	7.00
Germany	105.87	3.03
Estonia	89.56	12.49
Spain	183.40	26.14
Finland	81.56	7.70
France	102.04	7.98
Croatia	61.82	6.03
Hungary	61.97	7.81
Italy	100.19	11.90
Latvia	89.78	12.51
The Netherlands	176.01	16.43
Poland	41.94	9.59
Portugal	161.70	18.24
Romania	36.16	11.25
Sweden	117.15	10.79
Slovenia	77.40	14.71
Slovakia	40.43	6.63

The table contains country summary statistics of financial development as taken from the the World Bank on-line database.

from Amadeus. Table 5.7 provides an overview of this ranking of industries, where higher values indicate larger financial dependence.

All in all, our dataset consists of nearly 75,000 observations. For each firm we estimate the probability of exit as a pooled Probit (following Bernard and Jensen (2007)) :

$$Pr(D_{af_t} = 1|Z_t) = \Phi(\beta^{af} Z_t^{af} + \beta^{pa} Z_t^{pa} + \beta^{ind} Z_t^{ind} + c_{ctr} + c_{ind} + c_t)$$

where Z^{af} is a vector of affiliate characteristics and Z^{pa} a vector of parent characteristics. Furthermore, we cluster standard errors at the level of affiliate's country and include country, industry and year fixed effects.

5.3 Results

This section presents the results of our main estimations. We consider the impact that parent firm level characteristics have on the probability of affiliate exit as well as whether having parents located in countries with more developed financial markets improves survival rate of affiliates.

Table 5.7: Industry external finance dependence

Nace code	Industry	Ext. finance dep.
15	Food products and beverages	-0.117
16	Tobacco products	-0.124
17	Textiles	0.204
18	Wearing apparel	-0.382
19	Leather	-1.548
20	Wood and cork	0.061
21	Pulp and paper products	-0.02
22	Publishing and printing	-0.217
23	Coke and refined petroleum	0.658
24	Chemical products	1.821
25	Rubber and plastic	-0.082
26	Other non-metallic mineral prod.	-0.101
27	Basic metals	0.073
28	Fabricated metal products	-0.32
29	Machinery and equipment	-0.117
30	Office machinery and computers	0.886
31	Electrical machinery and communication equipment	0.417
32	Radio, tv, and communication equipment	0.84
33	Medical, precision and optical instruments	0.665
34	Motor vehicles	0.17
35	Other transport equipment	-0.117
36	Manufacturing n.e.c.	-0.141

The table contains industry values of external finance dependence calculated for European firms as taken from Bena and Ondko (2012).

5.3.1 Affiliate exit and firm structure

Table 5.8 shows the main estimation results, in the form of marginal effects. Controlling for country, industry and year fixed effects we observe that practically all of the included characteristics matter in predicting affiliate exit. While larger, older, more productive and more capital intensive affiliates seem to improve affiliate survival, distance from parent increases probability of shut-down, as expected. Vertical affiliates are also more likely to survive, suggesting that they might play a crucial role in the production network of the parent.

As far as the characteristics of the parents are concerned, we see that larger, more productive and listed parents are more likely to shut down their affiliates, while the more capital intensive parents have the opposite effect. This is somewhat in line with Bernard and Jensen (2007), as they find multinationals and multi-plant firms, who are usually larger, more productive and present on the stock exchange, to have a negative effect on plant survival.

Table 5.8: Affiliate exit and parent characteristics

	(1) 2004-2010	(2) 2004-2010	(3) 2004-2010	(4) 2004-2007	(5) 2008-2010
Affiliate characteristics					
Size	-0.00773***	-0.00710***	-0.00926***	-0.01076***	-0.00788***
	(0.00086)	(0.00087)	(0.00158)	(0.00170)	(0.00170)
Age	-0.00014**	-0.00015**	-0.00011	-0.00015**	-0.00005
	(0.00007)	(0.00007)	(0.00007)	(0.00007)	(0.00010)
Wage	0.00299	0.00493*	0.00280	0.00053	0.00417
	(0.00271)	(0.00279)	(0.00275)	(0.00301)	(0.00372)
Productivity	-0.00600***	-0.00570***	-0.00642***	-0.00407**	-0.00865***
	(0.00126)	(0.00126)	(0.00168)	(0.00182)	(0.00217)
Capital intensity	-0.00123*	-0.00134**	-0.00094	-0.00161**	-0.00051
	(0.00063)	(0.00063)	(0.00067)	(0.00070)	(0.00084)
Vertical	-0.01750***	-0.01705***	-0.00903***	-0.01375***	-0.00289
	(0.00257)	(0.00257)	(0.00221)	(0.00240)	(0.00516)
Distance	0.00152***	0.00241***	0.00145***	0.00143***	0.00152**
	(0.00043)	(0.00050)	(0.00042)	(0.00054)	(0.00075)
Parent characteristics					
Size			0.00776***	0.00899***	0.00609***
			(0.00141)	(0.00131)	(0.00220)
Age			-0.00022***	-0.00023***	-0.00021**
			(0.00005)	(0.00005)	(0.00009)
Wage			-0.00400*	-0.00043	-0.00750**
			(0.00238)	(0.00272)	(0.00307)
Productivity			0.00580***	0.00509***	0.00649***
			(0.00134)	(0.00175)	(0.00125)
Capital intensity			-0.00259***	-0.00295***	-0.00207***
			(0.00030)	(0.00057)	(0.00057)
Listed parent			0.03775***	0.02339	0.07170***
			(0.01412)	(0.01653)	(0.02677)
Multi-firm		-0.00348	-0.00210	-0.00433	0.00007
		(0.00247)	(0.00221)	(0.00327)	(0.00268)
Multi-national		-0.01618***	-0.02194***	-0.02406***	-0.01840***
		(0.00382)	(0.00345)	(0.00486)	(0.00495)
Diversification			-0.00562**	-0.00049	-0.01067**
			(0.00272)	(0.00517)	(0.00522)
Industry characteristics					
Entry costs	-0.67668***	-0.67776***	-0.63649***	-0.21194	-1.22361***
	(0.21183)	(0.21181)	(0.21669)	(0.21911)	(0.34630)
Country, Industry, Year controls	Y	Y	Y	Y	Y
Observations	69,418	69,418	69,418	37,915	31,413

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Looking at the status of the parent separately, we find little evidence that being either a multinational or owning more than one affiliate is detrimental to affiliate survival, even if we don't control for parent characteristics. To the contrary, multinational parents as well as the diversified ones, exhibit a positive effect of affiliate survival rate. This might be due to smaller multinational parents in our data, compared to the large multinationals in Bernard and Jensen (2007)'s sample, indicating that parents active in more than one industry or present in several countries can better diversify their exposure to shocks. Moreover, we do not find evidence that having more than one affiliate is beneficial or detrimental to affiliate survival.

In columns (4) and (5) we split our sample into pre- and crisis periods. It is worth noting that while most of the coefficients do not change significantly, the level of diversification of the parent does seem to improve affiliate survival especially during 2008-2010 period.

5.3.2 Affiliate exit and relative financial development

Next we look at the financial landscape that parents and affiliates are active in, in order to find out whether parents from countries with a more developed financial sector manage to protect their affiliates better than parents coming from countries with worse financial development. We do so for two samples, mainly all countries and excluding Spanish and Dutch affiliates and parents. We decide to exclude Spain due to it's extremely large levels of credit to private sector, the only country to exceed 200% level during the sample period. We also choose to drop the Netherlands since a large share of the credit to private sector in this country is actually due to private mortgage debt and thus does not proxy well firm access to credit. Moreover, since many of our affiliates are domestic, we include country financial indicators for affiliates and parents separately. Table 5.9 present the main findings for all countries, while the results for the sample without Spain and the Netherlands is shown in Table 5.10.

Table 5.9: Affiliate exit and financial constraints

	(1) 2004-2010	(2) 2004-2010	(3) 2004-2007	(4) 2004-2007	(5) 2008-2010	(6) 2008-2010
Affiliate characteristics						
Size	-0.00927***	-0.00921***	-0.01079***	-0.01074***	-0.00791***	-0.00789***
	(0.00235)	(0.00234)	(0.00245)	(0.00245)	(0.00195)	(0.00196)
Age	-0.00011***	-0.00011**	-0.00015***	-0.00015***	-0.00005	-0.00005
	(0.00004)	(0.00004)	(0.00005)	(0.00005)	(0.00007)	(0.00008)
Wage	0.00280	0.00258	0.00047	0.00050	0.00410	0.00421
	(0.00500)	(0.00491)	(0.00893)	(0.00895)	(0.00374)	(0.00376)
Productivity	-0.00643***	-0.00624***	-0.00410	-0.00402	-0.00863***	-0.00863***
	(0.00191)	(0.00186)	(0.00344)	(0.00343)	(0.00138)	(0.00137)
Capital intensity	-0.00093	-0.00098	-0.00160***	-0.00160***	-0.00049	-0.00050
	(0.00117)	(0.00116)	(0.00060)	(0.00060)	(0.00207)	(0.00206)
Vertical	-0.00904***	-0.00898***	-0.01383***	-0.01373***	-0.00306***	-0.00287***
	(0.00273)	(0.00266)	(0.00346)	(0.00341)	(0.00109)	(0.00109)
Distance	0.00145***	0.00144***	0.00141**	0.00143**	0.00149***	0.00150***
	(0.00054)	(0.00053)	(0.00069)	(0.00069)	(0.00042)	(0.00041)
Parent characteristics						
Size	0.00776***	0.00777***	0.00899***	0.00899***	0.00607***	0.00609***
	(0.00064)	(0.00063)	(0.00106)	(0.00106)	(0.00101)	(0.00101)
Age	-0.00022**	-0.00022**	-0.00023***	-0.00023***	-0.00021	-0.00021
	(0.00010)	(0.00010)	(0.00008)	(0.00008)	(0.00014)	(0.00014)
Wage	-0.00399***	-0.00407***	-0.00023	-0.00043	-0.00714***	-0.00755***
	(0.00127)	(0.00145)	(0.00150)	(0.00177)	(0.00144)	(0.00184)
Productivity	0.00580***	0.00584***	0.00510***	0.00507***	0.00640***	0.00653***
	(0.00117)	(0.00117)	(0.00141)	(0.00137)	(0.00192)	(0.00193)
Capital intensity	-0.00259***	-0.00260***	-0.00293***	-0.00295***	-0.00204***	-0.00207***
	(0.00045)	(0.00046)	(0.00063)	(0.00064)	(0.00053)	(0.00053)
Listed parent	0.03772***	0.03810***	0.02308	0.02371	0.07124**	0.07135**
	(0.01445)	(0.01433)	(0.01821)	(0.01815)	(0.03122)	(0.03137)
Multi-firm	-0.00209	-0.00220	-0.00433	-0.00439	0.00010	0.00015
	(0.00274)	(0.00274)	(0.00345)	(0.00348)	(0.00229)	(0.00226)
Multi-national	-0.02195***	-0.02192***	-0.02412***	-0.02406***	-0.01866***	-0.01831***
	(0.00350)	(0.00343)	(0.00309)	(0.00323)	(0.00477)	(0.00525)
Diversification	-0.00563	-0.00534	-0.00055	-0.00029	-0.01070**	-0.01101**
	(0.00702)	(0.00705)	(0.01006)	(0.01001)	(0.00439)	(0.00445)
Financial development						
Credit ratio affiliate		0.03755**		0.02614		0.10912
		(0.01796)		(0.01623)		(0.11909)
Credit ratio parent	-0.00041		-0.00781		-0.00819	
	(0.00891)		(0.01190)		(0.00895)	
Industry characteristics						
Entry costs	-0.63642*	-0.63605*	-0.20986	-0.22042	-1.22237**	-1.22836**
	(0.33531)	(0.33742)	(0.29668)	(0.29140)	(0.60545)	(0.61224)
Country, Industry, Year controls	Y	Y	Y	Y	Y	Y
Observations	69,418	69,418	37,915	37,915	31,413	31,413

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5.10: Affiliate exit and financial constraints - excluding Spain and the Netherlands

	(1) 2004-2010	(2) 2004-2010	(3) 2004-2007	(4) 2004-2007	(5) 2008-2010	(6) 2008-2010
Affiliate characteristics						
Size	-0.0102***	-0.0101***	-0.0127***	-0.0125***	-0.00763***	-0.00756***
	(0.00299)	(0.00296)	(0.00268)	(0.00273)	(0.00284)	(0.00284)
Age	-8.66e-05**	-8.22e-05**	-0.000149***	-0.000143***	3.43e-06	4.84e-06
	(4.14e-05)	(3.97e-05)	(4.78e-05)	(5.06e-05)	(7.07e-05)	(6.99e-05)
Wage	0.000958	0.00117	-0.00283	-0.00235	0.00418	0.00426
	(0.00593)	(0.00598)	(0.0107)	(0.0108)	(0.00442)	(0.00447)
Productivity	-0.00422**	-0.00428**	-0.000509	-0.000504	-0.00839***	-0.00838***
	(0.00169)	(0.00169)	(0.00264)	(0.00267)	(0.00205)	(0.00204)
Capital intensity	-0.00243***	-0.00247***	-0.00204***	-0.00207***	-0.00303**	-0.00304**
	(0.000536)	(0.000534)	(0.000764)	(0.000775)	(0.00122)	(0.00120)
Vertical	-0.0121***	-0.0117***	-0.0175***	-0.0172***	-0.00388**	-0.00366**
	(0.00178)	(0.00181)	(0.00322)	(0.00325)	(0.00157)	(0.00146)
Distance	0.00123*	0.00120*	0.00133	0.00128	0.00113**	0.00110**
	(0.000643)	(0.000665)	(0.000897)	(0.000929)	(0.000544)	(0.000559)
Parent characteristics						
Size	0.00800***	0.00795***	0.00974***	0.00965***	0.00576***	0.00580***
	(0.000830)	(0.000847)	(0.00183)	(0.00188)	(0.00147)	(0.00145)
Age	-0.000130**	-0.000144**	-0.000159***	-0.000170***	-9.11e-05	-9.74e-05
	(6.41e-05)	(6.32e-05)	(5.35e-05)	(5.52e-05)	(8.33e-05)	(8.31e-05)
Wage	-0.00458**	-0.00644***	-0.00172	-0.00396	-0.00783***	-0.00861***
	(0.00208)	(0.00243)	(0.00215)	(0.00282)	(0.00207)	(0.00236)
Productivity	0.00666***	0.00673***	0.00492***	0.00503**	0.00861***	0.00867***
	(0.00154)	(0.00156)	(0.00191)	(0.00196)	(0.00166)	(0.00167)
Capital intensity	-0.00285***	-0.00280***	-0.00314***	-0.00311***	-0.00243***	-0.00241***
	(0.000657)	(0.000639)	(0.000812)	(0.000810)	(0.000732)	(0.000720)
Listed parent	0.0498***	0.0492***	0.0431**	0.0418**	0.0717**	0.0712**
	(0.0121)	(0.0121)	(0.0175)	(0.0172)	(0.0338)	(0.0338)
Multi-firm	0.000323	0.000447	-0.000776	-0.000500	0.000999	0.00102
	(0.00299)	(0.00307)	(0.00344)	(0.00363)	(0.00308)	(0.00307)
Multi-national	-0.0213***	-0.0228***	-0.0242***	-0.0252***	-0.0179***	-0.0188***
	(0.00369)	(0.00368)	(0.00302)	(0.00382)	(0.00687)	(0.00670)
Diversification	-0.00930	-0.00916	-0.00687	-0.00701	-0.0114**	-0.0114**
	(0.00659)	(0.00671)	(0.0108)	(0.0112)	(0.00480)	(0.00473)
Financial development						
Credit ratio affiliate		-0.101		-0.108		0.0245
		(0.0698)		(0.0768)		(0.0760)
Credit ratio parent	-0.0409**		-0.0567**		-0.0161	
	(0.0181)		(0.0222)		(0.0219)	
Industry characteristics						
Entry costs	-0.981***	-0.980***	-0.527***	-0.529***	-1.836***	-1.844***
	(0.220)	(0.216)	(0.192)	(0.189)	(0.299)	(0.298)
Country, Industry, Year controls	Y	Y	Y	Y	Y	Y
Observations	49,202	49,202	27,555	27,555	21,600	21,600

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

We observe that for the entire sample the coefficient for the parent's financial development is not significant, while greater levels of domestic credit at home increase the probability of exit, primarily in the pre-crisis years. The results without the two aforementioned countries however seem more plausible, as there larger levels of credit do indicate a lower probability of exit, as predicted by the literature. In particular, parent's level of credit is significant, albeit not during the crisis years. This seems to suggest that investment is pro-cyclical and even parents from countries with better access to finance were not able to improve their affiliates' chances of survival during the crisis.

5.4 Robustness checks

In this section we perform a couple of robustness checks. We first consider the impact of the relative size of an affiliate compared to that of the parent. We do so to compare the importance of being large in the absolute versus in the relative terms for survival. Moreover, we also check whether our results are sensitive to the estimation methodology used. Mainly, we re-run our basic estimations using a Linear Probability Model (LPM). Furthermore, we rerun the main specification on foreign affiliates only. We do so due to the belief that foreign firms are generally less constrained than their domestic counterparts (see Buch et al. (2010)). Nonetheless, we would like to see if parents coming from countries with better access to finance improve the chances of survival of their foreign affiliates. Finally, we check if financial constraints differences play a stronger role in those sectors which are classified as dependent on external finance. As the ranking of industries with dependence on external finance is available for manufacturing firms only, we exclude sales affiliates and repeat the analysis on manufacturing affiliates only.

5.4.1 Exit and relative size of affiliate

To determine whether the relative or the absolute size play a role in affiliate survival, we construct an extra variable which represents the relative size of the affiliate to that of the parent. We employ the log of the ratio of employment in both entities, but results hold for a similar analysis when fixed capital assets of parent and affiliate are used.

Table 5.11: Affiliate exit and parent characteristics - the role of relative size

	(1) 2004-2010	(2) 2004-2010	(3) 2004-2010	(4) 2004-2007	(5) 2008-2010
Affiliate characteristics					
Size	-0.00361***	-0.00234**	-0.00184**	-0.00213*	-0.00212
	(0.00101)	(0.00102)	(0.00087)	(0.00115)	(0.00151)
Age	-0.00013*	-0.00014**	-0.00011	-0.00015**	-0.00006
	(0.00007)	(0.00007)	(0.00007)	(0.00007)	(0.00010)
Wage	0.00019	0.00242	0.00310	0.00094	0.00439
	(0.00268)	(0.00275)	(0.00273)	(0.00301)	(0.00370)
Productivity	-0.00616***	-0.00577***	-0.00643***	-0.00406**	-0.00867***
	(0.00123)	(0.00124)	(0.00169)	(0.00184)	(0.00218)
Capital intensity	-0.00117*	-0.00131**	-0.00095	-0.00164**	-0.00051
	(0.00062)	(0.00062)	(0.00066)	(0.00070)	(0.00083)
Vertical	-0.01264***	-0.01154***	-0.00933***	-0.01409***	-0.00313
	(0.00256)	(0.00255)	(0.00225)	(0.00238)	(0.00523)
Distance	0.00044	0.00150***	0.00152***	0.00152***	0.00157**
	(0.00043)	(0.00048)	(0.00042)	(0.00054)	(0.00075)
Parent characteristics					
Relative size affiliate	-0.00607***	-0.00670***	-0.00712***	-0.00823***	-0.00557**
	(0.00073)	(0.00074)	(0.00138)	(0.00130)	(0.00217)
Age			-0.00020***	-0.00021***	-0.00020**
			(0.00005)	(0.00005)	(0.00009)
Wage			-0.00445*	-0.00107	-0.00778**
			(0.00234)	(0.00265)	(0.00305)
Productivity			0.00593***	0.00526***	0.00658***
			(0.00133)	(0.00175)	(0.00125)
Capital intensity			-0.00258***	-0.00289***	-0.00209***
			(0.00029)	(0.00055)	(0.00057)
Listed parent			0.03919***	0.02541	0.07224***
			(0.01411)	(0.01670)	(0.02701)
Multi-firm		-0.00502**	-0.00197	-0.00414	0.00014
		(0.00240)	(0.00223)	(0.00327)	(0.00269)
Multi-national		-0.02049***	-0.02130***	-0.02336***	-0.01785***
		(0.00371)	(0.00339)	(0.00483)	(0.00496)
Diversification			-0.00517*	0.00011	-0.01034**
			(0.00273)	(0.00524)	(0.00517)
Industry characteristics					
Entry costs	-0.66386***	-0.66341***	-0.64166***	-0.21847	-1.22911***
	(0.21056)	(0.21048)	(0.21697)	(0.21998)	(0.34548)
Country, Industry, Year controls	Y	Y	Y	Y	Y
Observations	69,418	69,418	69,418	37,915	31,413

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5.12: Affiliate exit and financial constraints - excluding Spain and the Netherlands - the role of relative size

	(1) 2004-2010	(2) 2004-2010	(3) 2004-2007	(4) 2004-2007	(5) 2008-2010	(6) 2008-2010
Affiliate characteristics						
Size	-0.00261	-0.00253	-0.00336	-0.00319	-0.00227	-0.00217
	(0.00311)	(0.00310)	(0.00391)	(0.00398)	(0.00185)	(0.00187)
Age	-8.64e-05**	-8.20e-05**	-0.000147***	-0.000141***	1.61e-06	3.04e-06
	(4.08e-05)	(3.90e-05)	(4.70e-05)	(4.99e-05)	(7.07e-05)	(6.98e-05)
Wage	0.00129	0.00151	-0.00244	-0.00197	0.00446	0.00455
	(0.00592)	(0.00598)	(0.0108)	(0.0109)	(0.00446)	(0.00451)
Productivity	-0.00418**	-0.00425**	-0.000441	-0.000438	-0.00839***	-0.00837***
	(0.00168)	(0.00168)	(0.00264)	(0.00267)	(0.00204)	(0.00203)
Capital intensity	-0.00243***	-0.00248***	-0.00206***	-0.00209***	-0.00302**	-0.00303**
	(0.000529)	(0.000529)	(0.000780)	(0.000790)	(0.00120)	(0.00118)
Vertical	-0.0125***	-0.0121***	-0.0179***	-0.0177***	-0.00426***	-0.00403***
	(0.00176)	(0.00181)	(0.00327)	(0.00330)	(0.00148)	(0.00139)
Distance	0.00132**	0.00129*	0.00144	0.00140	0.00120**	0.00117**
	(0.000653)	(0.000676)	(0.000900)	(0.000931)	(0.000553)	(0.000568)
Parent characteristics						
Relative size affiliate	-0.00724***	-0.00720***	-0.00888***	-0.00880***	-0.00511***	-0.00515***
	(0.000849)	(0.000869)	(0.00188)	(0.00192)	(0.00142)	(0.00140)
Age	-0.000114*	-0.000128**	-0.000140***	-0.000151***	-7.76e-05	-8.41e-05
	(6.15e-05)	(6.12e-05)	(5.25e-05)	(5.43e-05)	(8.04e-05)	(8.06e-05)
Wage	-0.00508**	-0.00697***	-0.00245	-0.00464*	-0.00817***	-0.00898***
	(0.00200)	(0.00231)	(0.00207)	(0.00266)	(0.00202)	(0.00229)
Productivity	0.00681***	0.00688***	0.00511***	0.00521***	0.00871***	0.00877***
	(0.00152)	(0.00154)	(0.00189)	(0.00194)	(0.00168)	(0.00168)
Capital intensity	-0.00281***	-0.00276***	-0.00303***	-0.00301***	-0.00243***	-0.00241***
	(0.000651)	(0.000634)	(0.000820)	(0.000818)	(0.000724)	(0.000712)
Listed parent	0.0514***	0.0508***	0.0455**	0.0442**	0.0724**	0.0719**
	(0.0124)	(0.0123)	(0.0178)	(0.0175)	(0.0339)	(0.0339)
Multi-firm	0.000390	0.000516	-0.000625	-0.000351	0.000996	0.00102
	(0.00300)	(0.00308)	(0.00345)	(0.00364)	(0.00305)	(0.00304)
Multi-national	-0.0206***	-0.0221***	-0.0236***	-0.0245***	-0.0173**	-0.0183***
	(0.00372)	(0.00373)	(0.00308)	(0.00388)	(0.00685)	(0.00670)
Diversification	-0.00873	-0.00859	-0.00617	-0.00634	-0.0110**	-0.0109**
	(0.00662)	(0.00675)	(0.0108)	(0.0113)	(0.00485)	(0.00478)
Financial development						
Credit ratio affiliate		-0.104		-0.109		0.0238
		(0.0684)		(0.0751)		(0.0756)
Credit ratio parent	-0.0415**		-0.0556**		-0.0168	
	(0.0177)		(0.0220)		(0.0218)	
Industry characteristics						
Entry costs	-0.986***	-0.984***	-0.531***	-0.532***	-1.843***	-1.852***
	(0.221)	(0.217)	(0.190)	(0.188)	(0.298)	(0.297)
Country, Industry, Year controls	Y	Y	Y	Y	Y	Y
Observations	49,202	49,202	27,555	27,555	21,600	21,600

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Given that this ratio is highly correlated with employment levels of the parent, we replace this variable by the ratio. Table 5.11 presents the results and suggests that while the relatively larger affiliates have a higher chance of survival, the absolute size is still important, with bigger affiliates having lower exit rates. However, the absolute size advantage becomes insignificant during the crisis, potentially due to the severity of the economic downturn experienced by firms. When considering the financial development in countries of parents and affiliates in Table 5.12, while the relative size remains important, the absolute size is not significant any longer. This leads us to conclude that size in itself does not impact survival. However, compared to the size of the parent, larger affiliates are more likely to survive than smaller ones.

5.4.2 Predicting exit with LPM

Next we also consider whether using a Linear Probability Model impacts our results. We present the main estimation with parent firm-level characteristics as well as one on country financial development. The results in Table 5.13 confirm the findings from our Probit model. Mainly, we again see that larger and more capital intensive affiliates have a lower probability of exit. Moreover, as predicted by the literature, more productive firms will be able to survive longer, as they are using their inputs more efficiently. Vertical affiliates have a lower exit probability as well, probably due to their importance in the production chain of the parent. Moreover, we confirm that distance is positively associated with exit, as MNEs will close further away locations that more difficult to monitor than their closer by counterparts.

In terms of parent characteristics, we again find that larger, more productive and listed parents are more likely to shut down their affiliates, while the multinational status, age and capital intensity of the parent is actually improving survival. As mentioned earlier, this is probably due to the nature of our multinational firms. While very large listed companies are able to buy and sell affiliates at a fast pace depending on the business model they chose to follow and keeping their own productivity as priority, smaller parents are less able to do so, potentially suffering as a result.

In Table 5.14 we repeat the analysis by adding the financial development of parents and affiliates.

Table 5.13: Affiliate exit and parent characteristics - LPM

	(1) 2004-2010	(2) 2004-2010	(3) 2004-2010	(4) 2004-2007	(5) 2008-2010
Affiliate characteristics					
Size	-0.00777***	-0.00711***	-0.00917***	-0.01035***	-0.00803***
	(0.00085)	(0.00086)	(0.00149)	(0.00161)	(0.00165)
Age	-0.00012*	-0.00012**	-0.00009	-0.00013*	-0.00004
	(0.00006)	(0.00006)	(0.00006)	(0.00007)	(0.00009)
Wage	0.00306	0.00510	0.00258	0.00084	0.00346
	(0.00305)	(0.00311)	(0.00311)	(0.00337)	(0.00404)
Productivity	-0.00619***	-0.00585***	-0.00633***	-0.00387**	-0.00837***
	(0.00130)	(0.00130)	(0.00176)	(0.00186)	(0.00215)
Capital intensity	-0.00127*	-0.00138**	-0.00100	-0.00170*	-0.00056
	(0.00066)	(0.00066)	(0.00075)	(0.00083)	(0.00090)
Vertical	-0.01719***	-0.01683***	-0.00919***	-0.01427***	-0.00288
	(0.00255)	(0.00256)	(0.00224)	(0.00278)	(0.00531)
Distance	0.00154***	0.00246***	0.00151***	0.00147**	0.00164*
	(0.00044)	(0.00052)	(0.00045)	(0.00062)	(0.00081)
Parent characteristics					
Size			0.00752***	0.00850***	0.00602***
			(0.00136)	(0.00135)	(0.00215)
Age			-0.00020***	-0.00020***	-0.00020**
			(0.00004)	(0.00005)	(0.00008)
Wage			-0.00329	-0.00003	-0.00675**
			(0.00243)	(0.00256)	(0.00317)
Productivity			0.00503***	0.00425**	0.00582***
			(0.00124)	(0.00176)	(0.00124)
Capital intensity			-0.00222***	-0.00235***	-0.00188***
			(0.00028)	(0.00047)	(0.00058)
Listed parent			0.03127***	0.02230	0.05047***
			(0.01112)	(0.01527)	(0.01702)
Multi-firm		-0.00366	-0.00216	-0.00376	-0.00053
		(0.00250)	(0.00220)	(0.00329)	(0.00284)
Multi-national		-0.01648***	-0.02367***	-0.02701***	-0.01942***
		(0.00416)	(0.00416)	(0.00628)	(0.00555)
Diversification			-0.00505*	-0.00054	-0.00910**
			(0.00248)	(0.00510)	(0.00412)
Industry characteristics					
Entry costs	-0.66270***	-0.66083***	-0.64888***	-0.26081	-1.17969***
	(0.20994)	(0.20991)	(0.22201)	(0.21994)	(0.36079)
Country, Industry, Year controls	Y	Y	Y	Y	Y
Observations	69,418	69,418	69,418	37,915	31,413

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5.14: Affiliate exit and financial constraints - excluding Spain and the Netherlands - LPM

	(1) 2004-2010	(2) 2004-2010	(3) 2004-2007	(4) 2004-2007	(5) 2008-2010	(6) 2008-2010
Affiliate characteristics						
Size	-0.0102**	-0.00998**	-0.0124***	-0.0121***	-0.00778**	-0.00769**
	(0.00346)	(0.00345)	(0.00294)	(0.00300)	(0.00348)	(0.00348)
Age	-6.72e-05**	-6.32e-05*	-0.000122**	-0.000118**	1.08e-05	1.36e-05
	(3.03e-05)	(3.00e-05)	(4.48e-05)	(4.77e-05)	(7.01e-05)	(6.93e-05)
Wage	0.000339	0.000708	-0.00247	-0.00203	0.00297	0.00311
	(0.00680)	(0.00684)	(0.0123)	(0.0125)	(0.00496)	(0.00498)
Productivity	-0.00386*	-0.00391*	-0.000350	-0.000319	-0.00752***	-0.00750***
	(0.00191)	(0.00190)	(0.00284)	(0.00288)	(0.00204)	(0.00202)
Capital intensity	-0.00256***	-0.00261***	-0.00210***	-0.00212***	-0.00327*	-0.00330*
	(0.000681)	(0.000677)	(0.000688)	(0.000694)	(0.00160)	(0.00159)
Vertical	-0.0125***	-0.0121***	-0.0183***	-0.0181***	-0.00406**	-0.00382**
	(0.00236)	(0.00238)	(0.00406)	(0.00405)	(0.00164)	(0.00156)
Distance	0.00128*	0.00124	0.00139	0.00137	0.00124*	0.00120*
	(0.000691)	(0.000713)	(0.000925)	(0.000951)	(0.000586)	(0.000602)
Parent characteristics						
Size	0.00783***	0.00776***	0.00932***	0.00923***	0.00571**	0.00573**
	(0.000780)	(0.000783)	(0.00126)	(0.00129)	(0.00200)	(0.00198)
Age	-0.000123*	-0.000137**	-0.000143**	-0.000156**	-8.75e-05	-9.53e-05
	(6.12e-05)	(6.00e-05)	(5.64e-05)	(5.68e-05)	(7.57e-05)	(7.57e-05)
Wage	-0.00391**	-0.00586**	-0.00144	-0.00330	-0.00698***	-0.00799***
	(0.00178)	(0.00215)	(0.00183)	(0.00262)	(0.00205)	(0.00231)
Productivity	0.00598***	0.00609***	0.00440**	0.00449**	0.00787***	0.00794***
	(0.00141)	(0.00145)	(0.00187)	(0.00184)	(0.00211)	(0.00210)
Capital intensity	-0.00243***	-0.00240***	-0.00261***	-0.00258***	-0.00211**	-0.00210**
	(0.000593)	(0.000567)	(0.000708)	(0.000704)	(0.000774)	(0.000765)
Listed parent	0.0404***	0.0404***	0.0365**	0.0364**	0.0505*	0.0503*
	(0.0111)	(0.0108)	(0.0171)	(0.0168)	(0.0255)	(0.0254)
Multi-firm	-1.12e-05	5.50e-05	-0.000567	-0.000496	0.000501	0.000498
	(0.00290)	(0.00300)	(0.00324)	(0.00339)	(0.00320)	(0.00322)
Multi-national	-0.0233***	-0.0247***	-0.0270***	-0.0279***	-0.0199**	-0.0208**
	(0.00507)	(0.00488)	(0.00423)	(0.00491)	(0.00852)	(0.00835)
Diversification	-0.00828	-0.00815	-0.00595	-0.00583	-0.0103**	-0.0101**
	(0.00522)	(0.00542)	(0.00890)	(0.00954)	(0.00426)	(0.00430)
Financial development						
Credit ratio affiliate		-0.0898		-0.0564		0.0196
		(0.0609)		(0.0616)		(0.0800)
Credit ratio parent	-0.0439**		-0.0506*		-0.0207	
	(0.0162)		(0.0243)		(0.0222)	
Industry characteristics						
Entry costs	-0.972***	-0.972***	-0.559***	-0.554***	-1.662***	-1.675***
	(0.195)	(0.192)	(0.188)	(0.186)	(0.284)	(0.278)
Country, Industry, Year controls	Y	Y	Y	Y	Y	Y
Observations	49,202	49,202	27,555	27,555	21,600	21,600

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5.14 confirms our previous findings that parents from countries with easier access to credit decrease affiliate exit probability. However, as before, this relationship breaks down in the crisis years, suggesting that investment is pro-cyclical and that parents with access to more capital are not able to shield their affiliates from closing down.

5.4.3 Exit of foreign firms

Recent literature has suggested that both exporters and firms engaging in FDI experience less financial constraints than firms operating in a single market (see e.g. Wagner (2014)). This is however due to less constrained firms self-selecting into either exporting or engagement in multinational activity. In this subsection we test whether indeed foreign affiliates are less dependent on the credit environment of their countries as well as the importance of credit constraints in the country of the parent for probability of exit. Table 5.15 shows results for all firms but also for the sample excluding Spain and the Netherlands. The results suggest that there is little impact of country level credit on affiliate exit. This is consistent with the literature as far as foreign affiliates do not face similar credit constraints as their domestic counterparts. Moreover, the credit conditions in the parent's country also plays no significant role in foreign affiliate exit.

5.4.4 Exit of foreign manufacturing firms

Whilst in the previous example we estimated survival of both manufacturing as well as sales affiliates, here we zoom in on the manufacturing firms only. Compared to the entire sample, manufacturing affiliates might have tighter constraints due to their heavier reliance on capital for production. Therefore, in this subsection we further explore the impact of financial linkages on foreign affiliate survival in manufacturing sectors alone. We do so by interacting the credit variables with the dependence of industries on external finance, as first employed by Rajan and Zingales (1998). We expect that firms in industries relying on external finance are able to benefit more from looser credit environment in their own countries as well as the countries of their parents. Table 5.16 shows the results for the entire sample, while in Table 5.17 we again exclude Spain and The Netherlands.

Table 5.15: Foreign affiliate exit and financial constraints - with and without Spain and the Netherlands

	all			excluding ES and NL		
	2004-2010	2004-2007	2008-2010	2004-2010	2004-2007	2008-2010
Affiliate characteristics						
Size	-0.00919***	-0.01224***	-0.00674**	-0.00902***	-0.01348***	-0.00665
	(0.00229)	(0.00235)	(0.00344)	(0.00326)	(0.00226)	(0.00458)
Age	-0.00004	-0.00006	-0.00003	0.00005	-0.00002	0.00008
	(0.00017)	(0.00010)	(0.00034)	(0.00029)	(0.00012)	(0.00045)
Wage	-0.00432	0.00213	-0.00618	0.00355	0.01186	-0.00165
	(0.00664)	(0.00857)	(0.00857)	(0.01010)	(0.01120)	(0.01256)
Productivity	-0.01001***	-0.00447	-0.01652***	-0.01096***	-0.00904**	-0.01341**
	(0.00182)	(0.00317)	(0.00420)	(0.00230)	(0.00396)	(0.00611)
Capital Intensity	-0.00248**	-0.00299	-0.00285	-0.00393**	-0.00229	-0.00507*
	(0.00117)	(0.00220)	(0.00181)	(0.00200)	(0.00311)	(0.00301)
Vertical	0.00277	-0.00632	0.01000	-0.00152	-0.01232	0.00303
	(0.00529)	(0.00498)	(0.01037)	(0.00757)	(0.00930)	(0.01160)
Distance	-0.00357	-0.00555	-0.00173	-0.00311	-0.00356	-0.00397
	(0.00532)	(0.00619)	(0.00685)	(0.00530)	(0.00629)	(0.00607)
Parent characteristics						
Size	0.00449***	0.00793***	0.00178	0.00288***	0.00514*	0.00130
	(0.00081)	(0.00208)	(0.00194)	(0.00103)	(0.00268)	(0.00221)
Age	-0.00022***	-0.00004	-0.00049***	-0.00026***	-0.00005	-0.00056***
	(0.00006)	(0.00007)	(0.00011)	(0.00007)	(0.00007)	(0.00015)
Wage	0.00811	0.01335**	0.00073	0.00975**	0.01785**	0.00279
	(0.00619)	(0.00531)	(0.01140)	(0.00423)	(0.00757)	(0.00579)
Productivity	0.00343	0.00148	0.00493	0.00158	0.00456	-0.00196
	(0.00236)	(0.00297)	(0.00642)	(0.00123)	(0.00312)	(0.00432)
Capital Intensity	-0.00228	-0.00031	-0.00401*	-0.00186	0.00030	-0.00349
	(0.00168)	(0.00283)	(0.00221)	(0.00201)	(0.00263)	(0.00289)
Listed parent	0.01937**	0.01352	0.01482	0.02152*	0.02864	0.00877
	(0.00947)	(0.01259)	(0.03195)	(0.01140)	(0.01782)	(0.03606)
Multi-firm	-0.01122***	-0.01112*	-0.00973*	-0.01127*	-0.01616**	-0.00639
	(0.00424)	(0.00607)	(0.00587)	(0.00647)	(0.00673)	(0.01239)
Diversification	-0.00708	-0.00846	-0.00465	0.00117	-0.00415	0.00912
	(0.00804)	(0.01540)	(0.00945)	(0.01066)	(0.01827)	(0.01168)
Financial development						
Credit ratio affiliate	-0.42899	0.39586	-1.30596***	-0.06811	-0.07112	-0.15305
	(0.50264)	(0.48658)	(0.48535)	(0.06409)	(0.09815)	(0.18813)
Credit ratio parent	0.01292	-0.00637	0.04742	0.01555	0.00248	0.03553
	(0.02166)	(0.02338)	(0.17855)	(0.01794)	(0.02795)	(0.02751)
Industry characteristics						
Entry cost	0.00532	-0.00279	0.01049	-0.06225	0.38731	-0.81732
	(0.00357)	(0.01373)	(0.00890)	(0.66535)	(0.85584)	(0.87070)
Country, Industry, Year controls	Y	Y	Y	Y	Y	Y
Observations	7,622	3,593	3,800	4,563	1,990	2,365

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5.16: Foreign manufacturing affiliate exit and financial constraints

	2004-2010	2004-2010	2004-2007	2004-2007	2008-2010	2008-2010
Affiliate characteristics						
Size	-0.00782***	-0.00773***	-0.01097**	-0.01214***	-0.00616	-0.00578
	(0.00208)	(0.00206)	(0.00447)	(0.00431)	(0.00447)	(0.00480)
Age	-0.00011	-0.00011	-0.00031	-0.00033	-0.00006	-0.00004
	(0.00019)	(0.00018)	(0.00026)	(0.00027)	(0.00032)	(0.00033)
Wage	-0.00804	-0.00776	-0.01671	-0.01696	-0.00101	-0.00019
	(0.01170)	(0.01168)	(0.01659)	(0.01648)	(0.01306)	(0.01248)
Productivity	-0.00249	-0.00246	-0.00035	-0.00185	-0.00306	-0.00235
	(0.00394)	(0.00393)	(0.00785)	(0.00798)	(0.00775)	(0.00789)
Capital intensity	-0.00362	-0.00373	-0.00452	-0.00346	-0.00464	-0.00564*
	(0.00223)	(0.00234)	(0.00444)	(0.00458)	(0.00306)	(0.00324)
Vertical	-0.01392	-0.01416	-0.02346***	-0.02317***	-0.00695	-0.00831
	(0.00951)	(0.00953)	(0.00776)	(0.00767)	(0.01837)	(0.01805)
Distance	-0.00857*	-0.00856*	-0.00828	-0.00969	-0.00953	-0.00884
	(0.00496)	(0.00500)	(0.00674)	(0.00615)	(0.00678)	(0.00694)
Affiliate characteristics						
Size	0.00522**	0.00516**	0.00825***	0.00855***	0.00221	0.00200
	(0.00228)	(0.00230)	(0.00236)	(0.00218)	(0.00572)	(0.00566)
Age	-0.00037**	-0.00037**	-0.00039***	-0.00040***	-0.00038*	-0.00038*
	(0.00014)	(0.00015)	(0.00014)	(0.00014)	(0.00020)	(0.00019)
Wage	0.02610***	0.02600***	0.02515***	0.02525***	0.02622**	0.02650**
	(0.00392)	(0.00395)	(0.00763)	(0.00778)	(0.01209)	(0.01156)
Productivity	-0.00305*	-0.00317*	-0.00240	-0.00318	-0.00359	-0.00369
	(0.00172)	(0.00169)	(0.00415)	(0.00392)	(0.00445)	(0.00458)
Capital intensity	-0.00388	-0.00380	-0.00063	-0.00092	-0.00596	-0.00553
	(0.00293)	(0.00295)	(0.00398)	(0.00382)	(0.00411)	(0.00430)
Listed parent	0.02038	0.02044	0.02507	0.02587	0.01123	0.01146
	(0.01931)	(0.01941)	(0.03701)	(0.03762)	(0.04444)	(0.04258)
Multi-firm	0.00120	0.00129	-0.00733	-0.01129	0.00872	0.00806
	(0.01010)	(0.01018)	(0.02097)	(0.02036)	(0.00811)	(0.00840)
Diversification	-0.01665	-0.01668	0.00031	0.00733	-0.02674	-0.02456
	(0.01847)	(0.01852)	(0.02576)	(0.02606)	(0.01912)	(0.01981)
Financial development						
Credit ratio affiliate	-0.06078**	-0.05988**	0.00296	-0.01764	0.07149	0.08704
	(0.02832)	(0.02861)	(0.01983)	(0.01760)	(0.19805)	(0.20310)
Credit ratio parent	0.00977	0.00924	0.00392	-0.00001	0.01145	0.01517
	(0.00780)	(0.00871)	(0.01550)	(0.01720)	(0.01170)	(0.01399)
Credit ratio affiliate*		-0.00630		0.03078		-0.03449***
External finance dependence affiliate		(0.00696)		(0.02145)		(0.01286)
Credit ratio parent*	-0.00103		-0.04915**		0.02881	
External finance dependence affiliate	(0.01399)		(0.02074)		(0.02051)	
Industry characteristics						
Entry costs	0.23193	0.24205	0.66597	0.55288	-0.46308	-0.54490
	(0.54501)	(0.53554)	(0.78995)	(0.72491)	(0.45144)	(0.46259)
Country, Industry, Year controls	Y	Y	Y	Y	Y	Y
Observations	3,133	3,133	1,418	1,418	1,548	1,548

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5.17: Foreign manufacturing affiliate exit and financial constraints - excluding Spain and the Netherlands

	2004-2010	2004-2010	2004-2007	2004-2007	2008-2010	2008-2010
Affiliate characteristics						
Size	-0.00741**	-0.00751**	-0.01613***	-0.01645***	-0.00615	-0.00618
	(0.00356)	(0.00357)	(0.00465)	(0.00481)	(0.00840)	(0.00840)
Age	-0.00021	-0.00021	-0.00029	-0.00039	-0.00021	-0.00021
	(0.00015)	(0.00013)	(0.00055)	(0.00060)	(0.00043)	(0.00042)
Wage	0.00976	0.00963	-0.01063	-0.01459	0.02198	0.02158
	(0.01988)	(0.02010)	(0.03269)	(0.03391)	(0.02469)	(0.02451)
Productivity	-0.00620	-0.00625	-0.00527	-0.00515	-0.00918	-0.00907
	(0.00579)	(0.00582)	(0.01604)	(0.01639)	(0.01192)	(0.01188)
Capital intensity	-0.00257	-0.00248	-0.00263	-0.00263	0.00008	0.00011
	(0.00424)	(0.00414)	(0.00967)	(0.00990)	(0.00516)	(0.00512)
Vertical	-0.01619	-0.01602	-0.03476***	-0.03689***	-0.01122	-0.01084
	(0.01312)	(0.01347)	(0.01002)	(0.01105)	(0.02007)	(0.02007)
Distance	-0.01281**	-0.01299**	-0.00276	-0.00529	-0.02049**	-0.02034**
	(0.00619)	(0.00550)	(0.00534)	(0.00498)	(0.00848)	(0.00849)
Parent characteristics						
Size	0.00261	0.00259	0.00990***	0.00962***	-0.00221	-0.00218
	(0.00284)	(0.00269)	(0.00227)	(0.00200)	(0.00640)	(0.00636)
Age	-0.00033*	-0.00033*	-0.00037**	-0.00035**	-0.00034	-0.00033
	(0.00018)	(0.00018)	(0.00018)	(0.00018)	(0.00023)	(0.00024)
Wage	0.02299***	0.02316***	0.04173***	0.04448***	0.00875	0.00864
	(0.00777)	(0.00814)	(0.01208)	(0.01431)	(0.01130)	(0.01109)
Productivity	-0.00237	-0.00241	-0.00299	-0.00324	-0.00287	-0.00290
	(0.00396)	(0.00392)	(0.00550)	(0.00565)	(0.00648)	(0.00656)
Capital intensity	-0.00288	-0.00286	-0.00028	-0.00093	-0.00548	-0.00545
	(0.00314)	(0.00313)	(0.00437)	(0.00451)	(0.00419)	(0.00432)
Listed parent	0.03583***	0.03558***	0.10396***	0.09774***	0.01157	0.01048
	(0.01311)	(0.01326)	(0.03361)	(0.03199)	(0.04782)	(0.04659)
Multi-firm	-0.00490	-0.00512	-0.03450***	-0.03644***	0.01187	0.01174
	(0.00990)	(0.01005)	(0.00968)	(0.01033)	(0.01057)	(0.01046)
Diversification	-0.00171	-0.00198	0.05882***	0.05524***	-0.02262	-0.02247
	(0.01548)	(0.01543)	(0.01957)	(0.01924)	(0.01921)	(0.01913)
Financial development						
Credit ratio affiliate	-0.11411	-0.11390	0.02591	0.00624	-0.08213	-0.08388
	(0.11283)	(0.11395)	(0.07990)	(0.08341)	(0.20156)	(0.20463)
Credit ratio parent	0.02229	0.02272	-0.15972***	-0.16182***	0.10648***	0.10520***
	(0.02867)	(0.02840)	(0.04412)	(0.05037)	(0.03688)	(0.03648)
Credit ratio affiliate*		0.01661		-0.03932		0.01321
External finance dependence affiliate		(0.03627)		(0.05697)		(0.04445)
Credit ratio parent*	-0.03056		-0.20934***		0.04082	
External finance dependence affiliate	(0.09029)		(0.04789)		(0.15015)	
Industry characteristics						
Entry costs	-0.12104	-0.16089	0.31517	0.17868	-0.49126	-0.51728
	(0.77628)	(0.71953)	(0.91414)	(0.88283)	(0.57416)	(0.52136)
Country, Industry, Year controls	Y	Y	Y	Y	Y	Y
Observations	1,970	1,970	788	788	1,045	1,045

The table present estimation results in form of marginal effects. Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

We find that for the entire sample the domestic credit abundance variable matters for affiliate survival. Moreover, the interaction of domestic credit with financial dependence of the industry of affiliate is significant during the crisis, indicating that affiliates had to rely on home credit for their financial needs in order to improve survival. Nonetheless, this is not the case for the credit ease in the country of the parent. To make sure that the results are not driven by the inclusion of Spain or the Netherlands, we repeat the analysis without these two countries.

Looking at the sample without Spain and the Netherlands we observe that in this case only the credit ease in the parent's country matters, especially in pre-crisis years. Moreover, the interaction of parent financial development with affiliate's industry dependence on external finance is also significant and improves survival prior to 2008. However, we find that affiliates of parents from financially developed countries have had higher exit rates after 2008, suggesting a pro-cyclical behaviour of parents. We therefore cannot confirm that parents that have more access to credit were able to finance their affiliates throughout the crisis, therefore improving their chances of survival. If anything, there was a negative impact at work, with affiliates being closed down precisely from parents that are located in countries with higher levels of credit to consumers.

5.5 Conclusion

This paper examines the role of parent characteristics in predicting affiliate exit. First we look at whether multinational and multi-firm parents are more 'footloose' than their domestic and single-firm counterparts, taking into account traditional characteristics such as size, age, productivity and capital intensity of both the parent and the affiliate. Moreover, we test whether listed and industrially diversified parents impact affiliate survival. We find that older, more productive and more capital intensive affiliates are less likely to exit. Moreover, due to stronger relationship with the parent, vertical affiliates are also less likely to be shut down. However, even controlling for similar characteristics of the parent, we do not find any evidence that multinational and multi-firm parents are more 'footloose'. On the contrary, MNEs decrease affiliate exit due to potentially being better diversified against shocks. Moreover, parents with affiliates active in different industries are also better for affiliate survival especially during 2008-2010 period, again a sign of the benefits of diversification.

Conducting multinational activity requires access to capital, with financial linkages between parents and their affiliates playing a potentially important role. In a second step we check whether parents coming from relatively more financially developed countries can influence affiliate exit by providing assistance if need be. We test this by using the level of private credit as percentage of country GDP. We find that indeed, parents from countries with larger private credit ratios reduce probability of exit in a significant way. This effects is however primarily visible prior to 2008. Looking at foreign affiliates only, we confirm that these are not financially constrained and do not rely on the level of credit in own countries or that of their parent. Zooming in on foreign manufacturing affiliates, our results indicate that credit ratio's matter in the country of the parent, but again only prior to the crisis. We also find some evidence that firms in industries more dependent on external finance are more sensitive to the financial development of their parent's country. This suggests that investment is pro-cyclical and that while in good times more investment will result in lower exit rates, in times of crisis MNEs are not able to transmit the ease of getting credit at home to their affiliates abroad by improving their survival chances.

6

Carrying the (Paper) Burden: A Portfolio View of Systemic Risk and Optimal Bank Size*

6.1 Introduction

As a result of the financial crisis, the health and safety of the financial system is at the heart of many policy agendas. Concerns regarding the financial system tend to relate mostly to commercial banks and their parent holding companies. Policy discussions focus either on the riskiness of individual financial institutions, or on what is broadly termed systemic risk. Regarding individual banks, the key question debated is whether some banks are too big: too big to fail, too big with respect to their country's GDP (Bertay et al., 2013; Demirgüç-Kunt and Huizinga, 2013), too big to produce at minimum average costs (see e.g. Wheelock and Wilson, 2012; Hughes and Mester, 2013), or even too big to rescue. Systemic risk discussions are much broader, and may consider the stability of the financial system itself, the macro effects of a shock to that system, or the

*This chapter is based on joint work with Jaap W.B. Bos (Maastricht University) and Martien Lamers (University of Groningen).

optimal supervisory setup for dealing with and minimizing the likelihood of such a shock.

In this chapter, we combine these two discussions and investigate whether the size of the largest banks in the system has contributed to an increase in systemic risk. We do so by engaging the reader in a thought experiment. We imagine a bank supervisor as an investor holding a portfolio of banks. Each bank aims to maximize profits, but thereby incurs a certain amount of risk. Given that banks' profits are not all perfectly correlated, the risk-return relationship of the portfolio that the supervisor holds is expected to be better than that of the riskiest banks in the system on their own. Taking the long-term view, the bank supervisor not only wants to minimize risk but is certainly also interested in return, as high charter values may boost the stability of individual banks.

Although we consider our view of the bank supervisor a thought experiment, recent events have shown that its experimental nature is closer to the reality of a crisis than one may at first suspect. In theory, the bank supervisor mainly represents the interests of deposit holders and deposit insurance guarantees those interests to a large extent. However, during the recent crisis, most supervisors went above and beyond that objective. In the U.S., the Trouble Asset Relief Program (TARP) initially provided support in terms of bank equity share purchases valued at more than three times the total amounts of deposits in the system, although much of these funds were later reclaimed as shares were sold in the market. Moreover, many assets were purchased well above their actual value, resulting in an implicit subsidy of the banking sector (Office of the Special Inspector General for the Troubled Asset Relief Program, 2013). Finally, the Safe, Accountable, Fair & Efficient (SAFE) Banking Act proposed in 2012 gives regulators additional powers to limit bank size in order to lower systemic risk.

Nevertheless, unlike the typical investor, the bank supervisor is seriously limited in buying and selling assets in order to reach or remain at the optimal frontier as depicted in Figure 6.1. As the crisis has shown, even this highly constrained investor can rebalance the weights of the banks in the portfolio, through orderly liquidation and other interventions by the Financial Stability Board such as the capital surcharge for Systemically Important Financial Institutions (SIFIs).

Using this scenario, we pose three questions, each related to the situation depicted in

Figure 6.1: The supervisory view: Markowitz efficient frontier, systemic risk and bank size

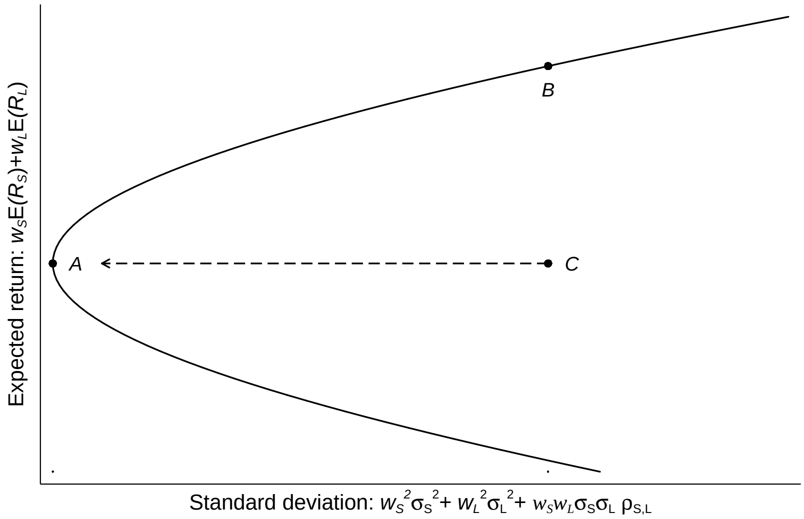


Figure 6.1. First, we ask whether large banks offer attractive investment opportunities for the bank supervisor or in other words, whether large banks are characterized by a risk-return relationship superior to that of the other banks in the system. This should establish whether the inclusion of large banks has brought the supervisory portfolio closer to the efficient frontier, e.g., moving from C to B in Figure 6.1.

Second, we examine what would happen to the portfolio's return if the bank supervisor held the minimum variance portfolio. In light of the example in Figure 6.1, would the return move from B to A, thereby requiring the supervisor to give up return in order to hold a less risky portfolio, or would it move from C to A, allowing the choice of a less risky portfolio without sacrificing returns?

Third, we examine whether the bank supervisor should reduce investments in large banks in order to achieve the minimum variance portfolio. In light of the example in Figure 6.1, we ask whether the supervisor has to increase or reduce the share of large banks (s_L) in the portfolio by moving to A. We also examine whether the differences in risk and return between the original portfolio held by the bank supervisor and the new minimum variance portfolio merely reflect a change in the weights of banks, or whether they are driven by high correlation of the returns of the largest banks in the original portfolio.

To perform our experiment, we examine developments in the U.S. banking market since 1984. Using quarterly data on banks' assets and profits, we construct two types of bank supervisory portfolios. For each Federal Reserve District, we include all unconsolidated commercial banks located within the district. For the U.S. as a whole, we construct a portfolio comprising all Bank Holding Companies. In both cases, we are interested in the design of the minimum variance portfolio and how it compares with the actual portfolio.

Our findings indicate that the current portfolios are not located on the efficient frontier, as risk can be reduced without sacrificing return in order to attain the minimum variance portfolio. Moreover, we find that the largest banks in the Federal Reserve portfolios consistently have a significantly lower weight in the hypothetical minimum variance portfolio. In addition, the minimum variance portfolio does not allow for large levels of concentration in the first place, with its weights being much more evenly spread across banks. These results hold even after relaxing some assumptions and allowing the correlation structure to change with the size of banks. In obtaining the minimum variance portfolio, we assume that the supervisor is able to alter the relative weights of the individual banks in the portfolio, along the lines of the Safe, Accountable, Fair & Efficient (SAFE) Banking Act proposed in 2012. The Act was designed to make banks small enough to fail without causing global panic, using regulatory caps. We argue that our results provide important insights into the optimal design of a portfolio of banks, held by a risk-averse supervisor who prefers to incur the least possible amount of risk. The findings suggest reducing the size of the largest banks in the financial system may not only make individual banks safer and easier to fail or rescue, but can also contribute to a reduction in the riskiness of the system as a whole.

This chapter proceeds as follows. In Section 6.2 we discuss our methodology and data, followed by Section 6.3 which contains our results. Extensions and policy implications are addressed in Section 6.4, after which we conclude.

6.2 Data and methodology

6.2.1 Methodology

We regard banks as assets having both a return and a risk component. By placing the banks in this risk-return framework, we analyze not only the profitability of the banks but also the risk inherent in profit maximization. When considering a portfolio of these assets and assuming that returns are not perfectly correlated, a supervisor holding this portfolio can diversify and enjoy a better risk-return trade-off. The diversification opportunities of the supervisor are constrained by the variety of banks available. Of course, banks themselves also individually diversify their loan portfolios, business lines or geographical markets served. However, the dimensions along which banks can diversify are correlated with their size, as larger banks can be assumed to have greater geographical reach and a different mix of activities compared with smaller banks. Moreover, if enough banks diversify along the same lines, the system as a whole becomes more susceptible to common shocks, even though individual banks themselves seem safer from a micro-prudential point of view. This observation is not new, being noted for instance by Rajan (2005), Wagner (2008), Acharya (2009), Ibragimov et al. (2011) and Allen et al. (2012), who also draw on Modern Portfolio Theory (Markowitz, 1952, 1959) and the insight that portfolio risk can be reduced by choosing assets that are not perfectly correlated with each other.

We contribute to this line of thinking by applying portfolio theory to the banking system in order to investigate the role of large banks in determining systemic (portfolio) risk. The supervisor in our case is able to change the portfolio's risk-return trade-off by altering the size of banks with respect to the system. Although this ability goes beyond the existing mechanisms in place (such as 'Cease and Desist' orders and other 'Prompt Corrective Actions'), we argue that this thought experiment can, at the very least, give us insights into the optimal design of a portfolio of banks. Applying portfolio theory to the banking system is not however all that straightforward, due to some of its strict assumptions. First, the return distribution is assumed to be fully characterized by the first two moments and disregards any tail dependence, even though financial market returns are found to be skewed and fat-tailed. To (partially) mitigate this issue, we use lower-frequency quarterly returns. Second, market participants are assumed to have no influence on prices and return structures of assets in their investment universe, regardless of the weight they are given. In reality, if a bank supervisor were to reduce the size of a bank, its risk-return trade-off would be bound to change as well, as would its correlation structure. Third, an investor is assumed to be able to purchase assets in parcels of any size, meaning that bank sizes could fluctuate heavily between the investor's decision moments. Under a more realistic scenario, the supervisor would be able to change the size of a bank in a limited way, e.g. by only a certain percentage of the bank's assets. Although we initially proceed under the strict setup, the last two assumptions are relaxed at a later stage.

In order to apply portfolio theory and build the regulator's portfolio, we first need to define the return and weight of the assets under consideration. Previous studies have relied mainly on market-based measures when assessing systemic risk (see e.g. De Jonghe, 2010; Adrian and Brunnermeier, 2011; Acharya, Engle and Richardson, 2012; Acharya, Pedersen, Philippon and Richardson, 2012; Bisias et al., 2012; Brownlees and Engle, 2012; Engle et al., 2014). Unlike these studies, we instead use the return-on-assets from book data for the returns of the banks. We do so for several reasons, the first being that the aggregated risk concerning the supervisor is not based on the returns and risks of the shareholders of banks, but rather on those of the (productive) assets that they hold. In the event that the regulator has to bail out a bank, saving or guaranteeing its liabilities will be equivalent to saving or guaranteeing its assets. Second, as shown by Allen and Carletti (2008), in financial crises market prices tend to reflect the amount of available liquidity instead of future earnings. Since these episodes are of particular interest to this analysis, market-based measures might not be appropriate as they could capture liquidity risk instead of systemic risk. Third, accounting data enable us to explore a more extensive sample since market data is only available for a small subset of banks. While listed banks do account for a large percentage of the total banking assets, small banks are potentially a source of (liquidity) contagion through the interbank market (see e.g. Furfine, 1999, 2003; van Lelyveld and Liedorp, 2006; Degryse and Nguyen, 2007). Finally, return-on-assets is a cleaner measure of the underlying profitability, as return-on-equity incorporates management choices with regards to leverage. While our baseline results are based on the book value of the return-on-assets, they are robust to using market-based measures.

The weight that the regulator holds in each bank is calculated as the bank-level total assets divided by the sum of all bank-level total assets available in the portfolio.

According to portfolio theory, the investor's return and risk are calculated as:

$$r_{p,t} = \mathbf{w}_t' \boldsymbol{\mu}_t \tag{6.1}$$

$$\sigma_{p,t}^2 = \mathbf{w}_t' \boldsymbol{\Sigma}_t \mathbf{w}_t \tag{6.2}$$

where $\mathbf{w}_t$ is a column vector representing the weights of all banks in period t and $\boldsymbol{\mu}_t$ represents the expected return of the banks, usually defined as the average return of the previous quarters. Furthermore, $\boldsymbol{\Sigma}_t$ represents the expected covariance matrix of these returns and is often replaced by its sample equivalent. The average expected return of the portfolio is given by $r_{p,t}$, while the variance of this set of returns is given by $\sigma^2_{p,t}$ and represent the measure of portfolio (or in our case: systemic) risk at time t.[1]

The supervisor is considered to be risk averse, and to prefer to hold the portfolio with the least amount of risk according to the objective function in Equation 6.2. Moreover, the supervisor is able to influence systemic risk by changing the weights in the portfolio, assuming that this does not impact the matrix $\boldsymbol{\Sigma}_t$.[2] Minimizing the objective function allows us to compare the differences in portfolio design between the initially realized and the hypothetical minimum variance portfolio.

To achieve the minimum variance portfolio (MVP), the supervisor solves for:

$$\begin{aligned} \underset{\mathbf{w}_t}{\operatorname{argmin}} \quad & \mathbf{w}_t'\boldsymbol{\Sigma}_t\mathbf{w}_t \\ \text{s.t.} \quad & \mathbf{w}_{i,t} \geq 0 \\ & \mathbf{1}'\mathbf{w}_t = 1 \\ & \mathbf{w}_t'\boldsymbol{\mu}_t \geq 0 \end{aligned} \tag{6.3}$$

with the addition of several further constraints. First, the supervisor cannot go short in a bank, i.e. no bank can have a negative weight. Second, the weights of banks have to add up to 1 as the existing assets are merely reshuffled, without any being created or destroyed. This is equivalent to assuming that the banks under consideration constitute the entire investment universe of the supervisor. Finally, a supervisor is also assumed to choose the weights such that the portfolio does not have negative returns.[3] It follows from these non-linear constraints that no analytical solution is possible, and

[1]Since the source of risk in a portfolio selection problem is correlation of returns, portfolio risk mainly measures the endogenous build-up of risk due to exposure to a common factor. This follows recent literature which focus on the trade-off between individual diversification and systemic diversification (see, e.g., Wagner, 2008; Acharya, 2009; Ibragimov, Jaffee and Walden, 2011). However, the portfolio risk measure is less suited to capture exogenous shocks to the financial system, since these shocks have to materialize before showing up in the data and can spillover to other parts of the system that are connected through interbank-linkages.)

[2]As the recent crisis has shown, it is not unusual for supervisors to intervene through liquidation, (hidden) bailouts or forcing banks to sell off assets to maintain a competitive environment.

[3]The inclusion of this constraint does impact our results as will be shown later.

we therefore rely on a numerical solution.

Using this approach, we investigate the following questions. First, we ask whether systemic risk can be reduced and if so, by how much. To do this, we compare the standard deviation of the initial portfolio with that of the MVP. Second, we investigate whether the supervisor would have to sacrifice returns in order to achieve a lower risk, by comparing the average return of the initial portfolio with that of the MVP. Within the framework described in the introduction, we therefore ask if, in order to reach the MVP, the supervisor has to move along the efficient frontier or shift towards it. Third, we compare the dispersion of weights within each of the two portfolios, by looking at the share of the largest 5% of banks in the initial portfolio and in the MVP.[4] Finally, we ask whether the largest 5% of banks in the initial portfolio have retained their relative importance in the MVP by comparing their initial share with the weight they receive in the MVP.

6.2.2 Data

We perform our analysis on the U.S. banking system, which has several regulatory bodies at different levels. Depending on location, membership status and type, a bank can be regulated by the Federal Reserve System (FED), the Office of the Comptroller of the Currency (OCC) and the Federal Deposit Insurance Corporation (FDIC). While the main supervisory task at district level is carried out by the 12 Federal Reserve Banks, depicted in Figure 6.2, at the national level the main regulatory task is performed by the Federal Reserve's Board of Governors. As a consequence of this division, we consider the regulatory portfolio both at national level and at district level.

Bank data are obtained on a quarterly basis from the Call Reports for Income and Condition provided by the Federal Reserve System. For the national (FED) portfolio, we consider consolidated Bank Holding Companies (BHCs) as the assets in which the regulator can invest. Data for the BHCs are obtained from the FR Y-9C Forms, between 1986Q3 and 2012Q1. We select only Holding Companies and exclude Insurance/Securities brokers, Utilities and other Non-Depository institutions. At the

[4] An alternative measure of concentration would be the Herfindahl-Hirschman Index, although using the portfolio weights of the largest banks is more intuitive in this setting. The 5% concentration measure is preferred, since it allows for a better comparison in different-sized banking systems (see Alegria and Schaeck, 2009).

Figure 6.2: The Federal Reserve Districts

Federal Reserve District (FRD) level, data on unconsolidated Commercial Banks are retrieved from the FFIEC 031/041 Forms between 1984Q1 and 2010Q4, excluding Savings/Cooperative/Industrial banks as well as Non Deposit Trust companies.

We use balance sheet data instead of financial market data, allowing us to consider all banks that are required to file reports and not only those that are listed on an exchange. Moreover, lower frequency returns are preferred to daily or even weekly returns, in order to comply with the assumption of normality of returns. We collect total assets (bhck/rcfd2170) and net income (bhck/riad4340), deflate both to 2005Q1 dollars using the Producer Price Index provided by the St. Louis Federal Reserve Bank, and filter out banks with return-on-assets exceeding +100% or -100%. This leaves us with 4,694 BHCs across 154,577 bank-year observations and 19,225 commercial banks over 1,132,425 bank-year observations.

Summary statistics for the BHCs and commercial banks are shown in Table 6.1. In any quarter in our sample, there are between 964 and 2,333 holding companies active in the United States. Due to inflation and a wave of consolidation, the total assets reporting threshold for BHCs was raised from \$150 million to \$500 million in 2006, causing a drop in the number of banks in the sample. Banks controlling less than \$500 million in total assets prior to 2006 are kept in the dataset, since, as will be shown at a later stage, their exit does not affect our results. Given that the distributions of assets and returns are highly unequal, we report percentiles instead of means and standard deviations. The median BHC controlled \$500 million in total assets and reported a net income of \$1.3 million. The table shows the skewness in the distribution of total assets, with the largest 5% of BHCs having total assets ranging between \$14.5 billion and \$2.1

trillion. While the net income of the median holding company is $1.3 million, again there is a large disparity: the highest earning 5% of BHCs recorded profits ranging from $44 million to $6.4 trillion. At the other end of the spectrum, losses are equally large, partly due to the recent financial crisis, with one holding company reporting a net loss of $15 trillion in the third quarter of 2008. Since the return-on-assets takes into account the size of the BHC, its values are less extreme compared with those of returns and assets separately, with the mean (0.257%), median (0.292%) and mode (0.325%) lying in close proximity.

The number of reporting commercial banks lies between 6,477 and 14,474 over the 12 districts. Regarding net income and total assets, commercials banks follow a similar pattern to that of Bank Holding Companies, although smaller on average. The total assets disparity is even larger than at the national level, with some banks dwarfing their competitors.

We proceed by placing every bank in its respective FRD portfolio, defining its weight as the total assets of the bank divided by the sum of total assets in the FRD. The BHCs are analyzed at the national level in a similar manner: the weights are calculated as the individual level of total assets divided by the sum of all total assets of the BHCs. As we have quarterly data over a period of 25 years at our disposal, we perform the analysis using a window of 8 consecutive quarters on which we calculate the expected return and sample covariance matrix, thereby taking into account time-varying correlation.[5] As a consequence, assets need to have posted data in each consecutive quarter of the window to be included in the analysis.

6.3 Results

In this section we present the results of our analysis. We first examine the risk-return trade-off between the FED portfolio and the MVP. In a second step, we look at the differences in portfolio allocation between the two systems before comparing their other features. We present the results of the analysis at BHC level graphically, referring the reader to the Online Appendix for the results on FRD level as they are quantitatively similar, and conclude this section with several robustness tests.

[5]In Section 6.3.5 we show that our results are robust to a different window size.

Table 6.1: Summary statistics

Panel A: Bank Holding Companies

		Percentiles					
	Min.	5^{th}	25^{th}	50^{th}	75^{th}	95^{th}	Max.
Net Income (in $ million)	-15132.20	-0.72	0.64	1.30	3.09	44.0	6414.61
Total Assets (in $ million)	5.33	156.32	284.62	500.37	1184.04	14456.24	2115728.50
Return on Assets (in %)	-39.29	-0.14	0.18	0.29	0.39	0.58	82.81

	Mean	Std. Dev.	Min.	Max.
Number of Banks	1502	387	964	2333

Panel B: Commercial Banks

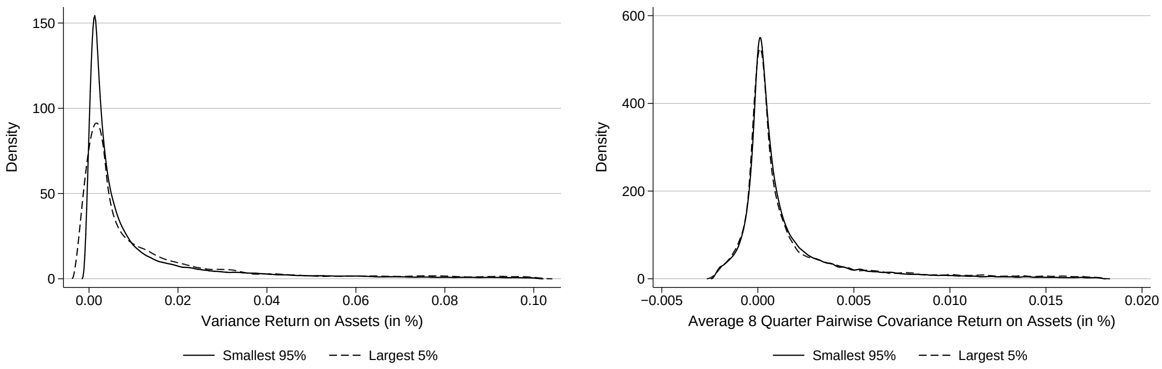
		Percentiles					
	Min.	5^{th}	25^{th}	50^{th}	75^{th}	95^{th}	Max.
Net Income (in million $)	-11168.98	-0.20	0.08	0.24	0.58	2.94	4682.32
Total Assets (in million $)	0.14	15.91	39.48	79.39	171.95	896.35	1594746.30
Return on Assets (in %)	-78.23	-0.34	0.19	0.33	0.46	0.72	90.85

	Mean	Std. Dev.	Min.	Max.
Number of Banks	10115	2655	6477	14474

The table presents summary statistics for Net Income (bhck/riad4340), Total Assets (bhck/rcfd2170) and the Return-on-Assets of the banks in the analysis. Panel A displays these descriptives for Bank Holding Companies between 1986Q3 and 2012Q1, while Panel B summarizes them for Commercial Banks between 1984Q1 and 2010Q4. Due to the highly skewed distributions, we summarize the data according to their percentiles as well as the minimum and maximum values.

Figure 6.3: Are large banks more risky?

(a) Variance **(b)** Covariance



The figure shows density plots of the average 8 quarter variance of the return-on-assets and average pairwise covariances of the smallest 95% and the largest 5% of Bank Holding Companies between 1988Q2 and 2012Q1. Due to long tails for both large and small banks, the variance is truncated at its 90$^{\text{th}}$ percentile as it is strictly positive, while the covariance is truncated at its 5$^{\text{st}}$ and 95$^{\text{th}}$ percentile for graphical purposes.

6.3.1 Are large banks more risky?

However, before reporting the results we first need to establish the similarities in the return structure between large and small banks. Should large banks have (co)variances different from those of small banks, the assumption that the covariance matrix is independent of size means we would impose an unrealistic structure when large banks are reduced in size, or small banks are made larger.[6] In Figure 6.3, we show the two dimensions of the covariance matrix by plotting the densities of the average 8 quarter variance and average 8 quarter pairwise covariance for the largest 5% and the smallest 95% of BHCs. From the Figure, it becomes clear that despite the differences in size, there is ample common support in the individual and common riskiness of bank returns as the distributions overlap almost entirely. These results hold for different time samples, and for both BHCs and Commercial Banks.[7]

Of course, even if large and small banks share a common support in the covariance matrix, it does not mean that a bank that changes size will maintain its return structure. At a later stage, we therefore look at banks that have seen large increases or decreases in

[6] In this respect, our approach is similar to verifying whether the assumption of a 'common support' holds in propensity score matching (Heckman et al., 1998). If there is enough 'common support' between small and large banks, the assumption that the supervisor could change assets without having these actions leading to a different return structure is more realistic.

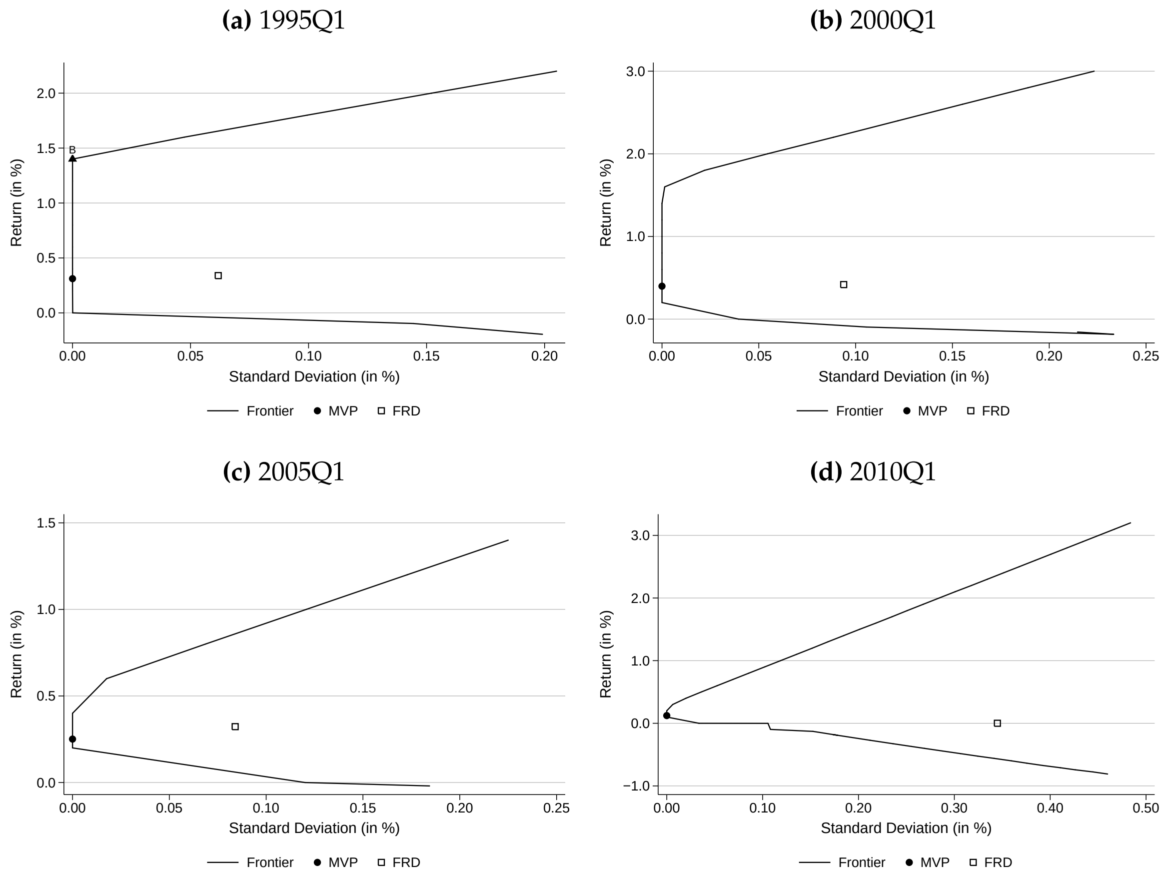
[7] Full results available upon request.

size, and analyze how this changed the elements of the covariance matrix. Using these average changes in turn allows us to alter the covariance matrix during the numerical optimization, leading to a more realistic portfolio allocation. However, since there are only a limited number of cases on which we can base this analysis, we first proceed by assuming that changing a bank's size does not influence the structure or level of its returns, and later revisit this assumption.

6.3.2 What role does inequality play in the risk/return trade-off?

It is quite straightforward to obtain the risk, return and weight distributions for the initial portfolios. By contrast, obtaining the respective MVPs is more cumbersome, as a minimum of 964 and a maximum of 2,333 BHCs are present during the sample period. The Chicago, Kansas City and Dallas portfolios typically contain well over 1,800 banks. The solution is computationally intensive, but is nonetheless obtained after a lengthy optimization process.

Figure 6.4: Efficient frontier - MVP and FRD Boston Portfolio



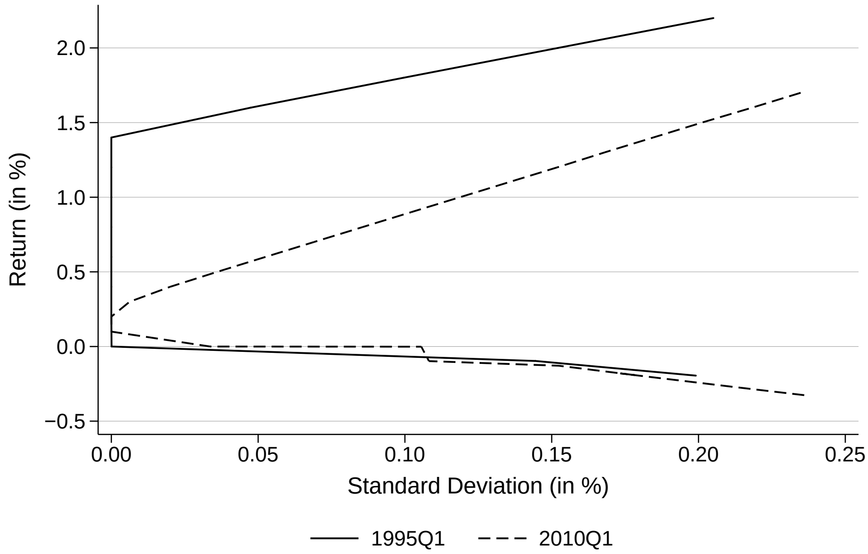
The figure shows the risk-return trade-off for the Federal Reserve District of Boston at 4 time periods. It plots the efficient frontier and places the MVP and FRD portfolios in the risk-return space.

Figures 6.4 and 6.5 show the risk/return trade-off for the Federal Reserve District of Boston in a selected number of years using the 8 quarter rolling window approach. For graphical purposes, we show the trade-off at 4 moments: 1995Q1, 2000Q1, 2005Q1 and 2010Q1. We choose the district of Boston for computational reasons as it has the least number of active banks throughout the sample period.[8]

Figure 6.4 plots the empirical efficient frontier and places the MVP and Boston FRD portfolio's in the mean-variance space. The shape of the frontier reveals that even when few banks are present, the maximization process is non-linear and a consistent frontier is difficult to obtain. We see that in each of the four periods the FRD portfolio is *not* located on the efficient frontier, as the MVPs have similar returns but a lower aggregate risk. Furthermore, the FRD portfolio's seem to be moving away from the MVPs over

[8]Plotting the efficient frontiers entails multiple minimizations while holding the return constant and is therefore computationally more intensive than simply finding the MVP.

Figure 6.5: Boston Portfolio Diversification Possibilities – Comparison 1995Q1 and 2010Q1



The figure plots the efficient frontier for the Federal Reserve District of Boston at the beginning and the end of the sample period.

time: in 1995 the MVP has a return of 0.31% and standard deviation of 0%, while the FRD has a return of 0.33% and a standard deviation of 0.06%. Meanwhile, in 2010 the MVP has a return of 0.12% and its standard deviation remains at 0%, while the FRD portfolio now has a return of essentially 0% and a standard deviation of 0.34%.

Two further remarks are in order here. First, the hump-shaped frontier in the beginning of the sample seems to indicate that, in point *B* in panel 6.4a, a portfolio return of 1.4% is possible with a portfolio standard deviation of essentially 0% compared to the MVP which has similar risk but a lower return. Analyzing the outcomes of the frontier, we find that the risk is nonetheless lowest in the MVP, albeit marginally. Moreover, the weight distributions between these two points are quite similar. The cumulative weight of the largest 5% of banks in point *B* is 12%, while it is 17% in the MVP.

The second remark is that when comparing the steepness in panel 6.4a and panel 6.4d, it seems that diversification possibilities were more plentiful in the beginning of the sample period. This is emphasized in Figure 6.5, where the efficient frontiers from 1995Q1 and 2010Q1 are plotted simultaneously. Figure 6.5 shows that the efficient frontier turns inward, indicating that the risk-return trade-off has worsened over time.

Considering the difficulty in obtaining the frontier, and the amount of time periods where we would have to perform these minimizations, we obtain only the MVP next to the respective FRD or FED portfolio in the rest of the analysis.

Figure 6.6 shows the results of the portfolio optimization based on BHC data: panels 6.6a and 6.6b display the risk-return trade-off for the FED portfolio and the MVP. They show that with different weights, portfolio risk is effectively eliminated in the MVP. Whereas the standard deviation of the FED portfolio spikes during the Savings-and-Loans and subprime mortgage crises, that of the MVP remains stable at around zero. One would expect a portfolio with lower (minimized) risk to have a lower return as well. However, this less risky system has positive returns throughout the sample period with values that closely track the actual returns. Therefore, we conclude that the initial portfolio does not lie on the efficient frontier, as risk is reduced while the level of returns has been maintained. Panel 6.6c shows that the lower risk is achieved in the MVP through a markedly lower concentration than in the FED portfolio. While in reality the weight of the largest banks lies between 65% and 90%, the concentration in the MVP is on average 13% and at most 52%. More interestingly, the largest banks in the FED portfolio see their cumulative weight reduced to at most 15% of total assets in the MVP.[9]

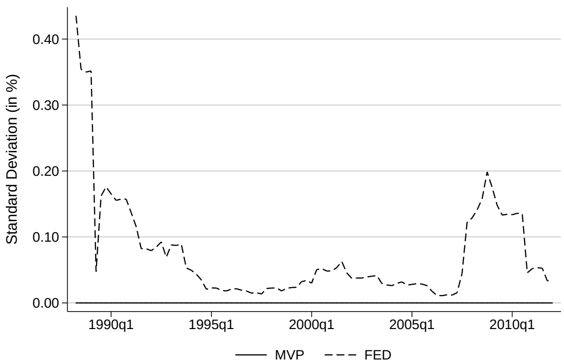
The same picture as for the BHC data emerges if we look at the separate FRD portfolios, reported in the Online Appendix. Table A.1 summarizes the differences in risk and return between the FRD portfolios and their MVPs. Evidence of returns over the whole period is mixed: some FRDs outperform their MVPs, whereas others exhibit lower returns. One interesting fact is that in the boom period of 1994Q1 - 2006Q4, we find that all FRDs outperform their MVP counterparts in terms of returns. Regarding size disparity, Table A.2 shows that the MVPs consistently have a much lower level of concentration compared to their FRD portfolios, the difference ranging on average between 44% and 78% throughout the sample period.

Two further remarks are in order. First, the portfolio standard deviation in Figure 6.6 seems to be higher during the S&L crisis than in the subprime crisis. This result can be explained by the fact that a standard deviation, unlike a correlation coefficient, is not a dimensionless number and can only be interpreted as a function of its mean. Given that the returns of the FED portfolio and the MVP are of similar size throughout the sample,

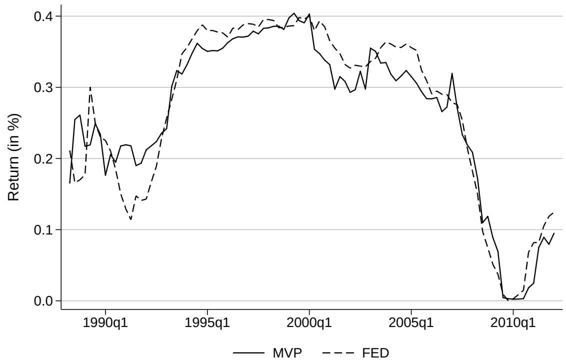
[9] Although not reported in the Figure, the Gini coefficient exhibits the same pattern, as the average coefficient of the FED portfolio is 0.9 compared to 0.2 for the MVP.

Figure 6.6: What role does inequality play in the risk/return trade-off?

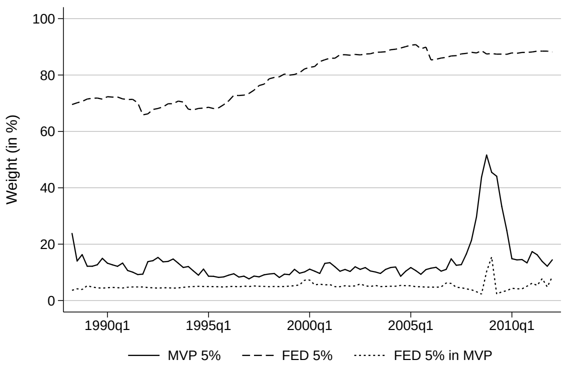
(a) How much is systemic risk reduced in the MVP?



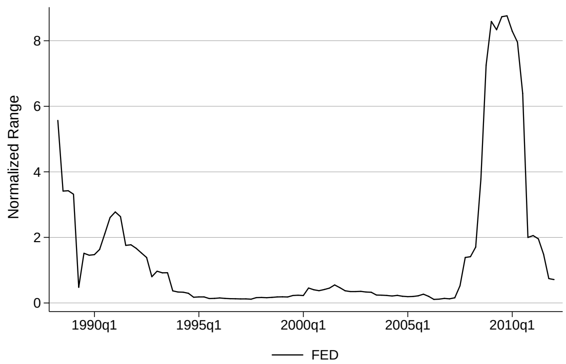
(b) Do returns have to be sacrificed to achieve this lower risk?



(c) How unequal is the MVP compared to the FED portfolio?



(d) How does the systemic risk measure evolve over time?



The figure shows the comparison between the FED portfolio and the hypothetical MVP: panel 6.6a and 6.6b display the difference in the risk and return of each portfolio. Panel 6.6c shows how the weights are distributed in each portfolio by plotting their concentration ratios, as well as the weights that the current largest banks have in the MVP. Panel 6.6d shows the demeaned risk measure from Panel 6.6a to allow for a comparison over time. Instead of using the standard deviation, which is a reflection of the mean, we show the normalized range defined as (max-min)/median and is independent of the mean return.

we can compare their standard deviation in each quarter, but not between quarters. To allow for comparison over time, we de-mean the risk measure, although this leads to a loss in direct interpretation.[10] Figure 6.6d displays the de-meaned risk measure, and portfolio risk now shows a higher peak in the subprime crisis than in the S&L crisis.

The second issue we want to address is the MVP's high concentration during 2008 and 2009. This spike can be explained by the fact that up to 87% of BHCs reported lower average returns than in the previous quarter and 40% of all BHCs recorded losses. Given the number of banks involved, it is possible that concentration rose because of this increase in correlation between average returns. However, it is also possible that the MVP weights are chosen to avoid violating the no-loss constraint. To test the latter possibility, we ran the analysis excluding the no-loss constraint but still find the same spike, indicating that, indeed, higher weights are given to banks that share a lower correlation. Since these are few in number, they therefore have to receive a higher weight in order to minimize portfolio risk.

The results here suggest that inequality and concentration play an important role in the risk-return trade-off with which a regulator is faced. In this simple exercise, reducing inequality drives down risk without significantly affecting returns at both FED and FRD level. These findings indicate that regardless of the regulatory level, supervisors need to be concerned when looking at the optimal design of their portfolio not only with a bank's individual size but also its size relative to the system. Moreover, the rebalancing of weights does not appear to be random. We find that in order to obtain a less risky portfolio, a supervisor has to reduce holdings of the currently largest banks and create a more equal system. In reality, the largest banks have had a much higher share in the portfolio compared with that in the MVP, and even increased their weight from 65% to 90% during the sample period. Moreover, the financial industry as a whole has also grown in relation to GDP, to the extent that Carvalho and Gabaix (2013) attribute the recent rise in macroeconomic volatility mainly to this growth in combination with idiosyncratic shocks to the largest banks. Indeed, the share of the largest banks in the current portfolio relative to GDP has increased from 35% at the beginning to 75% at the end of the sample. Given the evidence presented here,

[10]We do so by calculating the normalized range of returns. Since the portfolio standard deviation is the standard deviation of the weighted average returns during the last 8 quarters, we define the normalized range as the (maximum-minimum)/median of this set.

combined with the finding of Carvalho and Gabaix (2013), large banks do seem to play an increasingly important role not only in the banking system, but also in the broader economy and its volatility.

However, one important caveat needs to be acknowledged. In the methodology section, we assumed that the banks under consideration represent the entire investment universe for the supervisor, mimicking the current regulatory set-up. Therefore, we exclude financial institutions that fall outside of the regulator's jurisdiction, such as for instance investment banks and money market funds that played a big role in the propagation of risk during the 2008 crisis. Similarly, while U.S. subsidiaries of international banks have to file financial statements and are therefore included, their international parents are not considered. While on a global scale their activities can impact the profitability and risk of the banks under consideration, we are only able to observe their indirect effect on the U.S. system. As a consequence, if the supervisor would set his weights according to the minimum variance portfolio, he might force banks to move their activities abroad or into the shadow banking realm where they are unsupervised. Given the data at hand and our assumption that assets are redistributed and not moved outside of the portfolio, we can not consider this outcome. The results presented here and in the following sections therefore correspond only to a *partial* equilibrium.

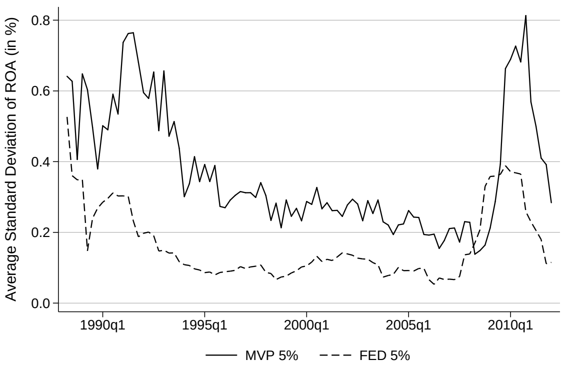
6.3.3 What are the other features of the Minimum Variance Portfolio?

Given that we have seen that lower portfolio risk is achievable in a less concentrated banking system, this raises the question of what causes the largest banks in the FED portfolio to have such consistently low weights in the MVP, and how they differ from the largest banks in the MVP. The two components that determine the weighting decision are on the one hand individual bank risk, as measured by the standard deviation of the returns, and on the other hand the correlation between these returns.[11] Figure 6.7a shows the average standard deviation of return-on-assets for the largest 5% of banks in the FED portfolio and MVP, while Figure 6.7b displays the average pairwise correlation among the largest 5% of banks in their respective portfolios. We see that the standard

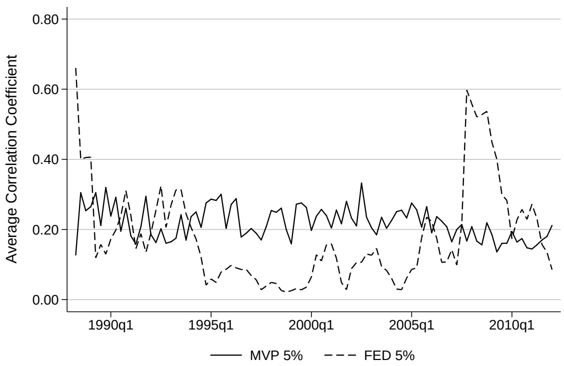
[11] In this section, we show the correlation of returns instead of the covariance as it is easier to interpret.

Figure 6.7: What are the other features of the Minimum Variance Portfolio?

(a) Is there a large share for low-risk banks?



(b) Is there on average less correlation among banks?



The figure shows the average bank level standard deviation of the return-on-assets, as well as the average pairwise correlation between these returns for the largest 5% of banks in both MVP and FED portfolio.

deviations of the returns of these largest banks are on average twice as high in the MVP, yet the average correlation coefficient is much more stable compared to the FED portfolio. In the two crisis periods, the FED portfolio pairwise correlation spikes to average values of 0.6, almost three times larger than the MVP. Nonetheless, even when the individual risk and correlation are lowest in the FED portfolio, its risk is still higher than that of the MVP.

These observations add to the evidence that in this context, weight plays a significant role in determining the level of risk, as it magnifies the effect of increased correlation and individual riskiness. When both components have low values, systemic risk is low, even in a highly concentrated market. However, when they increase in crisis periods, systemic risk increases dramatically if size inequality is high. As already shown in Gabaix (2011), individual shocks to firms have the potential to lead to aggregate volatility when the size distribution of an economy is heavy tailed, something that also holds for the banking system (see e.g. Janicki and Prescott, 2006; Blank et al., 2009). In terms of portfolio theory, the variances will dominate the covariances in crisis periods due to the large disparity in weight. Since the movements in correlation can be extremely volatile and difficult to control or even predict, the best tool for keeping systemic risk low in this context is to limit concentration.

We note, however, that these results do not necessarily imply a cap on the size of banks. On the one hand, a system with only small banks is subject to the Too-Many-To-Fail problem since they could herd, thereby acting as one large entity (Acharya and Yorulmazer, 2007; Claeys and Schoors, 2007; Brown and Dinç, 2011). In this setting, herding would be picked up via an increase in correlation of returns, posing a systemic threat despite a lack in concentration. Our results, however, indicate that the correlations of small banks are relatively stable over the sample period, and would therefore not pose a systemic threat. On the other hand, the system can also be diversified by limiting activities that banks can undertake and/or markets it can serve, provided they operate in their own (uncorrelated) niche. This point was also touched on by Loutskina and Strahan (2011), who found that increased geographic diversification went hand-in-hand with a decline in loan monitoring by lenders prior to the financial crisis.

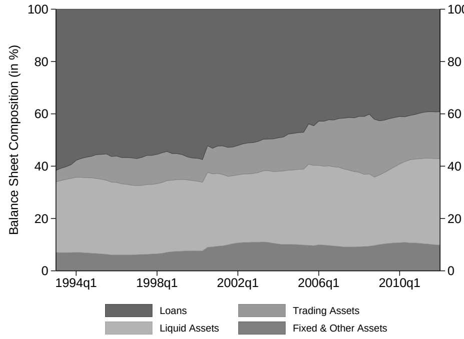
To determine the characteristics of banks which have been heavily reweighted, we construct a crude industry level balance sheet for both the FED portfolio and the MVP. We use the weights allocated to each bank to construct this weighted average balance sheet, which is shown in Figure 6.8. The allocation of assets in the FED portfolio shows the increasing importance of trading assets at the expense of loans, whereas this trend is less evident in banks favored in the MVP. While the FED balance sheet has less than 40% of assets invested in loans at the end of the sample, that of the MVP remains close to 50%. On the liabilities side, the FED portfolio is more reliant on non-deposit funding than the MVP balance sheet. We observe that at the end of the sample, the FED portfolio uses about 10% more of these non-deposit liabilities than the MVP, although this difference was much more apparent before the recent crisis.

As a reflection of the industry asset composition, the non-interest income/total income ratio for each portfolio is shown in Figure 6.8e. We observe that with exception of the crisis years, there has been a significant increase in reliance on non-interest income in the FED portfolio. On the other hand, banks favored in the MVP have a more constant share of non-interest income throughout the sample period. Notwithstanding the financial crisis, the gap between the portfolios has been steadily increasing since 1990.

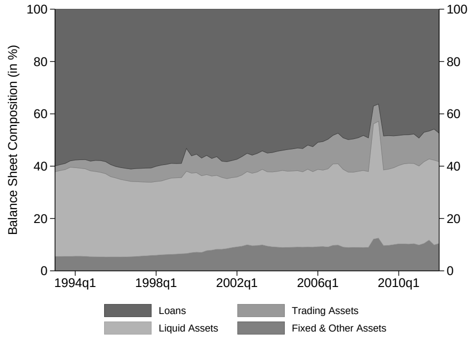
Given these results, we conclude that the fictitious banking industry in the MVP is characterized by retail banking, as higher weights are given to banks that are mainly

Figure 6.8: What happens to the intermediary role of banks in a safer banking system?

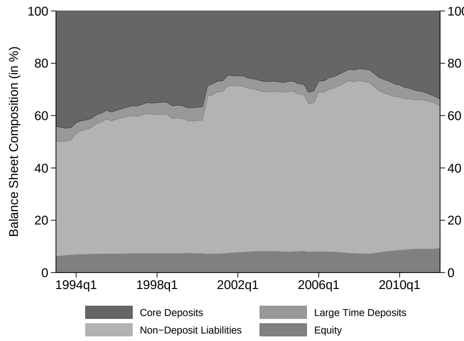
(a) FED - Assets



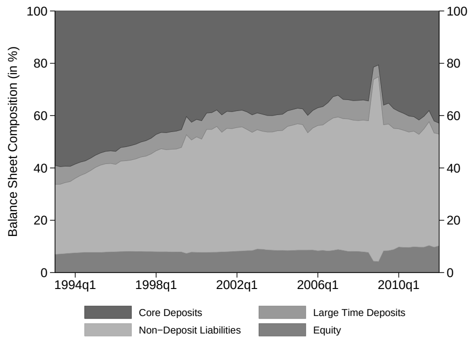
(b) MVP - Assets



(c) FED - Liabilities



(d) MVP - Liabilities



(e) Is there a difference in income characteristics?

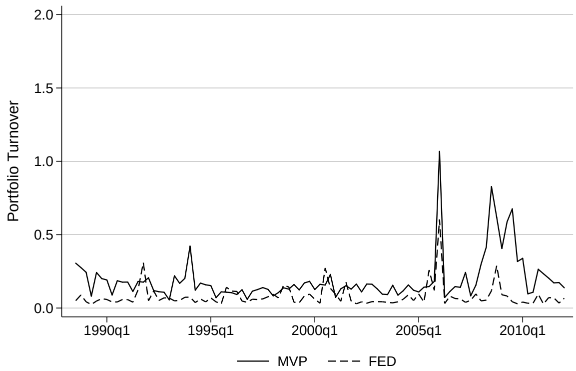
The figure compares the intermediary role of the FED portfolio and MVP, by showing weighted average industry level balance sheets in panels 6.8a-6.8d and the different weighted non-interest income/total income ratios in panel 6.8e. To construct the balance sheet we use the following data series: Loans (BHCK2122); Trading Assets up to 1994 (BHCK2146); Trading Assets after 1994 (BHCK3545); Liquid Assets up to 1994 (BHCK0081 + BHCK0395 + BHCK0397 + BHCK3365 + BHCK0390); Liquid Assets after 1994 (BHCK0081 + BHCK0395 + BHCK0397 + BHCK3365 + BHCK1754 + BHCK1773); Fixed & Other Assets (BHCK2145 + BHCK3163 + BHCK2160); Core Deposits (BHCB3187 + BHCB2389 + BHCB6648 + BHCB2210 + BHOD3187 + BHOD2389 + BHOD6648); Large Time Deposits (BHCB2604 + BHOD2604); Non-Deposit Liabilities (BHCK2948 - Core Deposits - Large Time Deposits); Equity (BHCK2948 + BHCK3210). Non-interest income (BHCK4079); Total income (BHCK4107).

funded by deposits, make loans, and therefore rely less on non-interest income. Since banking activities are not a direct input in the minimization of the portfolio risk, it is their influence on the behavior of the returns which drives these findings. Indeed, this is in agreement with a growing literature emphasizing the role of income diversification in financial instability. For instance, Stiroh (2004) and Stiroh and Rumble (2006) find that non-interest income reduces aggregate profits while increasing risk. More recent evidence by De Jonghe (2010) shows that systemic risk is exacerbated by banks diversifying into activities other than lending, due to increasing correlations between income streams. This finding was also corroborated by Adrian and Brunnermeier (2011), Brunnermeier et al. (2012), DeYoung and Roland (2001) and DeYoung and Torna (2013). Huang and Ratnovski (2011), meanwhile, argue that wholesale lenders have lower incentives for costly monitoring, leading to large (and inefficient) fluctuations of loans on negative public signals, a problem not encountered in relationship banks. Finally, Boot and Ratnovski (2012) find that although there are initial benefits for banks from starting trading activities, beyond a critical point inefficiencies dominate and trading becomes increasingly risky. On the funding side, Fahlenbrach et al. (2012) emphasize that banks with increasing balance sheets through the use of short term non-deposit liabilities performed poorly during the last crises.

6.3.4 How easily is the Minimum Variance Portfolio obtained?

In our baseline scenario, the supervisor is able to switch assets rapidly from one bank to another on a quarterly basis to obtain the MVP. Although reweighting also occurs naturally in the FED portfolio via mergers and acquisitions, bank entry and exit or bailouts, the MVP would not be a realistic approximation if reweighting was much higher than in reality. In order to assess how stable the MVP is over time compared with the FED portfolio, we therefore calculate both of their turnovers. Turnover is defined as the sum of absolute weight changes in the portfolio between period $t-1$ and t, taking values ranging from zero (no change) to two (where all assets that were held are sold, and all assets that were not held bought). Figure 6.9 plots the turnover for both portfolios. We observe that the MVP turnover is on average 3 times as high as that of the FED portfolio. The spike in turnover in 2006Q1 is due to changes in the reporting threshold, as the banks that reported in 2005Q4 were considered to be sold in

Figure 6.9: How much more intervention would be required for the MVP?



The figure shows the difference in reweighting between the FED and the MVP by plotting the turnover of each portfolio.

2006Q1 and proceeds reinvested in the other banks. The MVP can only achieve low risk through a higher level of reweighting, especially in the crisis period. For the district portfolios, the average MVP turnover is around 3 times higher than the actual portfolio, ranging from 2 times for the least concentrated to 6 times for the most concentrated districts. In Section 6.4, we therefore explore alternative MVPs where the reweighting is restricted, in order to achieve a more realistic turnover.

6.3.5 Robustness

To find out how robust our results are, we test several of our assumptions. The results of these tests are summarized in Table 6.3, where we evaluate how well risk was reduced while limiting concentration. To this end, we define the ratio $(\sigma_{\text{FED}} - \sigma_{\text{MVP}})/\sigma_{\text{FED}}$, which measures the relative difference in portfolio risk between the actual and minimum variance portfolio. A score of 1 indicates that the risk has been effectively eliminated, while a score of 0 indicates that no improvements were possible. A test is regarded as successful when this ratio averages 0.9 or higher throughout the sample period, and when the level of concentration of the largest 5% of banks in the MVP is on average below 50%.

Covariance matrix – 1 We begin with the assumption that the covariance matrix is independent of the size of banks. In Section 6.3.1 we showed that large and small banks share a common support in the variance and average pairwise covariance. This however, does not imply that a bank which changes in size will maintain the same level

Table 6.2: Changes in the return and its covariance matrix

		ROA		Variance ROA		Covariance ROA	
		Pre-Jump	Post-Jump	Pre-Jump	Post-Jump	Pre-Jump	Post-Jump
Positive Jump	Value	0.341	0.259	0.023	0.111	0.002	0.005
	Average Difference	-0.082		0.088		0.004	
	Percentage Difference	-23.97%		379.70%		233.70%	
	P-value T-test	0.000		0.007		0.010	
Negative Jump	Value	-0.142	0.213	0.281	3.249	0.007	-0.015
	Average Difference	0.355		2.967		-0.022	
	Percentage Difference	249.52%		1054.64%		-324.52%	
	P-value T-test	0.110		0.359		0.293	

The table shows the average differences in ROA, its variance and its average pairwise covariance pre- and post-jump for both negative jumps and positive jumps using the procedure described in the text and the Online Appendix.

of returns or the same structure with regard to other banks. If we knew how the return structure changes due to a change in size, we would be able to adjust the covariance matrix in each iteration of the optimization. To this end, we have identified 15 cases in which BHCs experience a negative jump in bank size, and 287 where they experience a positive jump in bank size.[12] A jump is defined as an increase/decrease of bank assets of 25% or greater from one quarter to the next, provided that the preceding and following 8 quarters did not show jumps larger than 10% in each of the quarters, nor a cumulative change in the preceding and following 8 quarters of 25%. These last two conditions are imposed to make sure that bank size before and after the jump was relatively stable and that the change in the elements of the covariance matrix can be chiefly attributed to the one-time jump. The banks receiving a negative shock lost 40% of their total assets on average, while banks receiving a positive shock gained 60%. The average changes in the return and covariance matrices are displayed in Table 6.2. T-tests show that banks experiencing a positive jump in assets have a statistically significant lower return-on-assets, which is likely due to the construction of the variable, and a higher average variance and covariance of these returns. On the other hand, banks experiencing a negative jump do not see changes in their average variance and covariance, and only see a marginally significant higher return-on-assets.

Using the statistically significant changes for positive and negative jumps, we incorporate the effect in the sample covariance matrix and the return matrix, such that

[12]Full details on this identification can be found in the Online Appendix.

these matrices change dynamically with the weights of the banks.[13] We perform two additional robustness tests. In the first test, the vector of expected returns is adjusted according to the changes in weight, and a new sample covariance matrix is estimated with which portfolio risk will be minimized. In the second test, we adjust both the return and covariance matrix based on the changes in weight. Changes in the return and covariance matrices are interpolated if the proposed change in portfolio weight lies between -40% and +60%. As there was no data on changes in assets larger than these bounds, we use any proposed changes beyond them as if they were -40% or +60%, i.e. while a bank can receive an increase in weight higher than +60%, its variance and covariance terms are adjusted as if the weight has only been increased by +60%.[14] In both cases, however, this adds more complexity to the optimization and indeed we find that there are cases in which no improvement in portfolio risk is found. Fortunately an improved solution is still possible in most of the sample period, as can be seen in Figure 6.10. Interestingly, the findings in the baseline specification seem to be robust and are not influenced by our original assumption. The concentration in the MVP remains much lower than that of the FED portfolio, while simultaneously maintaining a smaller portfolio risk and returns of a similar level.

Covariance matrix – 2 Another issue in our baseline setup is that optimization can be quite unstable when using the sample covariance matrix: small changes in the return structure can lead to large differences in the outcome of the portfolio choice. To counter these unstable solutions, Brodie et al. (2009) use regularization of the optimal mean-variance portfolio by including a penalty in the objective function. They show that while introducing this penalty can lead to a sparse choice of weights, opting for a high penalty is equivalent to a constraint that does not allow shorting. Since our supervisor is already assumed not to be able to go short in any banks, we therefore consider the optimization problem stabilized. However, as Jobson and Korkie (1980) have pointed out, a considerable amount of noise is introduced in a sample covariance matrix in a small T and large N setting like ours. In response, Ledoit and Wolf (2003, 2004) proposed a shrinkage based estimator of the covariance matrix to reduce this noise. We

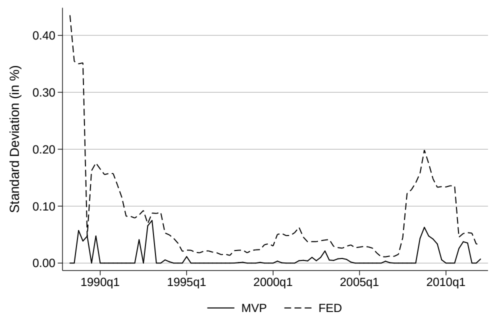
[13] As the change in return-on-assets for the banks experiencing a negative jump is marginally significant at 11%, we also regard this change as significant.

[14] A scenario under which reweighting of each bank was limited to -40% and +60% yielded no improvements over the actual portfolio. Results are available upon request.

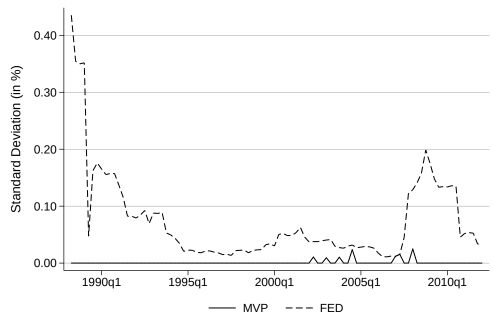
Figure 6.10: Can portfolio risk be minimized when taking into account a changing covariance matrix due to weight changes?

How much is systemic risk reduced in the MVP?

(a) Dynamic Return Matrix

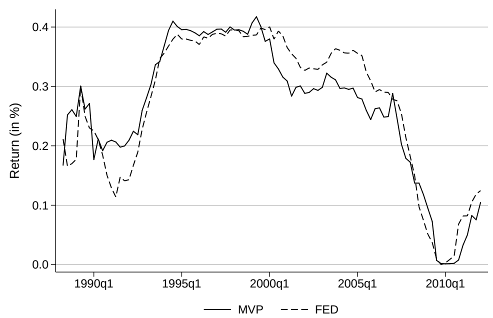


(b) Dynamic Covariance Matrix

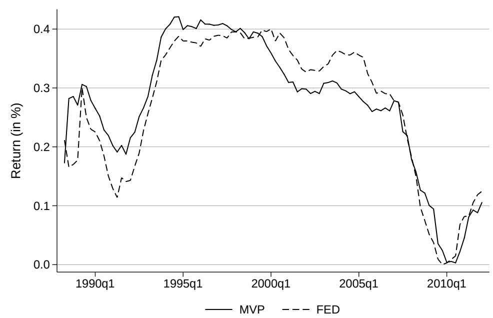


Do returns have to be sacrificed to achieve this lower risk?

(c) Dynamic Return Matrix

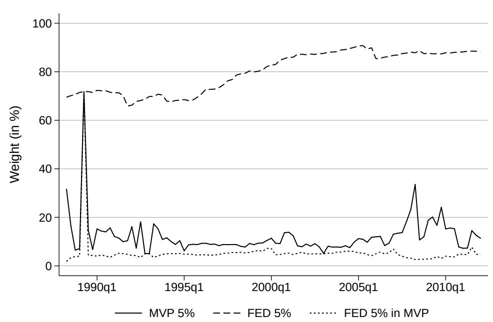


(d) Dynamic Covariance Matrix

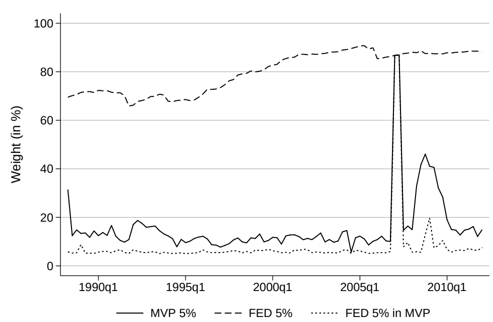


How unequal is the MVP compared to the FED portfolio?

(e) Dynamic Return Matrix



(f) Dynamic Covariance Matrix



The figure presents two robustness tests, in which the return and/or covariance matrices are dynamically updated. In the first robustness test (Dynamic Return Matrix), we adjust only the return matrix and then estimate a new sample covariance matrix. In the second test (Dynamic Covariance Matrix), we adjust both the return matrix and the covariance matrix with which the portfolio risk will be minimized. More information on how we perform these tests can be found in the Online Appendix. Panel 6.10a and 6.10b display the difference in the risk of each portfolio, while panel 6.10c and 6.10d display the difference in return. Panel 6.10e and 6.10f show how the weights are distributed in each portfolio by plotting their concentration ratios, as well as the weights that the current largest banks have in the MVP.

use their proposed estimator of the covariance matrix instead of its sample equivalent and re-run the analysis.[15] The full results are reported in the Online Appendix and are similar to the baseline specification. We find that the standard deviation of the MVP is now higher compared with the baseline specification, although still lower than that of the FED portfolio, while returns are at a similar level. The levels of concentration in the MVP are basically unchanged, as the largest banks on average still have a weight of 10%, while the largest banks in the FED portfolio have a weight in the MVP of 5%.[16]

Starting values Third, we explore the optimization starting values and choice of the length of the moving window. Given the fact that we are dealing with many banks, the minimization of the portfolio risk is likely to be a complex, highly nonlinear problem comprising multiple minima/solutions. The starting values, which are selected as the weights in the original portfolio, can have a substantial impact on whether a global or local minimum is found and in what direction the distribution of weights will move. To account for this possible bias, we run two robustness tests. In the first, we choose starting values based on an equally weighted portfolio. In the second, we run, for the BHC data only, 100 repetitions per quarter using randomized starting values.[17] Both tests show that the results are almost identical to the baseline specification, and we therefore refer the reader to the Online Appendix for the full results.

Length of rolling window Finally, we explore alternative lengths of the rolling window. So far we have taken an 8 quarter time frame to estimate the sample covariance matrix. However, it could be argued that using more data to estimate it would be less noisy and less prone to outliers. Taking this into account, we rerun the analyses using a 16 quarter window and report the full results in the Online Appendix. For both the BHCs and commercial banks, results follow similar patterns to those using 8 quarter windows: the largest banks are still shown to be consistently overweighted compared with their MVP counterparts, where lower risk is achieved while keeping returns at

[15] The code for estimating the covariance matrix is obtained from: `http://www.ledoit.net/honey_abstract.htm`

[16] As portfolio risk is not minimized by at least 90%, however, we do not consider this test to be successful in Table 6.3.

[17] The random starting values are drawn from a half-normal distribution and then divided by its sum, such that they add up to 1.

Table 6.3: Robustness tests

			Risk minimized without large banks?	
	Issue	Robustness test	BHC	COMM
(1)	Independence of covariance matrix	Dynamic return matrix	✓	n.a.
(2)	Independence of covariance matrix	Dynamic covariance matrix	✓	n.a.
(3)	Noise in sample covariance matrix	Shrinkage estimator	✗	0/12
(4)	Multiple minima	Equal starting weights	✓	12/12
(5)	Multiple minima	Randomized starting weights	✓	n.a.
(6)	Length of rolling window	Analysis on 16 quarters	✓	12/12

The table presents a summary of the three robustness tests that were performed, indicating whether portfolio risk was successfully minimized while keeping the levels of concentration of the largest 5% of banks low. To see if a test is successful we define the ratio $(\sigma_{\text{FED}} - \sigma_{\text{MVP}})/\sigma_{\text{FED}}$, which measures the relative difference in portfolio risk between the actual and minimum variance portfolio. A score of 1 indicates that the risk has been effectively eliminated, while a score of 0 indicates that no improvements were possible. A test is regarded as successful when this ratio averages 0.9 or higher throughout the sample period, and when the level of concentration of the largest 5% of banks in the MVP is below 50%. The second, fifth and sixth tests are only performed on BHC data due to their computationally intensive nature.

a comparable level. Table 6.3 summarizes the results of the robustness tests in this section.

6.4 Extensions and Policy implications

As we have seen in Section 6.3.4, the amount of turnover needed to lower systemic risk is three times higher in the MVP. In this section, we therefore explore some more realistic scenarios, and discuss implications for policy resulting from the analysis. We first look into several weighting alternatives. Besides analyzing these other weighting methods, we repeat our analysis on a smaller and more realistic sample of banks. Finally, we discuss whether optimization at district level also results in a lower countrywide systemic risk. Similar to Section 6.3.5, we summarize all results in Table 6.4, where we again define a test successful if the ratio $(\sigma_{\text{FED}} - \sigma_{\text{MVP}})/\sigma_{\text{FED}}$ averages 0.9 or higher throughout the sample period.

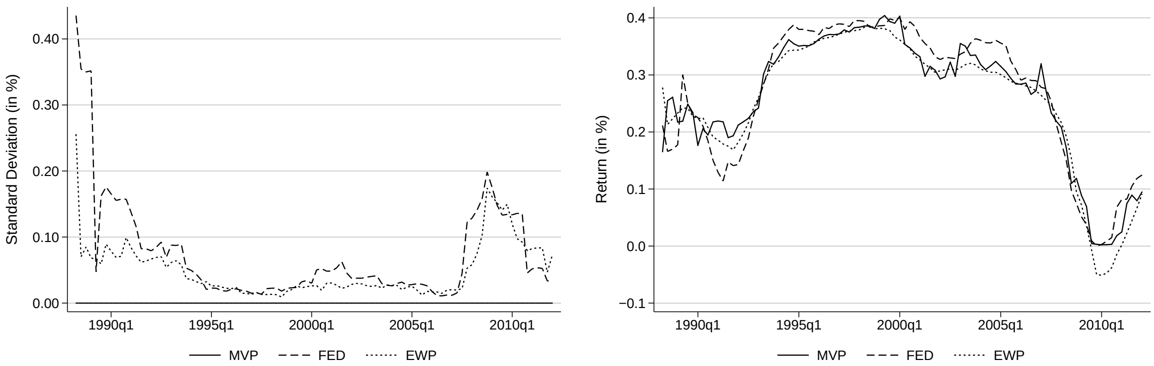
6.4.1 Can portfolio risk be minimized while limiting portfolio turnover?

Equally weighted portfolio Given that the MVP seems to favor a more equal distribution of assets, a natural course of action would be to analyze an equally weighted

Figure 6.11: What is the risk/return profile of an equally weighted portfolio?

(a) Can we reduce systemic risk in an equally weighted portfolio?

(b) Is an equally weighted system more profitable?



The figure presents results of the risk and return in an equally weighted portfolio, compared to the MVP and FED portfolio.

portfolio. As noted by DeMiguel et al. (2009), equally weighted portfolios still outperform many optimizing portfolio choice models and have a very low turnover. The turnover in our setup would indeed be lower than that of the FED portfolio, albeit not zero as bank entry and exit would still take place. Figure 6.11 shows the risk-return trade-off that the equally weighted portfolio (EWP) would have in comparison with the other two. In terms of returns, the EWP performs similarly to the MVP and FED portfolio, except for the last crisis period in which it records losses. Regarding risk, the EWP has levels similar to that of the FED portfolio, albeit marginally lower. All in all, this suggests that there is an optimal level of concentration, as neither a highly concentrated nor an equally weighted portfolio are able to significantly reduce systemic risk in the same way as the MVP.

Limited reweighting – 1 Since a high turnover is costly for the supervisor and therefore not very desirable, we consider several alternatives involving limited reweighting that could reduce turnover. We do this by setting lower and upper boundaries to the weights banks can take, conditional on their true weights. First, we allow banks to grow/shrink by 10% and 20% of their initial weight. Second, we construct a measure of asset growth in the previous quarter and allow changes equal to either the mean or standard deviation of this growth measure. Whereas the first constraint is static in nature, the second allows for business cycle effects to determine how much reweighting can take place. To ensure the no-loss constraint does not influence the results, we run the limited reweighting scenarios with and without this requirement. However, regardless of the specification, the risk in the MVP is practically unchanged compared with the FED portfolio in all time periods.[18]

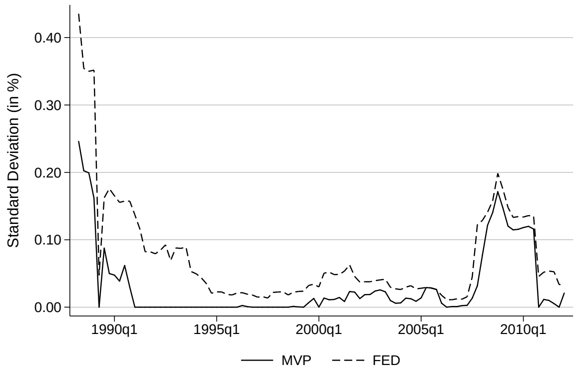
Limited reweighting – 2 Another way of reducing turnover would be to keep the largest 5% of banks at their current cumulative size, allowing unlimited reweighting of the remaining banks while still adhering to the no-shorting and no-loss constraint. The results of this exercise are shown in Figure 6.12, where we see that for most of the sample period it was possible to decrease systemic risk significantly while maintaining large banks. However, during the S&L and subprime crises periods, this MVP variant has a risk which is barely below that of the initial portfolio. This finding can be explained by the high correlation of returns for the largest banks, as seen in Figure 6.7b. During sudden increases in correlation between these largest banks, the high concentration of assets in a few banks will inevitably affect risk. In terms of optimal portfolio design, it seems to point to a trade-off between concentration and correlation: if the supervisor wants to keep the large banks at their current size, it would be necessary to ensure that correlation between them remains relatively low to avoid the Too-Many-To-Fail problem. This could, for instance, be achieved by limiting the geographical markets in which a bank can be active or the activities it can engage in, as used to be the case prior to e.g. the Riegle-Neal and the Gramm-Leach-Bliley acts.

Repeating this analysis on the FRD portfolios shows that there are 3 districts which are able to reduce systemic risk to a minimum in each of the three sub-samples.[19] Interestingly, a common feature is that they have the lowest levels of concentration among all districts. This relationship is shown in Figure 6.13, where we plot the average concentration ratios against the extent to which they are able to reduce systemic risk in three time periods. The horizontal axis represents the amount by which they reduce risk and is again constructed such that 1 stands for a reduction of risk effectively to zero,

[18] Since the standard deviation of asset growth was extraordinarily large in 1997Q4, this allowed the optimization to apply larger changes to the banks and therefore managed to reduce risk. A dynamic approach was also considered for the scenarios where banks are allowed to grow/shrink by 10% and 20% of their weights in the MVP in time $t-1$; however, portfolio risk was not significantly reduced. Full results can be found in the Online Appendix.

[19] Full results are shown in the Online Appendix.

Figure 6.12: How much can we lower systemic risk when we keep the largest banks at their actual size?



The figure displays how well the portfolio risk can be minimized while keeping the largest 5% of banks at their actual size and reweighting the remaining 95%.

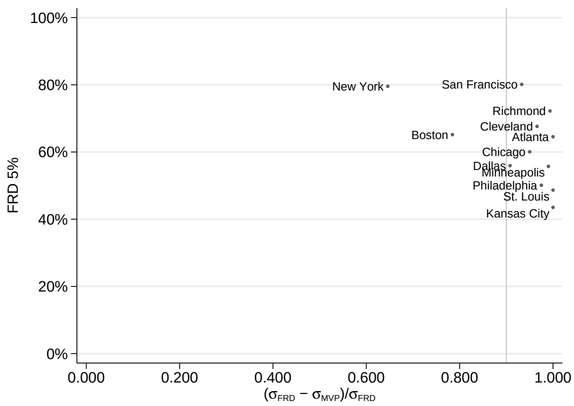
and 0 indicates that no improvements in risk are possible. The Figure clearly shows the trend in consolidation, with most districts becoming more concentrated over time. As they become more concentrated, they find themselves less able to achieve low risk while maintaining their largest banks.

Limited reweighting – 3 Since the weight of the 5% largest banks in the three successful districts never exceeds 60%, this leads us to a final test using limited reweighting. Is it possible to reduce portfolio risk while maintaining the cumulative weight of the largest 5% of banks between 50% and 60%? The results for the FED portfolio are plotted in Figure 6.14. We observe that under this limited reweighting scheme, it is possible to effectively eliminate risk while maintaining similar returns. Concerning the concentration in the portfolio, the share of the largest banks always hits the lower bound of 50%. Consequently, the level of concentration in the MVP also lies close to 50% except for 2008 and 2009, with similar results for the analysis at commercial bank level as shown in the Online Appendix. As in the previous scenarios, the high concentration in the MVP in 2008 and 2009 is likely due to increasing correlation of returns. To minimize portfolio risk, a small number of banks need to receive a higher weight, such that their concentration approaches that of the FED portfolio.

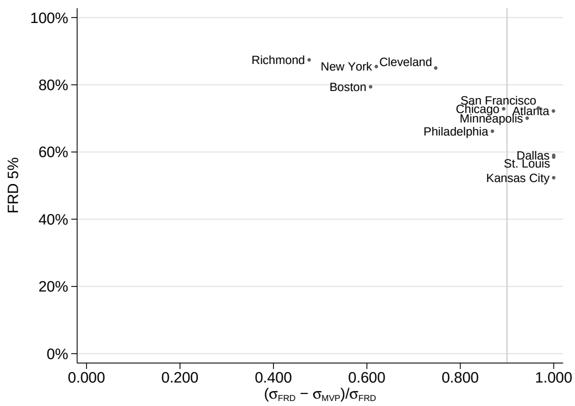
We have shown in Section 6.3 that minimizing systemic risk requires an extremely powerful and active regulator, who would have to intervene three times more than is currently the case. In practice, this could only be achieved by increasing the regulator's

Figure 6.13: Can inequality explain why some FRDs can reduce systemic risk while keeping the largest banks at their actual size?

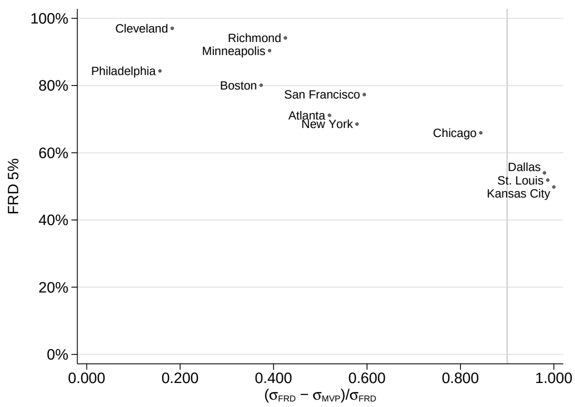
(a) 1984Q4 - 1993Q4



(b) 1994Q1 - 2006Q4



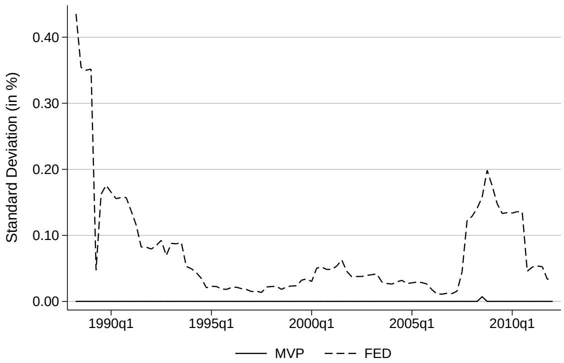
(c) 2007Q1 - 2010Q4



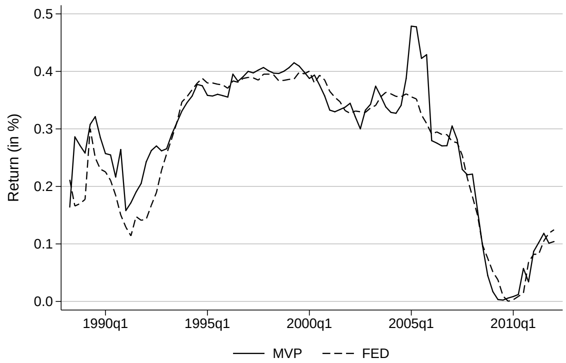
The figure shows the relationship over time between the average weight of the largest 5% of banks and the ability of the FRDs to reduce portfolio risk when the largest banks are allowed to keep their initial weight. $(\sigma_{\text{FRD}} - \sigma_{\text{MVP}})/\sigma_{\text{FRD}}$ measures the relative difference in portfolio risk between the actual and minimum variance portfolio. A score of 1 indicates that the risk has been effectively eliminated, while a score of 0 indicates that no improvements were possible. The threshold for successfully minimizing portfolio risk is set at 0.9. The weights of the largest 5% of banks are averaged over each time period.

Figure 6.14: By how much is the risk minimized when the largest banks hold between 50% and 60% of assets?

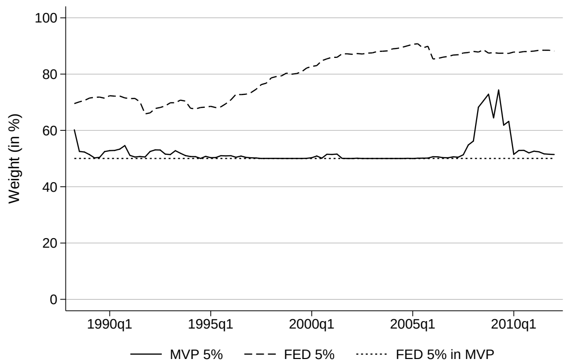
(a) How much is systemic risk reduced in the MVP?



(b) Do returns have to be sacrificed to achieve this lower risk?



(c) How large are the largest banks in the MVP?



The figure shows the risk and return characteristics of an MVP where the largest 5% of banks in the FED portfolio are kept between 50% and 60% of assets, and how the weights in this MVP are distributed.

discretionary power. Such a proactive position was also included in the proposed SAFE Banking Act of 2012, under which a maximum bank size relative to the system would be imposed. Notwithstanding a range of limited reweighting schemes, our results indicate that in our setup, systemic risk could not be reduced while maintaining the current size of the largest banks. However, in terms of optimal portfolio design, bringing their cumulative weight down from 90% to 50% yielded a significant improvement in systemic risk.

6.4.2 Does the result hold for a system where only listed banks are considered?

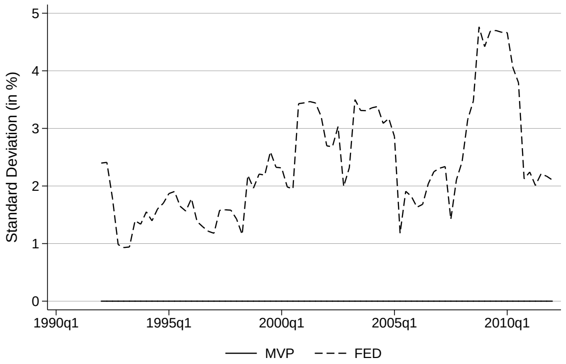
One assumption we have consistently made is that the regulator is able to move substantial amounts of assets from large banks to very small ones. However, small banks might not be able to sustain such an increase in assets in the first place. Moreover, previous research has shown that start-up banks only behave as mature banks after their first nine years of existence (see e.g. DeYoung and Hasan, 1998). Because these small banks might not be realistic investments for the supervisor, we select only those BHCs which have publicly traded equity, using the CRSP-FRB link provided by the Federal Reserve Bank of New York (2013). We thereby also remove those banks that do not file reports after 2006 due to the increase in the reporting threshold. The selected banks are considered to be the entire portfolio in which the supervisor can invest. The results in the Online Appendix show that removing these banks does not quantitatively or qualitatively change our results: It is possible to minimize portfolio risk by relocating assets from the largest listed banks to smaller listed ones. Moreover, in this new MVP, the actual largest 5% of banks would still receive a very low weight.[20]

Next, we use the CRSP-FRB link to obtain the market valuation of assets for listed BHCs by downloading equity prices as well as the number of outstanding shares from CRSP, and match them to the Call Reports. The quarterly market valuation of assets is obtained by adding the book value of the liabilities to the average market capitalization during that quarter. Compared with book value, which gives information on the past performance of a bank, the market valuation should indicate what market participants believe to be the value of the bank going forward, notwithstanding liquidity

[20] A scenario in which the smallest 60% of banks were removed yielded similar results.

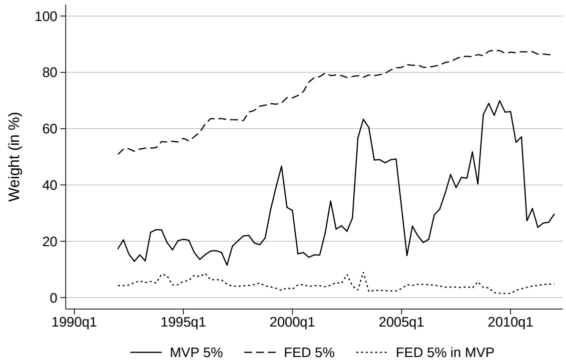
Figure 6.15: What role does inequality play in the risk/return trade-off when taking into account the market valuation of assets?

(a) How much is systemic risk reduced in the MVP?



(b) Do returns have to be sacrificed to achieve this lower risk?

(c) How unequal is the MVP compared to the FED portfolio?



The figure shows the comparison between the FED portfolio and the hypothetical MVP: panel 6.15a and 6.15b display the difference in the risk and return of each portfolio. Panel 6.15c shows how the weights are distributed in each portfolio by plotting their concentration ratios, as well as the weights that the current largest banks have in the MVP.

considerations during crises (Allen and Carletti, 2008). To perform our analysis on this subsample, we define returns as the quarter-to-quarter percentage changes in the market valuation of assets, and a banks' weight as the relative share in the portfolio. As in the baseline scenario, we obtain the MVP using an eight quarter moving window. If a merger takes place during this time, the assets of the acquired bank are added to the acquiring bank before the merger takes place for the appropriate quarters, while the acquired bank is removed from the investment universe.[21] The results are reported in Figure 6.15.

[21] Similar to the analyses on book data, banks that report a return below -100% or above +100% are not considered.

Table 6.4: Extensions and policy implications tests

	Issue	Policy test	Risk minimized? BHC	COMM
(1)	High turnover	Equally Weighted Portfolio	✗	0/12
(2)	High turnover	Reweighting limited to 10% or 20% of assets[a]	✗	0/12
(3)	High turnover	Reweighting limited to mean/std. dev. of growth[a]	✗	0/12
(4)	High turnover	Largest 5% keep their weight[a]	✗	5/12
(5)	High turnover	Largest 5% are weighted between 50% and 60%	✓	12/12
(6)	Small/DeNovo banks	Only BHCs with publicly traded equity – book value	✓	n.a.
(7)	Small/DeNovo banks	Only BHCs with publicly traded equity – market value	✓	n.a.

The table presents a summary of the seven policy tests that were performed and how well they worked in reducing systemic risk on FED and FRD level. To see if a test is successful we define the ratio $(\sigma_{\text{FED}} - \sigma_{\text{MVP}})/\sigma_{\text{FED}}$, which measures the relative difference in portfolio risk between the actual and minimum variance portfolio. A score of 1 indicates that the risk has been effectively eliminated, while a score of 0 indicates that no improvements were possible. A test is regarded as successful when this ratio averages 0.9 or higher throughout the sample period.
[a] To avoid the no portfolio loss constraint driving these results, we also performed the tests without the no-loss constraint. The results however do not change.

Figure 6.15 shows spikes in the portfolio risk of the FED portfolio at the beginning of the 2000s and during the subprime crisis, whereas that of the MVP is essentially zero. While the returns of the MVP are slightly higher prior to 2005, they are similar to the FED portfolio thereafter. Regarding the weight distribution of both portfolios, we again see a steady increase in the weights of the largest banks in the FED portfolio. The MVP based on market valuation shows a higher concentration than before, as the largest banks on average are assigned 33% of the assets compared with 13% in the baseline scenario. Similar to the baseline scenario, concentration in the MVP peaks in the subprime crisis when the largest banks have a weight of 69%. However, it appears that the largest banks in the actual portfolio are still overweighted, as they see their weight reduced to an average of 4.4% in the MVP. Table 6.4 summarizes the results of all scenarios in this section.

6.5 Conclusion

The last two decades have seen a major wave of consolidation and concentration of assets in the banking industry. In the same period, the sector has experienced two major crises with a significant impact on the real economy, of which the subprime crisis had global repercussions. As a consequence of recent bailouts and government-forced sales, the sector is now even more concentrated than before the crises. In the light of moral hazard and Too-Big-To-Fail banks, we have investigated how the high concentration in the industry impacts systemic risk.

In the absence of counterfactuals, we consider a thought experiment in which we view the supervisor as a constrained investor in a portfolio of banks. As profit maximization by banks is inherently risky, but is not perfectly correlated with that of other banks, the portfolio of the supervisor will have a better risk-return profile. By applying elements of Modern Portfolio Theory, we derive a hypothetical distribution of weights that the supervisor should have held to arrive at the minimum variance portfolio in order to give us insights into the optimal design of the banking system.

Our results consistently show that the hypothetical minimum variance portfolio had a lower risk than the actual portfolio, achieved by reducing the level of concentration in the portfolio. Moreover, it was not necessary to sacrifice returns in order to achieve this lower risk. The minimum variance portfolio favors more traditional banks as measured by the non-interest income/total income ratio and balance sheet items such as loans, trading assets and deposits. In contrast, an equally-weighted portfolio would perform similarly to the actual, concentrated, system. These findings are robust to different starting values, time windows, covariance matrices and the exclusion of the smallest banks.

However, to achieve lower risk, the supervisor would have to adjust weights in each quarter, leading to a portfolio turnover three times higher than that of the real portfolio. Since this might not be possible or even desirable within the current regulatory framework, we tested several alternatives involving limited reweighting which were largely unsuccessful. These findings indicate that in times of crisis, an increase in systemic risk was unavoidable while keeping the concentration at current levels. Nonetheless, our analysis did show - ceteris paribus - that when the weight of the largest banks was kept at a sufficiently low level, systemic risk was reduced significantly in the hypothetical minimum variance portfolio.

The policy implications flowing from these findings are that supervisors should seriously consider the effects of concentration on systemic risk. A reduction in disparity of size could create a more competitive environment, similar to provisions of the proposed Safe, Accountable, Fair & Efficient (SAFE) Banking Act of 2012, which would limit individual banks' funding strategies to 10% of the total industry. Other measures could include imposing higher equity capital demands for the large banks as are currently being implemented, or separating investment from retail banking as proposed by Paul Volcker, and the Vickers and Liikanen reports in Europe. Forcing large banks to hold more capital could lead to a relative reduction in their size only if their assets are 'redistributed' to smaller banks in order to maintain a safer and more competitive environment. Our findings show that we should not only consider the size of each bank individually, but also consider each bank's size with respect to the whole system. However, given the data at hand which mirrors the current regulatory set-up, the results presented here can only be interpreted as a partial equilibrium effect as we can not consider the shifting of bank activities abroad or into the shadow banking realm.

References

Abraham, F., T. Goesaert, and J. Konings (2014). Staying Home or Moving Away? Restructuring Efforts within Multinational Enterprises. *The World Economy 37*(6), pp. 765–782.

Acharya, V. V. (2009). A theory of systemic risk and design of prudential bank regulation. *Journal of Financial Stability 5*(3), pp. 224–255.

Acharya, V. V., R. F. Engle, and M. P. Richardson (2012a). Capital Shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review 102*(3), pp. 59–64.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. P. Richardson (2012b). Measuring systemic risk. CEPR Discussion Papers 8824, C.E.P.R. Discussion Papers.

Acharya, V. V. and T. Yorulmazer (2007). Too many to fail – An analysis of time-inconsistency in bank closure policies. *Journal of Financial Intermediation 16*(1), pp. 1–31.

Ackerberg, D. A., K. Caves, and G. Frazer (2008). Structural identification of production functions. mimeo URL:http://www.econ.ucla.edu/ackerber/ACF20withtables.html.

Adrian, T. and M. K. Brunnermeier (2011). CoVaR. Working Paper 17454, National Bureau of Economic Research.

Aghion, P., R. Burgess, S. Redding, and F. Zilibotti (2005). Entry liberalization and inequality in industrial performance. *Journal of the European Economic Association 3*(2-3), pp. 291–302.

Aitken, B., A. Harrison, and R. E. Lipsey (1996). Wages and foreign ownership: A comparative study of Mexico, Venezuela, and the United States. *Journal of International Economics 40*, pp. 345–371.

Aitken, B. J. and A. E. Harrison (1999). Do domestic firms benefit from direct foreign investment? Evidence from Venezuela. *American Economic Review 89*(3), pp. 605–618.

Alegria, C. and K. Schaeck (2009). On measuring concentration in banking systems. *Finance Research Letters 5*, pp. 59–67.

Alfaro, L. and M. X. Chen (2012). Surviving the global financial crisis: Foreign ownership and establishment performance. *American Economic Journal: Economic Policy 4*(3), pp. 30–55.

Allen, F., A. Babus, and E. Carletti (2012). Asset commonality, debt maturity and systemic risk. *Journal of Financial Economics 104*(3), pp. 519–534.

Allen, F. and E. Carletti (2008). Mark-to-market accounting and liquidity pricing. *Journal of Accounting and Economics 45*(2-3), pp. 358–378.

Altomonte, C. and I. Colantone (2008). Firm heterogeneity and endogenous regional disparities. *Journal of Economic Geography 8*, pp. 779–810.

Altomonte, C. and E. Pennings (2009). Domestic plant productivity and incremental spillovers from foreign direct investment. *Journal of International Business Studies 40*, pp. 1131–1148.

Alvarez, R. and H. Görg (2007). Multinationals as stabilizers?: Economic crisis and plant employment growth. Working Papers Central Bank of Chile 412, Central Bank of Chile.

Antonescu, D. (2012). Identifying Regional Disparities in Romania: A Convergence Process Perspective in Relation to European Union's Territorial Structures. *Procedia Economics and Finance 3*(0), pp. 1144–1155.

Antràs, P. and E. Helpman (2004). Global Sourcing. *Journal of Political Economy 112*(3), pp. 552–580.

Arnold, J. M. and B. S. Javorcik (2009). Gifted kids or pushy parents? Foreign direct investment and plant productivity in Indonesia. *Journal of International Economics 79*(1), pp. 42 – 53.

Beck, T., H. Degryse, R. de Haas, and N. van Horen (2014). When Arm's Length is Too Far : Relationship Banking over the Business Cycle. Discussion Paper 2014-042, Tilburg University, Center for Economic Research.

Beck, T., A. Demirgüç-Kunt, and R. Levine (2000). A new database on the structure and development of the financial sector. *The World Bank Economic Review 14*(3), pp. 597–605.

Becker, S., P. Egger, and M. von Ehrlich (2010). Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics 94*, pp. 578–590.

Becker, S., P. Egger, and M. von Ehrlich (2012). Too much of a good thing? on the growth effects of the EU's regional policy. *European Economic Review 56*(4), pp. 648–668.

Becker, S. O., K. Boeckh, C. Hainz, and L. Woessmann (2011). The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy. CEPR Discussion Paper 8288.

Bems, R., R. C. Johnson, and K.-M. Yi (2012). The Great Trade Collapse. NBER Working Papers 18632, National Bureau of Economic Research, Inc.

Bena, J. and P. Ondko (2012). Financial development and the allocation of external finance. *Journal of Empirical Finance 19*(1), pp. 1–25.

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

Bernard, A. B. and J. B. Jensen (2007). Firm Structure, Multinationals, and Manufacturing Plant Deaths. *Review of Economics and Statistics 89*(2), pp. 193–204.

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 1995*, pp. 67–119.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2014). Production networks, geography and firm performance. Working paper, Tuck School of Business.

Bernard, A. B. and F. Sjöholm (2003). Foreign owners and plant survival. Working Paper 10039, National Bureau of Economic Research.

Bertay, A. C., A. Demirgüç-Kunt, and H. Huizinga (2013). Do we need big banks? Evidence on performance, strategy and market discipline. *Journal of Financial Intermediation 22*(4), pp. 532–558.

Bezemer, D. (2014). Schumpeter might be right again: the functional differentiation of credit. *Journal of Evolutionary Economics 24*(5), pp. 935–950.

Bisias, D., M. Flood, A. W. Lo, and S. Valavanis (2012). A survey of systemic risk analytics. *Annual Review of Financial Economics 4*(1), pp. 255–296.

Bitzenis, A. (2006). Decisive FDI barriers that affect multinationals' business in a transition country. *Global Business and Economics Review 8*, pp. 87–118.

Blalock, G. and P. J. Gertler (2008). Welfare gains from foreign direct investment through technology transfer to local suppliers. *Journal of International Economics 74*(2), pp. 402–421.

Blanchard, P., E. Dhyne, C. Fuss, and C. Mathieu (2012). (Not so) easy come, (still) easy go? Footloose multinationals revisited. Working Paper Research 223, National Bank of Belgium.

Blank, S., C. M. Buch, and K. Neugebauer (2009). Shocks at large banks and banking sector distress: The banking granular residual. *Journal of Financial Stability 5*(4), pp. 353–373.

Blundell, R. and S. Bond (1998). Initial conditions and moment restriction in dynamic panel data models. *Journal of Econometrics 87*, pp. 115–143.

Boot, A. and L. Ratnovski (2012). Banking and Trading. IMF Working Paper 12/238.

Brodie, J., I. Daubechies, C. De Mol, D. Giannone, and I. Loris (2009). Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences 106*(30), pp. 12267–12272.

Brown, C. O. and I. S. Dinç (2011). Too many to fail? Evidence of regulatory forbearance when the banking sector is weak. *Review of Financial Studies 24*(4), pp. 1378–1405.

Brownlees, C. T. and R. F. Engle (2012). Volatility, correlation and tails for systemic risk measurement. Working paper. <http://ssrn.com/abstract=1611229>.

Brunnermeier, M. K., G. Dong, and D. Palia (2012). Banks' non-interest income and systemic risk. Working paper. <http://ssrn.com/abstract=1786738>.

Buch, C. M., I. Kesternich, A. Lipponer, and M. Schnitzer (2010). Exports versus FDI revisited: does finance matter? Discussion Paper Series 1: Economic Studies 2010,03, Deutsche Bundesbank, Research Centre.

Bureau van Dijk Electronic Publishing (various issues). Amadeus database. Brussels, Belgium.

Cantwell, J. and L. Piscitello (2007). Attraction and deterrence in the location of foreign-owned R&D activities: the role of positive and negative spillovers. *International Journal of Technological Learning, Innovation and Development 1*(1), pp. 83–111.

Carvalho, V. and X. Gabaix (2013). The great diversification and its undoing. *American Economic Review 103*(5), pp. 1697–1727.

Castellani, D. and G. Giovannetti (2010). Productivity and the international firm: dissecting heterogeneity. *Journal of Economic Policy Reform 13*(1), pp. 25–42.

Caves, R. E. (1974). Multinational firms, competition, and productivity in host-country markets. *Economica 41*(162), pp. 176–193.

Cecchetti, S. G. and E. Kharroubi (2015). Why does financial sector growth crowd out real economic growth? BIS Working Papers 490, Bank for International Settlements.

Cetorelli, N. and L. S. Goldberg (2012). Liquidity management of U.S. global banks: Internal capital markets in the great recession. *Journal of International Economics 88*(2), pp. 299–311.

Chen, N. (2004). Intra-national versus international trade in the European Union: why do national borders matter? *Journal of International Economics 63*(1), pp. 93–118.

Claeys, S. and K. Schoors (2007). Bank supervision Russian style: Evidence of conflicts between micro- and macro-prudential concerns. *Journal of Comparative Economics 35*(3), pp. 630–657.

Comin, D. A., M. Dmitriev, and E. Rossi-Hansberg (2012). The spatial diffusion of technology. NBER Working Papers 18534, National Bureau of Economic Research, Inc.

CompNet (2014). Micro-Based Evidence of EU Competitiveness: The CompNet Database. Working Paper No. 1634, European Central Bank, Frankfurt.

Crespo, N. and M. P. Fontoura (2007). Determinant factors of FDI spillovers - What do we really know? *World Development 35*(3), pp. 410–425.

Crespo, N., M. P. Fontoura, and I. Proença (2009). FDI spillovers at regional level: Evidence from Portugal. *Papers in Regional Science 88*(3), pp. 591–607.

Cristea, A. D. (2011). Buyer-seller relationships in international trade: Evidence from U.S. States' exports and business-class travel. *Journal of International Economics 84*(2), pp. 207–220.

Damijan, J. P., M. Rojec, B. Majcen, and M. Knell (2013). Impact of firm heterogeneity on direct and spillover effects of fdi: Micro-evidence from ten transition countries. *Journal of Comparative Economics 41*(3), pp. 895–922.

Davies, S. W., L. Rondi, and A. Sembenelli (2001). Are multinationality and diversification complementary or substitute strategies?: An empirical analysis on european leading firms. *International Journal of Industrial Organization 19*(8), pp. 1315–1346.

Davis, D. R. and J. I. Dingel (2012). A spatial knowledge economy. NBER Working Papers 18188, National Bureau of Economic Research, Inc.

de Haas, R. and I. Lelyveld (2014). Multinational Banks and the Global Financial Crisis: Weathering the Perfect Storm? *Journal of Money, Credit and Banking 46*, pp. 333–364.

de Haas, R. and I. van Lelyveld (2010). Internal capital markets and lending by multinational bank subsidiaries. *Journal of Financial Intermediation 19*(1), pp. 1–25.

De Jonghe, O. (2010). Back to the basics in banking? A micro-analysis of banking system stability. *Journal of Financial Intermediation 19*(3), pp. 387–417.

Degryse, H. and G. Nguyen (2007). Interbank exposures: An empirical examination of contagion risk in the Belgian banking system. *International Journal of Central Banking 3*(2), pp. 123–171.

DeMiguel, V., L. Garlappi, and R. Uppal (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies 22*(5), pp. 1915–1953.

Demirgüç-Kunt, A. and H. Huizinga (2013). Are banks too big to fail or too big to save? International evidence from equity prices and CDS spreads. *Journal of Banking & Finance 37*(3), pp. 875–894.

DeYoung, R. and I. Hasan (1998). The performance of de novo commercial banks: A profit efficiency approach. *Journal of Banking & Finance 22*, pp. 565–587.

DeYoung, R. and K. P. Roland (2001). Product mix and earnings volatility at commercial banks: Evidence from a degree of total leverage model. *Journal of Financial Intermediation 10*(1), pp. 54–84.

DeYoung, R. and G. Torna (2013). Nontraditional banking activities and bank failures during the financial crisis. *Journal of Financial Intermediation 22*(3), pp. 397–421.

Dikova, D., R. Smeets, H. Garretsen, and H. V. Ees (2013). Immediate responses to financial crises: A focus on US MNE subsidiaries. *International Business Review 22*(1), pp. 202 – 215.

Disney, R., J. Haskel, and Y. Heden (2003). Entry, Exit and Establishment Survival in UK Manufacturing. *The Journal of Industrial Economics 51*(1), pp. 91–112.

Dunne, P. and A. Hughes (1994). Age, Size, Growth and Survival: UK Companies in the 1980s. *The Journal of Industrial Economics 42*(2), pp. 115–140.

Dunne, T., M. J. Roberts, and L. Samuelson (1988). Patterns of Firm Entry and Exit in U.S. Manufacturing Industries. *RAND Journal of Economics 19*(4), pp. 495–515.

Egger, P., M. Fahn, V. Merlo, and G. Wamser (2014). On the genesis of multinational foreign affiliate networks. *European Economic Review 65*(C), pp. 136–163.

Engle, R. F., E. Jondeau, and M. Rockinger (2014). Systemic risk in Europe. *Review of Finance 19*(1), pp. 145–190.

European Commission (2011). Regions in the European Union—nomenclature of territorial units for statistics nuts 2010. Working Paper ISSN 1977-0375, European Commission, Brussels.

European Council (1993). Regulation on the statistical units for the observation and analysis of the production system in the community. *Official Journal of the European Communities L76*, pp. 1–11.

Fahlenbrach, R., R. Prilmeier, and R. M. Stulz (2012). This time is the same: Using bank performance in 1998 to explain bank performance during the recent financial crisis. *The Journal of Finance 67*(6), pp. 2139–2185.

Federal Reserve Bank of New York (2013). CRSP-FRB Link.

Fosfuri, A., M. Motta, and T. Rønde (2001). Foreign direct investment and spillovers through workers' mobility. *Journal of International Economics 53*(1), pp. 205–222.

Foster, L., J. C. Haltiwanger, and C. J. Krizan (2001). Aggregate productivity growth. lessons from microeconomic evidence. In C. R. Hulten, E. R. Dean, and M. J. Harper (Eds.), *New Developments in Productivity Analysis*, pp. 303–372. University of Chicago Press.

Furfine, C. H. (1999). The microstructure of the Federal Funds market. *Financial Markets, Institutions & Instruments 8*(5), pp. 24–44.

Furfine, C. H. (2003). Interbank exposures: Quantifying the risk of contagion. *Journal of Money, Credit and Banking 35*(1), pp. 111–128.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), pp. 733–772.

Gal, P. N. (2013). Measuring Total Factor Productivity at the Firm–Level using OECD–ORBIS. Working Paper No. 1049, Organisation for Economic Co-operation and Development, Paris.

Gibson, J. K. and R. I. D. Harris (1996). Trade Liberalisation and Plant Exit in New Zealand Manufacturing. *The Review of Economics and Statistics 78*(3), pp. 521–529.

Girma, S. (2005). Absorptive Capacity and Productivity Spillovers from FDI: A Threshold Regression Analysis. *Oxford Bulletin of Economics and Statistics 67*(3), pp. 281–306.

Giroud, A., B. Jindra, and P. Marek (2012). Heterogeneous FDI in transition economies - a novel approach to assess the developmental impact of backward linkages. *World Development 40*(11), pp. 2206–2220.

Giroud, X. (2012). Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics 128*(2), pp. 861–915.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer (1992). Growth in cities. *Journal of Political Economy 100*(6), pp. 1126–1152.

Glass, A. J. and K. Saggi (1998). International technology transfer and the technology gap. *Journal of Development Economics 55*, pp. 369–398.

Glass, A. J. and K. Saggi (2002). Multinational firms and technology transfer. *Scandinavian Journal of Economics 104*(4), pp. 495–513.

Godart, O., H. Görg, and A. Hanley (2011). Surviving the Crisis: Foreign Multinationals vs Domestic Firms in Ireland. IZA Discussion Papers 5882, Institute for the Study of Labor (IZA).

Görg, H. and D. Greenaway (2004). Much ado about nothing? Do domestic firms really benefit from foreign direct investment? *World Bank Research Observer 19*(2), pp. 171–197.

Görg, H. and F. Ruane (2001). Multinational companies and linkages: Panel-data evidence for the Irish electronics sector. *International Journal of the Economics of Business 8*(1), pp. 1–18.

Görg, H. and E. Strobl (2005). Spillovers from foreign firms through worker mobility: An empirical investigation. *The Scandinavian Journal of Economics 107*(4), pp. 693–709.

Görg, H., A. Hijzen, and B. Muraközy (2009). The role of production technology for productivity spillovers from multinationals: Firm-level evidence for Hungary. Kiel Working Papers 1482, Kiel Institute for the World Economy.

Görg, H. and E. Strobl (2002). 'Footloose' Multinationals? CEPR Discussion Papers 3402, C.E.P.R. Discussion Papers.

Griliches, Z. and J. Mairesse (1995). Production functions: The search for identification. NBER Working Paper No. 5067.

Halpern, L. and B. Muraközy (2007). Does distance matter in spillover? *Economics of Transition 15*, pp. 781–805.

Harding, T. and B. Smarzynska Javorcik (2011). Roll out the red carpet and they will come: Investment promotion and FDI inflows. *Economic Journal 121*(557), pp. 1445–1476.

Havránek, T. and Z. Irŝová (2011). Estimating vertical spillovers from FDI: Why results vary and what the true effect is. *Journal of International Economics 85*(2), pp. 234–244.

Havránek, T. and Z. Irŝová (2013). Determinants of horizontal spillovers from FDI: Evidence from a large meta-analysis. *World Development 42*(0), pp. 1–15.

Havranek, T. and Z. Irsova (2015). Do Borders Really Slash Trade? A Meta-Analysis. William Davidson Institute Working Papers Series WP1088, William Davidson Institute at the University of Michigan.

Head, K. and T. Mayer (2002). *Illusory Border Effects: Distance mismeasurement inflates estimates of home bias in trade*, Volume 1. CEPII Paris.

Hebous, S. and A. Weichenrieder (2010). Debt financing and sharp currency depreciations: wholly versus partially-owned multinational affiliates. *Review of World Economics (Weltwirtschaftliches Archiv) 146*(2), pp. 281–302.

Heckman, J. J., H. Ichimura, and P. Todd (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies 65*(2), pp. 261–94.

Helpman, E., M. Melitz, and S. R. Yeaple (2004). Export versus FDI with heterogeneous firms. *American Economic Review 94*, pp. 300–316.

Hilber, C. A. L. and I. Voicu (2010). Agglomeration economies and the location of foreign direct investment: Empirical evidence from Romania. *Regional Studies 44*(3), pp. 355–371.

Huang, R. and L. Ratnovski (2011). The dark side of bank wholesale funding. *Journal of Financial Intermediation 20*(2), pp. 248–263.

Hughes, J. P. and L. J. Mester (2013). Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation 22*(4), pp. 559–585.

Ibragimov, R., D. Jaffee, and J. Walden (2011). Diversification disasters. *Journal of Financial Economics 99*(2), pp. 333–348.

Ivarsson, I. (2002). Transnational corporations and the geographical transfer of localised technology: a multi-industry study of foreign affiliates in Sweden. *Journal of Economic Geography 2*(2), pp. 221–247.

Ivarsson, I. and C. G. Alvstam (2005). Technology transfer from TNCs to local suppliers in developing countries: A study of AB Volvo's truck and bus plants in Brazil, China, India, and Mexico. *World Development 33*(8), pp. 1325–1344.

Janicki, H. P. and E. S. Prescott (2006). Changes in the size distribution of U.S. banks: 1960-2005. *Economic Quarterly 92*(4), pp. 291–316.

Javorcik, B. S. (2004). Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages. *American Economic Review 94*(3), pp. 605–627.

Javorcik, B. S. and S. Poelhekke (2014). Former Foreign Affiliates: Cast Out and Outperformed? CESifo Working Paper Series 5111, CESifo Group Munich.

Javorcik, B. S. and M. Spatareanu (2008). To share or not to share: Does local participation matter for spillovers from foreign direct investment? *Journal of Development Economics 85*(1-2), pp. 194–217.

Jobson, J. D. and B. Korkie (1980). Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association 75*(371), pp. 544–554.

Kamal, F. (2014). Does firm ownership affect spillover opportunities? Evidence from Chinese manufacturing. *Journal of Regional Science 54*(1), pp. 137–154.

Keller, W. (2002). Geographic localization of international technology diffusion. *American Economic Review 92*(1), pp. 120–142.

Keller, W. and S. R. Yeaple (2013). The gravity of knowledge. *American Economic Review 103*(4), pp. 1414–1444.

King, R. G. and R. Levine (1993). Finance and Growth: Schumpeter Might Be Right. *The Quarterly Journal of Economics 108*(3), pp. 717–37.

Landier, A., V. B. Nair, and J. Wulf (2009). Trade-offs in staying close: Corporate decision making and geographic dispersion. *Review of Financial Studies 22*(3), pp. 1119–1148.

Ledoit, O. and M. Wolf (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance 10*(5), pp. 603–621.

Ledoit, O. and M. Wolf (2004). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management 30*(4), pp. 110–119.

Lenaerts, K. and B. Merlevede (2012). Supply chain fragmentation and spillovers from foreign direct investment. Technical report, Ghent University, Faculty of Economics and Business Administration.

Lenaerts, K. and B. Merlevede (2015). Firm size and spillover effects from foreign direct investment: the case of Romania. *Small Business Economics 45*(3), pp. 595–611.

LeSage, J. (2014). What regional scientists need to know about spatial econometrics. *The Review of Regional Studies 44*, pp. 13–32.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies 70*(2), pp. 317–341.

Liu, Z. (2008). Foreign direct investment and technology spillovers: Theory and evidence. *Journal of Development Economics 85*, pp. 176–193.

Loutskina, E. and P. E. Strahan (2011). Informed and uninformed investment in housing: The downside of diversification. *Review of Financial Studies 24*(5), pp. 1447–1480.

Manova, K. (2013). Credit Constraints, Heterogeneous Firms, and International Trade. *Review of Economic Studies 80*(2), pp. 711–744.

Manova, K., S.-J. Wei, and Z. Zhang (2015). Firm exports and multinational activity under credit constraints. *Review of Economics and Statistics 97*(3), pp. 574–588.

Marin, A. and M. Bell (2006). Technology spillovers from foreign direct investment (FDI): the active role of MNC subsidiaries in Argentina in the 1990s. *Journal of Development Studies 42*(2), pp. 678–697.

Mariotti, S., M. Mutinelli, M. Nicolini, and L. Piscitello (2014). Productivity spillovers from foreign multinational enterprises to domestic manufacturing firms: To what extent does spatial proximity matter? *Regional Studies 49*(10), pp. 1639–1653.

Mariotti, S. and L. Piscitello (1995). Information costs and location of FDIs within the host country: Empirical evidence from Italy. *Journal of International Business Studies 26*(4), pp. 815–841.

Mariotti, S., L. Piscitello, and S. Elia (2010). Spatial agglomeration of multinational enteprises: the role of information externalities and knowledge spillovers. *Journal of Economic Geography 10*, pp. 519–538.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), pp. 77–91.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.

Markusen, J. R. (1995). The Boundaries of Multinational Enterprises and The Theory of International Trade. *Journal of Economic Perspectives 9*(2), pp. 169–189.

McCallum, J. (1995). National Borders Matter: Canada-U.S. Regional Trade Patterns. *American Economic Review 85*(3), pp. 615–23.

McCann, P. and Z. J. Acs (2011). Globalization: Countries, cities and multinationals. *Regional Studies 45*(1), pp. 17–32.

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

Merlevede, B. and V. Purice (2015). Distance, time since foreign entry, and productivity spillovers from foreign direct investment. *Papers in Regional Science*, forthcoming.

Merlevede, B. and K. Schoors (2007). FDI and the consequences towards more complete capture of spillover effects. William Davidson Institute Working Paper No. 886.

Merlevede, B., K. Schoors, and M. Spatareanu (2014). FDI Spillovers and Time since Foreign Entry. *World Development 56*, pp. 108–126.

Moran, T. H., E. M. Graham, and M. Blomström (2005). *Does foreign direct investment promote development?* Peterson Institute.

Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *Review of Economics and Statistics 72*(2), pp. 334–338.

Narula, R. and A. Marin (2005). Foreign direct investment spillovers, absorptive capacities and human capital development : evidence from Argentina. ILO Working Papers 378212, International Labour Organization.

Navaretti, G. B., D. Checchi, and A. Turrini (2003). Adjusting Labour Demand: Multinational versus National Firms- A Cross-European Analysis. *Journal of the European Economic Association 1*(2-3), pp. 708–719.

Nicolini, M. and L. Resmini (2010). FDI spillovers in new EU member states: Which firms create them and which firms really benefit? *Economics of Transition 18*(3), pp. 487–511.

Norback, P.-J., A. Tekin-Koru, and A. Waldkirch (2013). Multinational Firms and Plant Divestiture. MPRA Paper 45957, University Library of Munich, Germany.

Office of the Special Inspector General for the Troubled Asset Relief Program (2013, October 29). Quarterly report to congres. Technical report, US Treasury.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), pp. 1263–1297.

Partridge, M. D., M. Boarnet, S. Brakman, and G. Ottaviano (2012). Introduction: Whither spatial econometrics? *Journal of Regional Science 52*(2), pp. 167–171.

Petrin, A. and J. Levinsohn (2012). Measuring Aggregate Productivity Growth using Plant-Level Data. *The RAND Journal of Economics 43*(4), pp. 705–725.

Poelhekke, S. (2015). Do global banks facilitate foreign direct investment? *European Economic Review 76*(0), pp. 25–46.

Rajan, R. G. and L. Zingales (1998). Financial dependence and growth. *The American Economic Review 88*(3), pp. 559–586.

Rajan, R. S. (2005). Foreign Direct Investment and the Internationalization of Production in the Asia-Pacific Region: Issues and Policy Conundrums. *Asia–Pacific Trade and Investment Review 1*(1), pp. 3–26.

Rodríguez-Clare, A. (1996). Multinationals, linkages, and economic development. *American Economic Review 86*(4), pp. 852–873.

Romanian National Statistical Office (RNSO) (1998 - 2014). TEMPO-Online Time Series. Bucharest, Romania.

Romanian National Statistical Office (RNSO) (2005). Statistical yearbook. Bucharest, Romania.

Ronnås, P. (1984). *Urbanization in Romania. A Geography of Social and Economic Change Since Independence*. Ph. D. thesis, Stockholm School of Economics.

Rousseau, P. L. and P. Wachtel (2011, 01). What Is Happening To The Impact Of Financial Deepening On Economic Growth? *Economic Inquiry 49*(1), pp. 276–288.

Schoors, K. and B. van der Tol (2002). Foreign direct investment spillovers within and between sectors: Evidence from Hungarian data. Ghent University Faculty of Economics and Business Administration, Ghent University Working Paper: No. 02/157.

Sinani, E. and K. E. Meyer (2004). Spillovers of technology transfer from FDI: the case of Estonia. *Journal of Comparative Economics 32*(3), pp. 445–466.

Spies, J. (2010). Network and border effects: Where do foreign multinationals locate in Germany? *Regional Science and Urban Economics 40*(1), pp. 20–32.

Stiroh, K. J. (2004). Diversification in banking: Is noninterest income the answer? *Journal of Money, Credit and Banking 36*(5), pp. 853–882.

Stiroh, K. J. and A. Rumble (2006). The dark side of diversification: The case of US financial holding companies. *Journal of Banking & Finance 30*, pp. 2131–2161.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature 49*(2), pp. 326–365.

Teece, D. J. (1977). Technology Transfer by Multinational Firms: The Resource Cost of Transferring Technological Know–how. *Economic Journal 87*(346), pp. 242–261.

Timmer, M. P. (2012). The World Input-Output Database (WIOD): Contents, Sources and Methods. Working Paper No. 10, World Input-Output Database, Groningen. (Downloadable at http://www.wiod.org/publications/papers/wiod10.pdf).

Timmer, M. P., A. A. Erumban, B. Los, R. Stehrer, and G. J. de Vries (2014). Slicing up global value chains. *Journal of Economic Perspectives 28*(2), pp. 99–118.

Urata, S. and H. Kawai (2000). Intrafirm technology transfer by Japanese manufacturing firms in Asia. In T. Ito and A. O. Krueger (Eds.), *The Role of Foreign Direct Investment in East Asian Economic Development*, Volume 9, pp. 49–77. University of Chicago Press.

van Lelyveld, I. and F. Liedorp (2006). Interbank contagion in the Dutch banking sector: A sensitivity analysis. *International Journal of Central Banking 2*(2), pp. 99–133.

Vandenbussche, H., C. Altomonte, and T. Sonno (2013). Firm-level productivity and exporting: Diagnosing the role of financial constraints. In *Product Market Review 2013 - Financing the real economy*. European Commission.

Vienna Institute for International Economic Studies (2008). Industrial database Eastern Europe. Vienna, Austria.

Wagner, J. (2014). Credit constraints and exports: a survey of empirical studies using firm-level data. *Industrial and Corporate Change 23*(6), pp. 1477–1492.

Wagner, W. (2008). The homogenization of the financial system and financial crises. *Journal of Financial Intermediation 17*(3), pp. 330–356.

Wen, Y. (2014). The spillover effect of FDI and its impact on productivity in high economic output regions: A comparative analysis of the Yangtze River Delta and the Pearl River Delta, China. *Papers in Regional Science 93*(2), pp. 341–365.

Wheelock, D. C. and P. W. Wilson (2012). Do large banks have lower costs? New estimates of returns to scale for U.S. banks. *Journal of Money, Credit and Banking 44*(1), pp. 171–199.

Wooldridge, J. M. (2009). On Estimating Firm-Level Production Functions using Proxy Variables to Control for Unobservables. *Economics Letters 104*, pp. 112–114.

Yasar, M., P. Garcia, C. Nelson, and R. Rejesus (2007). Is there Evidence of Learning-by-Exporting in Turkish Manufacturing Industries? *International Review of Applied Economics 21*(2), pp. 293–305.

Zhang, Y. A., Y. Li, and H. Li (2013). FDI spillovers over time in an emerging market: the roles of entry tenure and barriers to imitation. *Academy of Management Journal 57*(3), pp. 698–722.

Nederlandse samenvatting

Dit proefschrift beantwoordt drie vragen met betrekking tot de invloed van multinationale ondernemingen op lokale ondernemingen en de rol van de ontwikkeling van financiële markten hierin. Ten eerste onderzoekt het proefschrift de impact van nationale grenzen en afstand op indirecte effecten van de aanwezigheid van multinationale ondernemingen op lokale, binnenlandse bedrijven. Ten tweede kijkt het naar de directe impact van multinationale ondernemingen op hun dochterondernemingen. Ten slotte kijkt het naar de rol die het bankwezen hierin speelt, en hoe de gevolgen van financiële instabiliteit kunnen worden beperkt.

Om deze vragen empirisch te beantwoorden wordt grotendeels gebruik gemaakt van de database Amadeus, welke winst- en verliesrekeningen en balans-informatie bevat voor bijna 21 miljoen Europese ondernemingen. Deze database wordt verzameld en onderhouden door Bureau van Dijk Electronic Publishing. Vanuit Amadeus worden twee nieuwe datasets gecreëerd: *AUGAMA*, een panel van Europese bedrijven gedurende de periode 1996-2011, en *EUMULNET*, een dataset van netwerken van Europese multinationale ondernemingen. Hoofdstuk twee van het proefschrift beschrijft hoe deze datasets vanuit de ruwe Amadeus data voor 26 landen opgebouwd worden, en toont aan dat *AUGAMA* en *EUMULNET* de structuur van de Europese economie zeer goed benaderen.

Het derde en vierde hoofdstuk analyseren de indirecte effecten van multinationale ondernemingen op lokale, binnenlandse bedrijven. Hier ligt de nadruk op de studie van mogelijke productiviteitswinsten van binnenlandse bedrijven als gevolg van de vestiging van een multinationale onderneming in de buurt. Voorgaande literatuur heeft gemengde resultaten opgeleverd, in sommige gevallen zijn de indirecte effecten positief, in andere gevallen negatief. Welke kant deze effecten opgaan blijkt afhankelijk van de relatie van de multinationale onderneming tot de lokale onderneming. Als concurrent bijvoorbeeld kan de multinationale onderneming binnenlandse bedrijven dwingen om meer productief te worden, terwijl zij als afnemer van intermediaire goederen op de binnenlandse markt hogere kwaliteitseisen kan stellen. Waar voorgaande literatuur er van uitging dat deze indirecte effecten zich onmiddellijk en voor elke binnenlandse onderneming zouden manifesteren, nuanceert hoofdstuk 3 deze impliciete assumptie. Meer bepaald toont hoofdstuk drie aan dat binnenlandse bedrijven die dicht bij een nieuwe vestiging van een multinational gelegen zijn snel productiviteitswinsten laten optekenen. Binnenlandse bedrijven die echter verder weg gelegen zijn ervaren pas later én in mindere mate productiviteitswinsten ten gevolge van de aanwezigheid van de multinationale onderneming.

Gelet op de belangrijke rol voor afstand inzake de omvang van de indirecte effecten op de productiviteit van lokale ondernemingen van multinationals gaat hoofdstuk vier na of indirecte effecten door landsgrenzen beïnvloed worden. Voorgaande literatuur heeft enkel effecten binnen landsgrenzen geanalyseerd. Er is echter geen reden om a priori aan te nemen dat een lokale onderneming niet beïnvloed kan worden door een multinationale onderneming die zich voldoende dichtbij, doch in een ander land bevindt. Hoofdstuk vier toont aan dat landsgrenzen echter niet poreus genoeg zijn, en dat mogelijke productiviteitswinsten gehinderd worden door landsgrenzen. Hoofdstuk vier toont echter ook aan dat grenzen tussen landen die deel uitmaken van het Schengengebied productiviteitswinsten door technologietransfer niet significant verminderen, terwijl niet-Schengen grenzen dit wel doen.

Terwijl hoofdstuk drie en vier indirecte effecten van multinationale ondernemingen behandelen, bestudeert hoofdstuk vijf directe effecten. Met name wordt in hoofdstuk vijf geanalyseerd welke factoren een rol spelen bij de beslissing van multinationale hoofdkwartieren om tot sluiting van een dochteronderneming over te gaan. In tegenstelling tot voorgaande literatuur bestudeert hoofdstuk vijf deze beslissing voor, tijdens én na de financiële crisis, en wordt er ook rekening gehouden met de prestaties van andere dochterondernemingen binnen hetzelfde multinationale netwerk. De analyses in hoofdstuk vijf tonen aan dat multinationale ondernemingen met een groot netwerk van gediversifieerde dochterondernemingen in het algemeen minder snel hun dochterondernemingen zullen sluiten. Daarnaast zullen multinationale ondernemingen met een meer geavanceerd financieel systeem in hun thuisland hun dochterondernemingen minder snel sluiten in tijden van hoogconjunctuur. Tijdens de financiële crisis speelt het financieel systeem in het thuisland geen rol meer.

Ten slotte, waar hoofdstuk vijf vind dat een gezond financieel systeem gunstige gevolgen heeft voor economische groei, multinationale ondernemingen en productiviteitsgroei, analyseert hoofdstuk zes de rol van alsmaar groeiende banken in diepere economische integratie en de gevolgen hiervan voor financiële stabiliteit. De resultaten in hoofdstuk zes suggereren dat net kleinere banken goed zijn voor financiële stabiliteit, en - door meer leningen te verschaffen aan bedrijven en huishoudens - ook voor de reële economie.