Trade and Capital Movements in a Globalizing World

Jonas Gade Christensen

Dissertation for the degree philosophiae doctor (PhD)
at the University of Bergen

2011

To my family

Acknowledgements

After getting my masters degree in 2006 some people suggested that I should apply for a PhD scholarship. I thought that was a ridiculous idea, as I wanted to get out in the real world and deal with real world problems. And so I did. The idea of putting 'Dr.' in front of my name started to grow on me though, and after about a year I came lurking back and started my life as a PhD student. I had four specific research ideas in my project description, none of which ever saw the light, but somehow I managed to put together a dissertation, and there are many people who deserve thank for helping me do that.

First and foremost I would like to thank my main supervisor, Gaute Torsvik. In spite of specializing in quite different topics than mine, Gaute has always been able to come up with good comments and intelligent questions that have helped me move forward. His reviews have also contributed greatly to making my papers more readible and structurally more consistent. I would also like to thank my second supervisor, Eirik Gaard Kristiansen, who has been an important discussion partner as well, even when I have happened to send him old versions of my work.

I also wish to thank people at the Department of Economics. I have very much enjoyed the way the PhD students are integrated with the staff. This has made me feel an integrated and respected part of the department. There is never a bad moment to knock on someone's door to ask for help with a problem, or advice on a subject. Also, some of the professors have never seemed to mind bursting into the PhD students' offices, possibly not so often seeking advice, but to socialize, and disrupt the hard work of the young and promising. Those interruptions have been most welcome. I would also like to thank the other PhD students for making the last years and enjoyable time of my life. Especially I want to thank Julie Riise Kolstad for being my office partner during most of my time at the department. I enjoyed our talks and discussions a lot, and you always contributed a lot to bring a cheerful mood to the hallways at the fourth floor at the department. The previously mentioned interruptions by certain professors dropped significantly after you left, which might be an indication that not only I enjoyed your company. Katrine Løken has also been invaluable during these years. You have been more than helpful with any problem I have encountered at the department, whether it was of a practical or of an academic character. The rest of the fourth floor; Lene Lunde, Jurgita Januleviciute, Roar Gjelsvik, Karsten Marshall and Sigurd Birkeland also deserve thanks for always being ready for a

coffee-break, a fantasy-football-break, or just a talk-about-whatever-break. Bjørn Olav Johansen has also been an important person for me during these last years. Bjørn Olav never accepts a proposition without proof, and has an amazing ability for critical and inpendent assesments of any economical subject. Furthermore he is always a steady companion for after-work activities, whether it is a game of football, or a Friday afternoon beer. To all the newer PhD students I would also like to say thank you, good luck, and make sure to enjoy the reinvigorating effects of a good "coffee-talk" every now and again. Other friends, economists and non-economists have also been valuable discussion partners. Anders Gundersen deserves special thanks for volutarily reading through the entire dissertation, even claiming to having enjoyed it, and giving me important feedback. Andreas Tveito has often shown interest in the topics I have covered, and have been a good discussion partner on various occations.

From July 2008 to July 2009 I spent a year at Universidad Carlos III de Madrid. I learned a lot from some very talented people there, and I also had a great time in one of the greatest cities in the world. In particular I have to thank Nieves Valdes for helping me get settled and for sharing a mate when needed, the professors for teaching me a lot and for being available for discussions when needed, and not the least a thank you to Alberto Sánchez Esteban for improving my backhand significantly during that academic year.

Although this dissertation is a product of academic work, it would not have been possible without the support and encouragement I have received from my family and friends. My wife, Dani, has always had more faith in me than I have ever had, and has been of invaluable importance in pushing me through the hardest parts of this process. You and our unborn daughter are the main reasons that I have finished this. I love you very much..

To my parents, who have always encouraged and taken pride in my accomplishments, but never pressured me into pursuing anything that I did not want myself; you are examples that I wish to follow as a parent. I also wish to thank my sister, Nicca, for her heartfelt congratulations when this work was completed: "Nerd!".

Finally, this dissertation would have been a lot more difficult to complete without any financing, so I am very grateful for the grants provided by the Meltzer foundation, Norges Bank's fond, The Faculty of Social Sciences, and the Department of Economics at the University of Bergen.

Abstract

This dissertation consists of an introductory chapter and three essays on international trade and capital movements. Global trade has gone through many different phases over the centuries, and has changed character many times during that time. The introductory chapter gives a brief description of this, and also of how the academic literature has tried to explain different aspects of globalization. In this chapter I also characterize different approaches to trade and international capital movements within economic theory I show how my contributions fit into this literature.

All three essays consider different aspects of why and how capital move between countries, but approach the issue from quite different angles. The first essay discusses aggregated forces that determine industry structures and competition in countries, the second governments' incentives to respect foreign investor rights, while the last focuses on the individual firm's internationalization strategy in production. Whereas the first essay is purely theoretical, the last two include empirical testing of the predictions generated by my models.

The first essay, "Capital constraints, trade and the crowding out of southern firms", follows the tradition of the influential papers that won Paul Krugman his Nobel memorial prize in economic sciences. By introducing financial constraints into this framework I can determine how large a share of the population in each country has the possibility to start up a company, and thus how industrialized the country will be. When goods can be traded internationally, the relatively more industrialized country will have a higher total income, and more goods will be produced locally in this country. Both of these effects mean that the social welfare in this country will be higher than in the less industrialized country. The higher number of local firms, however, also increase competition between the firms. When firms have the possibility to move production to another country, there are thus two opposing factors at work: the larger demand in the most industrialized country makes it more attractive, whereas the larger number of competitors weighs towards making it less attractive. If trade becomes less costly, this affects the last of these factors. Cheaper goods from abroad increases competitive pressure, but since there are more firms in the most industrialized country, this effect is stronger in the least industrialized country. This means that lower trade costs increase the incentives for firms to move from the less to the more industrialized country. I discuss some important factors for when this phenomenon will increase or decrease welfare in the two countries.

The second essay, "Democracy and expropriations", considers govern-

ments' incentives to expropriate foreign investments. Sometimes a country depends on the technology of foreign firms in order to realize potential investment opportunities, especially in extractive sectors when moving into the less accessible reservoirs of these resources. However, once a mine has been dug, or an oil field has been developed, a lot of the costs have been sunk, whereas the revenues are still to come. This represents an opportunity for the local government to expropriate the investment and reap the benefits without having to bear the costs. Usually this would lead to a loss of reputation, making it more costly to attract future investors. This trade-off between short-term gains and long-term costs is essential for the government when deciding whether to expropriate or not. I show how the political system in a country can lead to discrepancies between policy makers and the voters in a country. Using data from actual expropriations over the last decades, I find that expropriations are least likely in countries with either very low or very high degree of political competition for office. My interpretation of this is that when government faces little or no risk of losing power, their horizon becomes more long-term, and more in line with voters. For the most democratic countries the effect at work is that the political competition forces the government to obey the wishes of the voters in order to stay in power, thus also acting in accordance with voter preferences. For the countries with intermediate levels of political competition governments have a shorter time horizon than voters, but not enough democratic discipline to force them to obey the voters' preferences.

In the last essay, "Productivity, size and the disintegration of industrial production", I discuss how firms organize their production in order to maximize profits. Taking a large investment cost by establishing a production unit in a foreign country may sometimes reduce unit costs of production, and thus making the firm more profitable. Similarly, firms will consider different options of producing inputs themselves, or buying them from independent providers, either domestically or abroad. My contribution with this paper is to show how these decisions may lead a firm to source inputs through several different channels simultaneously. I also show how these decisions are connected with the productivity levels of the firms. Evidence from Spanish manufacturing firms' behaviour supports the predictions from my model of the mapping of sourcing strategies as functions of productivity.

Contents

I

Chapter 1

Introduction

Introduction

"Theory is the poetry of science. It is simplification, the essential abstraction, the exaggeration of truth. Through simplification theory creates a caricature of reality. Through deduction the premises of the caricature are translated into empirical - and therefore refutable - generalizations." Mundell (1964: p. 421)

The history of trade is as old as the history of mankind. Even international trade can be traced as far back as around 3500 B.C. The earliest 'multinational enterprises' were based on trade, and took the shape of trading companies.[1] Probably the first of these appeared in the Assyrian kingdom as early as 2000 B.C. (Moore and Lewis, 1999). More recent examples are the English and Dutch East India Companies from 1600 and 1602, respectively. These companies, although they started out as pure trading houses, developed into vertically integrated multinationals more like the ones we see today. According to Carlos and Nicholas (1988), the Dutch East India Company opened a salpeter factory in Bengal in 1641, textile print works in the same place ten years later; by 1717 it employed more than four thousand silk spinners in Kaimbazar. This shows an interesting development in the nature of multinational enterprises. The motives for the multinational activity were originally of a resource-seeking character, as spices and silk had to be sourced where they were available, but later the decisions to refine these natural resources locally were driven by cost-reduction incentives. These motives for multinational activities are still widespread today, and are the main drivers behind

---

[1] It can be argued that the term 'multinational' would be somewhat misleading, since the notion of nation states is usually considered to have its origin with the peace of Westphalia in 1648. However, these trading companies did have offices in areas under different political jurisdictions, and can as such be argued to be the forefathers of modern multinationals.

vertical multinationals. Resource-seeking oil and mining companies are among the biggest multinational enterprises in the world, and a number of manufacturing firms produce parts or undertake assembly tasks in low-wage countries. Another more recent driver of foreign direct investments (FDI) is market seeking. Partly to avoid trade barriers such as quotas and tariffs and partly to reduce transport costs, some firms decide to employ a horizontal multinational structure where the same production process is undertaken simultaneously in several locations, and the output is sold locally. Where vertical multinationals seek cheap inputs, often in less developed countries, horizontal multinationals seek large markets, usually in larger, more developed countries.

From an academic point of view these different kinds of FDI pose a number of interesting questions as to which factors determine the direction and magnitude of such flows, and this dissertation is an aim to shed light on some of these. In this introduction I will briefly discuss some of the central contributions to the international economics literature, and then give a short description of how the different chapters in this dissertation fit into this body of literature and in which ways they contribute to the understanding of the different determinants of international capital flows.

1 International capital flows in economic literature

International capital flows were for a long time the problem child of international economics. Prior to 1980 the literature on international trade largely ignored the aforementioned drivers behind FDI, and viewed international capital movements purely within the framework of traditional capital theory, where FDI was no different from the flow of portfolio capital. In this line of reasoning, Robert Mundell (1957) introduced capital mobility to the classical Heckscher-Ohlin model of trade and found it to be a substitute to trade in commodities, implying that factors would flow from the country where they were abundant to the country where they were scarce if there where impediments to trade in commodities. This was perhaps not very surprising as in the Heckscher-Ohlin model it is the factors "embodied

in the goods" that are in fact traded. Similarly, the neoclassical growth models predicted that capital would flow from capital-rich developed countries to capital-poor developing countries. However, as Robert Lucas (1990) pointed out, this is not the case. Indeed, most foreign direct investments flow between the already capital-rich countries (Markusen, 2002).[2] In his seminal paper, Lucas points out several potential explanations for why FDI flows in the real world do not seem to follow the predictions from classical trade theory and neoclassical growth models. The failure of these models to correctly predict investment flows should, however, not be used to completely falsify the basis of the models. There are still strong indications that factor endowments contribute in shaping the market structure of a country, which again may determine trade and factor flows. As Lucas pointed out, there are other mechanisms that affect how attractive it is to invest in a country, and I will argue that there exist many more than the ones Lucas discusses.

The reason traditional trade theory displays major shortcomings when it comes to explaining capital flows is mainly its macroeconomic approach to the firm. In the following I briefly discuss how international economics "went micro" and developed a more sophisticated modeling framework for explaining international flows of goods and capital.

1.1 The new trade theory

Almost two hundred years after Ricardo wrote his *Principles*, economists are still trying to explain why countries trade with each other, and what the effects of international trade are. Up until the 1980s the theory of international trade was dominated by variations on the Ricardian- and Heckscher-Ohlin-type of trade models. Trade was either a result of differences in either technology or in relative factor endowments.

Although these models provide an elegant general equilibrium approach to international trade that generates important insights into the phenomenon, their shortcomings became clearer as time went by. Helpman and Krugman (1985)

---

[2]Developing countries as a whole are net recipients of FDI, but receive only between 20 and 30% of total inward FDI each year. Further, the least developed countries, the most capital-scarce countries in the world attract very little FDI, with a ratio of $\frac{\text{share of world inward FDI}}{\text{share of world GDP}} \approx 0.5$ (Markusen, 2002)

identified four areas where traditional trade theory seemed inadequate: it failed to explain volume of trade, composition of trade, FDI and intra-firm trade, and the welfare effects of trade. The traditional trade theory had limited room for multinational enterprises in its perfect competition approach to the world. In defense of traditional trade theory, it should be argued that one of the main reasons for its shortcomings was that the importance of multinationals grew significantly in the second half of the twentieth century, and thus well after the development of these theories.

Either way, it became ever more evident that the theory of international trade needed a new 'paradigm'. Enter the industrial organization approach to international trade. One of the first steps was to reconsider the perfect competition assumption. Economies of scale could easily explain some of the intra-industry trade that was observed between similar countries, which should have limited bilateral trade according to the classic theories.

Paul Krugman's seminal 1979 paper has largely been regarded as the start of the New Trade Theory (NTT). This new approach was based on increasing returns to scale in production within each firm as well as a form of the Spence-Dixit-Stiglitz model of monopolistic competition (Spence, 1976; Dixit and Stiglitz, 1977), and Samuelson's 'melting iceberg' trade costs (Samuelson, 1954). Introducing love of variety into the consumers' utility functions, Krugman was able to formally show how this could generate trade, and gains from trade between countries that were identical in consumer tastes, technology and factor endowments.[3] Over the following decade a number of papers were produced along these same lines, with Krugman as one of the dominating figures. His 1980 and 1981 papers extended the original model to include transportation costs and factor endowment differences, respectively. The latter paper went a step towards reconciling the new trade theory with the traditional Heckscher-Ohlin approach, but Krugman showed that instead of necessarily reducing trade, more similar factor endowments would shift trade from an inter-industrial to an intra-industrial character. Wilfred Ethier

---

[3]It should be noted that Bertil Ohlin had discussed increasing returns to scale as a source for trade already in his *Interregional and International Trade* in 1933 (Ohlin, 1933), but this had been largely ignored in the literature and in textbooks until Krugman's paper (Krugman, 1999). Krugman was also first in formally modeling increasing returns to scale at the firm level in this setting.

(1982) took the homogeneous good assumption from traditional trade theory, and introduced a constant elasticity of scale production function in intermediate inputs, and was able to reproduce some of the findings from the new trade theory in a more traditional framework. Helpman and Krugman (1985) integrated much of this new literature in their influential book that influenced the teachings of international trade for many years. This importance of increasing returns in trade theory helped reconcile the theory with some of the empirical evidence, but it would take another decade before the literature took an 'obvious' step further.

1.2   The new economic geography

If there are increasing returns to scale for certain productive factors, this would naturally cause a 'pull' effect towards locations where these are concentrated. In the aforementioned literature this was not considered, as factors were merely assumed to be immobile. Krugman (1991) was again central in the new development. In his own words:

> "This observation is [...] obvious in retrospect; but it certainly took me a while to see it. Why exactly I spent a decade between showing how the interaction of transport costs and increasing returns at the level of the plant could lead to the 'home market effect' (Krugman, 1980) and realizing that the techniques developed there led naturally to simple models of regional divergence (Krugman, 1991) remains a mystery to me. The only good news was that nobody else picked up that $100 bill lying on the sidewalk in the interim (Krugman, 1999)."[4]

The knowledge that people and production have a tendency for agglomeration was by no means a new topic. Already in 1920 Alfred Marshall had argued that technological spillovers between firms would lead to agglomeration, but these new models around 1990 were the first to formally show how scale and market effects could also cause agglomeration, and this new trend was termed the New Economic

---

[4]Whether or not nobody had picked up the $100 bill is probably a topic for debate, as the February 1988 issue of *Regional Science and Urban Economics* included papers explaining spacial agglomeration by increasing returns to scale and monopolistic competition from Abdel-Rahman (1988), Fujita (1988), and Rivera-Batiz (1988).

Geography (NEG). These trade models with internationally mobile factors finally took a step towards explaining some of the observed trends in international capital movements, such as why capital does not necessarily flow to the capital-scarce countries as Mundell (1957) had predicted. A number of variations on Krugman's 'Core-Periphery' model, and later Martin and Roger's (1995) 'Footloose Capital' model, explained agglomeration forces under a number of different scenarios, both through market size effects, externalities in intermediate input production, and forward and backward linkages. This early stage of the literature is nicely summed up in Ottaviano and Puga (1998).

It took another decade before the international economics literature saw the next important step forward. Marc Melitz's (2003) influential paper introduced firm heterogeneity into Krugman's (1980) trade model, and showed how more productive firms would self-select into exporting. This was the next step towards placing the individual firm at the center of attention in trade theory, a trend that quickly spread to the NEG literature. Helpman et al. (2004) extended Melitz's model to include the option of FDI, and Melitz and Ottaviano (2008) refined the model further to discuss different trade policies. Baldwin and Okubo (2006) introduced a Melitz-type heterogeneity into the footloose capital model, and showed that agglomeration forces would affect the more productive firms more strongly, and thus that one should expect the more productive firms to concentrate in the larger market.

With this development over the last 30 years international trade theory has been able to explain a lot more of the observed flows of FDI and its interaction with trade in commodities. The other important aspect of the development of trade theory over this time was the shift towards the industrial organization literature. However, the role of multinational firms in trade models was still mostly to explain trade patterns, and partly factor movements, whereas the real motives of the firms themselves to undertake FDI were still largely ignored among trade economists. More or less at the same time, academics working more within the industrial organization school were also studying the phenomenon of multinational enterprises, and were making progress towards answering this question.

1.3 The international organization of the firm

The alternative approach to the topic of FDI differed from the trade approach first and foremost in that it started from within the firm. In his 1971 paper Caves discussed intuitively, some of the reasons firms with specific characteristics choose either vertical or horizontal FDI.[5] He also ventured into the discussion of how the multinational behavior of firms was interconnected with market characteristics, thus taking the discussion into an industrial organization framework. Later, John Dunning (1977, 1988, 1993) developed the "OLI" framework to explain the existence of multinational firms. Although not a formal theory, this eclectic approach proved fruitful for economists in industrial organization trying to explain FDI. 'OLI' is named for the three potential advantages that may motive a firm to go multinational; Ownership, Location, and Internalization. The first of these, ownership, refers to certain advantages that make FDI an optimal decision for this firm, but not for others. Although a multidimensional variable, it is usually treated as the firm's overall productivity. Location focuses on characteristics of the host country that receives the FDI, generally its market demand for horizontal FDI and low factor costs for vertical FDI. The last, internalization, deals with the decision of establishing an integrated subsidiary abroad, versus dealing with an independent partner through arm's-length dealings. In this last strand the focus has mainly been on possible spillovers to local competitors, and on contract incompleteness between independent partners in vertical dealings.

Many formal contributions to this literature have had to focus on one or maybe two of Dunning's three legs for FDI. Many also use traditional trade models to frame the setting they are describing. One example of this is Helpman's (1984) model for vertical FDI in a general equilibrium, Heckscher-Ohlin framework. A similar model was presented by Markusen (1984). In both models firms produce some firm-specific headquarter service that can be used in production facilities in other countries. This was the start of a number of models, later grouped together as the 'Knowledge-Capital model', that identified intangible assets, such as technological advantages, as a driver for FDI. Yeaple (2003) extended these models

---

[5]More specifically Caves discussed traits "intrinsic to the industry", but qualitatively the arguments are the same as when focusing on the individual firm.

to include 'complex integration'; when firms choose to become multinational both along the horizontal axis, by producing locally in other developed countries, and vertically to take advantage of lower factor prices in developing countries.

All of these models, as well as many others in the same line of research, thus consider ownership and location advantages, but leave the internalization question out of the model by simply assuming away all other options, like most other traditional trade models. Ethier (1986) recognized that trade theory had answered many of the important questions around multinational firms, but argued that the internalization decision was still a 'black box,' and identified the task of explaining it as the "central task of general equilibrium theory of the multinational corporation". Motivated by the public good characteristics of technology, his model showed how firms may choose FDI to internalize the principal-agent problem with arm's-length dealing, and thus also costs associated with the incomplete contracts environment.

The contributions within this OLI-framework over the following decades include the widely used textbook by Markusen (2002) as well as seminal papers from Markusen (1995), Markusen and Venables (1998), and many others. Grossman and Helpman (2002) and Antrás (2003) introduced the Grossman-Hart-Moore (Grossman and Hart, 1986; Hart and Moore, 1990) approach to contract incompleteness and firm integration into a Helpman-Krugman type of trade environment to explain the internal motives for firms to become multinational. This branch of the literature is also considered a part of the new trade theory, and is as such closely related to many of the works that are referred to above. It was also brought an important step forward by the new way of modelling firm heterogeneity introduced by Melitz (2003). Taking a similar approach to the one used by Helpman et al. (2004), but from a cost-minimizing point of view, Antrás and Helpman (2004) show how firm productivity determines sourcing strategies for the firms, identifying four types of firms; foreign and domestic sub-contracting firms, and foreign and domestic integrating firms. This model became a pioneering model for theoretical work on the international organization of the firm, and hence also a very important part of understanding international capital flows, and was followed by other important contributors as Antràs (2005), Acemoglu et al. (2007), and others.

Today's theoretical literature on foreign direct investment has thus come from

two strands: the trade literature, taking an outside-in approach to the topic, and the industrial organization literature, taking an inside-out approach. Together these two aspects have been able to give a more complete picture of factors within the firm driving certain internationalization decisions, while simultaneously maintaining the general equilibrium perspective from trade theory that is important in order to understand the mechanics that generate the conditions under which the firms make their decisions. The chapters in this dissertation fit in, and contribute to, different parts of this literature.

2 Chapter summaries

In this section I will give brief summaries of the different chapters in this dissertation, and relate them to the literature discussed up until now.

2.1 Capital constraints, trade and crowding out of southern firms

This chapter concentrates on the mobility of factors, or more specifically on capital mobility. I mentioned above the large literature around Martin and Roger's (1995) 'Footloose capital' model and its extensions. Several of these models operate with identical countries, except for the fact that "one region for some reason has a larger manufacturing sector" (Krugman and Venables, 1995: p. 861). In my model this asymmetry is generated by differences in the strengths of the countries' financial systems, which ensures that capital is more efficiently allocated in the more developed country, meaning that more potential entrepreneurs are actually able to establish firms, and the country will be more industrialized. The number of firms in this model is not determined by a zero cutoff condition, as in many similar models. This means that existing firms may earn positive profits. In spite of this, the setup replicates many of the results from other contributions in the 'new economic geography' literature. One of these results is that the 'core' country will have a more than proportionate share of manufacturing firms. Whereas in Krugman and Venables (1990) increasing profits in the core lead new firms to enter the market here, and thus increase the share of manufacturing firms located

in this country, it is firms reallocating from the periphery to the core in my model that generate the same result. In both cases, it is the attractiveness of producing close to the larger market that drives this agglomeration. Unlike many of the aforementioned models, however, both countries start out as equally sized in my model. Both labor and capital endowments, and thus also the relative endowments, are identical in the two countries. It is the better allocation of financial capital in the 'north' that converts this into a larger amount of real capital in my model. This again increases total production, and thus also the market size.

However, the main contribution of this paper is not along the lines of firm agglomeration as such. Introducing inter-generational savings *à la* Banerjee and Newman (1993) , I show that the long-term number of firms is determined not only by credit constraints, but also by equilibrium profits, thus reincorporating a similar effect as a zero cutoff condition; falling profits drive down the equilibrium number of firms. This new mechanism causes a link between competition from integrating international markets and the degree of credit constraints in a country. I find that in equilibrium, internationalization will lead to a deindustrialization of the 'south,' as credit constraints will become binding for a larger share of the population, as competition from northern firms drive down profits. This leads to a drop in national income. Contrary to many other models in the literature, internationalization may thus lead to a drop in real wages, and also in general welfare in the less developed country in this model, compared to the autarky equilibrium.

2.2   Democracy and expropriations

In his famous paper on why capital does not flow from rich to poor countries, Lucas (1990) points to capital market imperfections as one of several potential explanations. The imperfections he had in mind were, however, of a very different kind than the ones discussed in Chapter 1 of this dissertation. Lucas referred to his concept of capital market imperfections as "political risk," and described a macroeconomic situation where foreign investments can be seen as a loan which is repaid with future flows of goods. With limited enforceability of contracts in this international climate this will easily end up in a classic 'hold-up' situation.

Chapter 2 in this dissertation addresses the hold-up problem of foreign invest-ments and expropriation risks. Numerous contributions have been made on the topic of foreign investments and the hold-up problem in the literature on contract theory, the most influential of which are probably Eaton and Gersovitz (1984), Thomas and Worrall (1994); also, Schnitzer (1999), Konrad and Lommerud (2001) and Engel and Fischer (2008) have given important insights into this problem of political risk. This contract theory approach to the hold-up problem of foreign investment generally treats the host country as a united, coherent and rational unit. I argue that this is not the case in many of the countries that have expro-priated foreign investment over the last decades. I therefore introduce political competition and an internal struggle for resources within the country to enrich the framework of my theory.

The chapter consists of both a theoretical model and empirical testing of the model's predictions. The modeling framework takes the form of a 'social conflict'-type election model along the lines of Padró I Miquel (2007) and Beasley and Ku-damatzu (2007), where both voters and candidates are divided into social groups. Rulers will enforce policies that benefit their own group, and voters from this group may thus prefer to support a poorly performing candidate from their own group to the risk of getting a ruler from the opposing group. In my model the performance measure is whether the president will expropriate foreign investments or not. For-eign investors act as in the traditional hold-up problem, and will be cautious to invest due to uncertainties concerning the ruler's propensity to expropriate. If he expropriates, he will reveal his willingness to do so, which will lead to a drop in future investments and thus future income and welfare in the country. Rulers have, due to the risk of being ousted from power, a different time perspective than voters, and may rationally want to expropriate foreign investments to increase short-term income. This time-inconsistency problem is affected by the political institutions in the country as well. Stronger political competition increases the risk of los-ing power, which increases incentives to expropriate today, whereas it may also make voters more responsive to the ruler's actions, thus reducing expropriation incentives.

The political institutions in the country also include an aspect of division of power and political constraints. The country may have political bodies that can

veto decisions by the executive power, and if any of these are controlled by the opposing social group, it will rationally do so when a ruler moves to expropriate a foreign investment. By dividing the concept of democracy in these two dimensions I can discuss political risks along both of them.

Probably the most striking result from this model is that when letting political constraints be constant, increasing political competition will initially increase the risk of expropriation as the ruler will discount future income more due to the increased risk of being ousted from power. However, for sufficiently high levels of political competition the voters' responsiveness to expropriation will start to dominate the net effect, and further increases in political competition will reduce expropriation risks.

In the last part of the chapter I test these predictions, as well as other predictions from the model, on data on actual expropriations of mineral mines in a set of 27 developing countries. This testing shows that expropriations are on average, more likely in newly independent countries with high government expenditure, weak integration into international trade, and weak political constraints. The data also exhibit a non-linear relationship between political competition and the expropriation risks, where competition has a positive but concave effect on the probability of expropriation. For political competition levels that correspond to the top 20% of my observations, increased competition reduces expropriation risks, as predicted by the model.

2.3 Productivity, size and the disintegration of industrial production

The previous two chapters in this dissertation show firms' incentives to invest in foreign countries as driven by market access (chapter 1) or resource-seeking (chapter 2). In this chapter I take on a third motivation for firms to invest abroad: cost reduction. The chapter does not limit the focus to foreign investments, as I develop a sourcing model where firms may acquire intermediate inputs from vertically integrated plants, or through arm's-length trading, either domestically or from abroad. This chapter is thus a step towards an increased microeconomic focus of the firm.

Increased global integration both through trade in goods and throug easier communication has lowered costs of 'slicing up the value chain.' While the River Rouge factory in Michigan received ore and delivered finished Model A Fords (Winnewisser, 2005), the typical "American" car is now produced in the United States (37%), Korea (30%), Japan (17.5%), Germany (7.5%), Taiwan (4%), Singapore (4%), the United Kingdom (2.5%), Ireland (1.5%), and Barbados (1.5%) (Antràs and Helpman, 2004). This, and other anecdotal evidence point to a trend that the production process is increasingly split up, and undertaken at different locations. However, at the same time there seems to be a great degree of heterogeneity between firms, even within the same industry, as to how they choose to organize their production chain. Grossman and Helpman (2002), Antràs (2003), Antràs and Helpman (2004, 2008), Grossman and Rossi-Hansberg (2008), and others have provided valuable contributions by showing how productivity, and contractibility may affect the sourcing behavior of firms. These models, however, often suffer from frameworks that permit firms to choose only one sourcing option. In combining the productivity dimension of Antràs and Helpman (2004) with the tradeable task dimension from Grossman and Rossi-Hansberg (2008), I am able to construct a model that permits the simultaneous use of a range of sourcing options.[6]

I use a variation on the Melitz model to generate a set of heterogeneous firms that differ in their productivity levels. More productive firms are, all else equal, larger, both in terms of output and in their use of input, and also enjoy higher profits. When organizing their production process, firms weigh the total costs associated with each individual task for each possible production organization of that task, and choose the one that minimizes costs. When there are fixed costs associated with starting production of each individual task, either through search costs for finding an appropriate supplier or through building a production plant, the optimal decision of weighing variable costs against fixed costs will differ between firms with different productivity levels. This generates testable predictions about firms' optimal sourcing strategies as functions of their productivity. While reproducing the findings from Antràs and Helpman (2004, 2008) that the least pro-

---

[6]In the Grossman and Rossi-Hansberg (2008) model firms may undertake tasks both domestically and abroad simultaneously, but their model takes a more aggregated approach, and does not capture the productivity dimension.

ductive firms source domestically, whereas the most productive undertake foreign direct investments and produce in vertically integrated plants abroad, my model generates further predictions concerning the intensities with which each sourcing option is used as functions of productivity. Also, it predicts an ordering of sourcing strategies as different combinations of sourcing options.

Using firm-level survey data from Spanish manufacturing firms, I am able to test the predictions from the model, and overall I find strong empirical support for the theoretical model.

References

Acemoglu, D., Antràs, P., & Helpman, E. 2007. Contracts and Technology Adoption. *American Economic Review*, **97**(3), 916–943.

Adbul-Rahman, H. 1988. Product Differentiation, Monopolistic Competition and City Size. *Regional Science and Urban Economics*, **18**(1), 69–86.

Antràs, P. 2003. Firms, Contracts, and Trade Structure. *Quarterly Journal of Economics*, **118**(4), 1375–1418.

Antràs, P. 2005. Incomplete Contracts and the Product Cycle. *American Economic Review*, **95**(4), 1054–1073.

Antràs, P., & Helpman, E. 2004. Global Sourcing. *Journal of Political Economy*, **112**(3), 552–580.

Antràs, P., & Helpman, E. 2008. Contractual Frictions and Global Sourcing. *In:* Helpman, E., Marin, D., & Verdier, T. (eds), *The Organization of Firms in a Global Economy*. Harvard University Press: Cambridge, MA.

Baldwin, R., & Okubo, T. 2006. Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting. *Journal of Economic Geography*, **6**, 323–346.

Banerjee, A., & Newman, A. 1993. Occupational Choice and the Process of Development. *Journal of Political Economy*, **101**(2), 274–298.

Besley, T., & Kudamatsu, M. 2007. *Making Autocracy Work*. mimeo. London School of Economics.

Carlos, A., & Nicholas, S. 1988. "Giants of Earlier Capitalism": The Chartered Trading Companies as Modern Multinationals. *The Business History Review*, **62**(3), 398–419.

Caves, R. 1971. International Corporations: The Industrial Economics of Foreign Investment. *Economica*, **38**(149), 1–27.

Dixit, A., & Stiglitz, J. 1977. Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, **67**(3), 297–308.

Dunning, J. 1977. Trade, Location of Economic Activity and the Multinational Enterprise. *In:* Ahlin, B., Hesselborn, P., & Wijkman, P. (eds), *The International Allocation of Economic Activity*. London: MacMillan.

Dunning, J. 1988. The Eclectic Paradigm of International Production: A Restatement. *Journal of International Business Studies*, **19**(1), 1–32.

Dunning, J. 1993. Trade, Location of Economic Activity and the Multinational Enterprise: A Search for an Eclectic Approach. *In:* Dunning, J. (ed), *The Theory of Transnational Corporations*. London: Routledge.

Eaton, J., & Gersovitz, M. 1984. A Theory of Expropriation and Deviations from Perfect Capital Mobility. *The Economic Journal*, **94**(373), 16–40.

Engel, E., & Fischer, R. 2008. *Optimal Resource Extraction Contracts under Threat of Expropriation*. NBER Working Paper Series 13742. National Bureau of Economic Research.

Ethier, W. 1982. National and International Returns to Scale in the Modern Theory of International Trade. *American Economic Review*, **72**(3), 389–405.

Ethier, W. 1986. The Multinational Firm. *Quarterly Journal of Economics*, **101**(4), 805–833.

Fujita, M. 1988. A Monopolistic Competition Model of Spatial Agglomeration: Differentiated Product Approach. *Regional Science and Urban Economics*, **18**(1), 87–124.

Grossman, G., & Helpman, E. 2002. Integration versus Outsourcing in Industry Equilibrium. *Quarterly Journal of Economics*, **117**(1), 85–120.

Grossman, G., & Rossi-Hansberg, E. 2008. Trading Tasks: A Simple Theory of Offshoring. *American Economic Review*, **98**(5), 1978–1997.

Grossman, S., & Hart, O. 1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, **94**(4), 691–719.

Hart, O., & Moore, J. 1990. Property Rights and the Nature of the Firm. *Journal of Political Economy*, **98**(6), 1119–1158.

Helpman, E. 1984. A Simple Theory of International Trade with Multinational Corporations. *Journal of Political Economy*, **92**(3), 451–471.

Helpman, E., & Krugman, P. 1985. *Market Structure and Foreign Trade*. Cambridge: MIT Press.

Helpman, E., Melitz, M., & Yeaple, S. 2004. Export Versus FDI with Heterogenous Firms. *American Economic Review*, **94**(1), 300–316.

Konrad, K., & Lommerud, K. 2001. Foreign Direct Investment, Intra-Firm Trade and Ownership Structure. *European Economic Review*, **45**, 475–494.

Krugman, P. 1979. Increasing Returns, Monopolistic Competition, and International Trade. *Journal of International Economics*, **9**(4), 469–479.

Krugman, P. 1980. Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review*, **70**(5), 950–959.

Krugman, P. 1981. Intraindustry Specialization and the Gains from Trade. *Journal of Political Economy*, **89**(5), 959–973.

Krugman, P. 1991. Increasing Returns and Economic Geography. *Journal of Political Economy*, **99**(3), 483–499.

Krugman, P. 1999. Was it all in Ohlin? *Unpublished notes*.

Krugman, P., & Venables, A. 1990. Integration and the Competitiveness of Peripheral Industry. *In:* Bliss, C., & Braga de Macedo, J. (eds), *Unity with Diversity in the European Community*. Cambridge: Cambridge University Press.

Krugman, P., & Venables, A. 1995. Globalization and the Inequality of Nations. *Quarterly Journal of Economics*, **110**(4), 857–880.

Lucas, R. 1990. Why Doesn't Capital Flow from Rich to Poor Countries? *American Economic Review*, **80**(2), 92–96.

Markusen, J. 1984. Multinationals, Multi-Plant Economies, and the Gains from Trade. *Journal of International Economics*, **16**(3-4), 205–226.

Markusen, J. 1995. The Boundaries of Multinational Enterprises and the Theory of International Trade. *Journal of Economic Perspectives*, **9**(2), 169–189.

Markusen, J. 2002. *Multinational Firms and the Theory of International Trade.* Cambridge: The MIT Press.

Markusen, J., & Venables, A. 1998. Multinational Firms and the New Trade Theory. *Journal of International Economics*, **46**(2), 183–203.

Marshall, A. 1920. *Principles of Economics.* London: MacMillan.

Martin, P., & Rogers, C. 1995. Industrial Location and Public Infrastructure. *Journal of International Economics*, **39**, 335–351.

Melitz, M. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, **71**(6), 1695–1725.

Melitz, M., & Ottaviano, G. 2008. Market Size, Trade, and Productivity. *Review of Economics Studies*, **75**, 295–316.

Moore, K., & Lewis, D. 1999. *Birth of the Multinational: 2000 Years of Ancient Business History - From Ashur to Augustus.* Copenhagen: Copenhagen Business School Press.

Mundell, R. 1957. International Trade and Factor Mobility. *American Economic Review*, **47**(3), 321–335.

Mundell, R. 1964. A Reply: Capital Mobility and Size. *The Canadian Journal of Economics and Political Science*, **30**(3), 421–431.

Ohlin, B. 1933. *Interregional and International Trade.* Cambridge: Harvard University Press.

Ottaviano, G., & Puga, D. 1998. Agglomeration in the Global Economy: A Survey of the 'New Economic Geography'. *The World Economy*, **21**(6), 707–731.

Padró i Miquel, G. 2007. The Control of Politicians in Divided Societies: The Politics of Fear. *Review of Economic Studies*, **74**, 1259–1274.

Ricardo, D. 1817. *On the Principles of Political Economy and Taxation*. London: John Murray.

Rivera-Batiz, F. 1988. Monopolistic Competition, and Agglomeration Economies in Consumption and Production. *Regional Science and Urban Economics*, **18**(1), 125–153.

Samuelson, P. 1954. The Transfer Problem and Transport Costs, II: Analysis of Trade Impediments. *Economic Journal*, **64**, 264–289.

Schnitzer, M. 1999. Expropriation and control rights: A dynamic model of foreign direct investment. *International Journal of Industrial Organization*, **17**, 1113–1137.

Spence, M. 1976. Product Selection, Fixed Costs, and Monopolistic Competition. *Review of Economic Studies*, **43**(2), 217–235.

Thomas, J., & Worrall, T. 1994. Foreign Direct Investment and the Risk of Expropriation. *Review of Economic Studies*, **61**, 81–108.

Winnewisser, P. 2005. *The Legendary Model A Ford: The Ultimate History of One of America's Great Automobiles*. Iola, WI: KP Books.

Yeaple, S. 2003. The Complex Integration Strategies of Multinationals and Cross Country Dependencies in the Structure of Foreign Direct Investment. *Journal of International Economics*, **60**(2), 293–314.

Chapter 2

Capital Constraints, Trade and the Crowding Out of Southern Firms

II

Capital Constraints, Trade and Crowding Out of Southern Firms

Jonas Gade Christensen*

University of Bergen

Abstract

Introducing capital market imperfections to a 'footloose capital' model, I show how such distortions may explain the observed phenomena of an industrialized north and an underdeveloped south. Further, I show that with inter-generational savings internationalization will cause a crowding out of manufacturing firms in the south, increasing the share of the southern population that are credit-constrained, and also reducing total income in the country. This should not, however, be taken as an argument for protectionism, as welfare may indeed be higher with trade than in autarky, if trade costs are sufficiently low.

1   Introduction

Globalization is much more than just the increased movement of goods between countries. Although the economic literature has traditionally had a much stronger focus on trying to explain international trade in goods, there has over the last decades grown a more vivid debate around the phenomenon of foreign direct investment (FDI), and rightly so. Until it dropped as a consequence of the international financial crisis in 2007, FDI had significantly outgrown international trade in goods over the last decades (UNCTAD, 2008). This aspect of globalization has been studied on many different levels, from balance of payments and capital account issues, to micro-level labor market determinants and effects. Another very interesting question in this literature is the interplay between international trade and foreign direct investments. Although often ignored, this interaction is not new to the literature. Mundell's (1957) modification of the Heckscher-Ohlin model predicted that factor flows were substitutes for trade in goods. In this model world, it is relative differences in factor endowments that drive incentives for both trade in goods and for factor flows, and Mundell showed that impediments to trade would increase factor movements, and vice versa. Trade and factor mobility were substitutes in the sense that the more mobile of the two would always work towards equalizing the factor prices. However, real-world observations have shown that relatively little capital moves from capital-rich countries to capital-poor countries, and if the model were true, Prebisch's (1950) observation of industrialized 'core-countries' selling their manufactured goods in exchange for primary goods from the underdeveloped 'periphery' would not be the result. Recently Antràs and Caballero (2009) reapproached this question. Introducing financial frictions into a Heckscher-Ohlin-Mundell model of trade, they find that "in less developed economies (South), trade and capital mobility are compliments in the sense that trade integration increases [...] the incentives for capital to flow to South." I argue that the Heckscher-Ohlin framework is not very well suited for discussing such effects of trade integration on factor movements. The reason for this is that in these models one generally observes only one-way trade flows in each sector, which means that a trade liberalization in one sector is in fact a unilateral lowering of trade costs, which may lead to different conclusions than when trade integration

happens through a bilateral lowering of trade costs. When trade liberalization is *de facto* an increased market access for firms in one country without a reciprocal compensation, these firms will gain an advantage, and locating in this country will be more attractive than before the trade liberalization. For competing firms in the other country, the effect will be exactly the opposite, and firms will thus have incentives to relocate to the country that has gained increased market access to the other country. Another reason I find this kind of unilateral trade integration less interesting is the fact that trade integration over the last century has been dominated by multilateral liberalization through GATT and the WTO, and regionalism (NAFTA, EU and others), while unilateral trade liberalization has mainly been associated with developing countries opening up their economies for trade and investment "due to the demise of the socialist model" (Janeba and Schjelderup, 2003).

One of the 'stylized facts' about the globalized world is that there exist core- and periphery-countries, where the former mainly produce and export manufactured goods, while the latter specialize in commodities. Further, it has been well documented that while this specialization determines the relative net trade flows in each industry, there is also a significant amount of intra-industry trade, where different varieties of the same type of product are being traded in both directions. The 'new economic geography' (NEG) literature has lately shown mechanisms that may explain both of these 'facts.' Inter-regional models where labor is assumed to be mobile between regions has shown how a small difference in market sizes may start a process of concentration that "feeds on itself" through attracting more firms, which attract more workers, which again increases the differences in market sizes etc. until all production takes place at the same point in space (see for example Krugman, 1991). International trade models rarely assume labor to be mobile, but may still generate similar results. Krugman and Venables (1995) show how linkages between intermediate and final goods production can generate a circular process that leads to a core and a periphery in a similar way. In their model the periphery may gain or lose from globalization depending on the level of trade costs.

In this paper I combine and extend the aforementioned literature by building a model with both international trade and capital movements, as well as introduc-

ing another important factor to determine the degree of industrialization: credit constraints. Borrowing constraints have been argued to limit investments in both human and physical capital, and thus work as a major hindrance to economic growth in high-income countries, but more so in low-income countries. In Krugman and Venables (1995) the two countries are initially identical, except that "one region for some reason has a larger manufacturing sector." I argue that credit constraints may be one reason relative sector sizes may differ between two countries that are otherwise identical. In this paper I develop a theoretical model that shows how credit constraints may determine the initial degree of industrialization in a country, but also how they may interact with the effects of globalization. Especially I focus on how credit market imperfections in a developing country may lead to a deindustrialization of the country when it becomes more integrated in the global economy. My model also predicts that a trade integration will lead to a concentration of industry in the country with the more developed financial markets. However, this may lead to an increase or a decrease in international trade, depending on the initial situation. As such, trade and capital flows may both be complements or substitutes.

My basic 'workhorse' model replicates results from the NEG literature where sufficiently free trade will cause agglomeration of manufacturing production in one country.[1] In this simple north-south model the country with the better functioning financial markets will start out as more industrialized, and through this generate higher total income, making it a more attractive market, *ceteris paribus*. On the other hand, the more developed north will have more than its proportional share of firms, thus making competition harder than in the less developed south. For high levels of trade costs, this competition effect will dominate the market size effect, firms' profits will be larger in the less developed country, and firms will have incentives to move south. As trade costs fall firms start to experience competition from abroad, and the market size effect will gradually become more important, relative to the competition effect. For sufficiently free trade firms will prefer to locate in the more developed country, and there will be capital movements from the south

---

[1] In addition to the above cited papers by Krugman (1991) and Krugman and Venables (1995), see for example Ciccone and Hall (1996), Duranton and Overman (2005) and Baldwin and Okubo (2006) for theoretical and empirical contributions. For a comprehensive review of the agglomeration literature, see Puga (2010).

to the north. There also exists a level of trade costs, below which all firms in the model will locate in the developed country, and the developing country will be completely deindustrialized. In this basic version of the model the real wages in the south will always be higher in the international equilibrium than in autarky for all levels of trade costs, and thus also for trade costs that lead to complete deindustrialization in the south. However, this result depends crucially on the assumption that the number of firms is determined by the initial wealth distribution and is constant after that. Introducing intergenerational wealth dynamics into the model, I show that credit constraints and international competition will lead to a drop in the number of southern-owned firms. Tougher competition will cause a consolidation with fewer but larger firms, and this will happen through stricter credit constraints in the south, leading to a deindustrialization there. If trade is sufficiently costly, this may lead to a drop in real wages when going from autarky to a globalized equilibrium with trade and capital movements.

1.1   Previous literature

This paper draws on two strands of economic literature. On one hand it extends the work of Banerjee and Newman (1993), Ranjan (2001), Das (2006), Chesnokova (2007), and others who study how credit constraints determine sector sizes and industry structure, and how these again may be affected by international trade. The first one of these does not focus on trade at all, whereas the other three all assume two, homogeneous, internationally traded goods. This rules out intra-industry trade by assumption, which I will show can have important implications on policy implications derived from such trade models. None of the models consider capital mobility between countries. My model shows how international trade and international capital movements interact, and also how both of them may, individually and in combination, affect the degree of credit constraints in a country.

The other strand of relevant literature that I follow is the new trade theory, and new economic geography literature discussed above. More specifically my model is a variation on Dupont and Martin (2006), which again builds on the 'Footloose Capital' models by Martin and Rogers (1995) and Krugman (1991). These models describe the interplay between international trade and capital movements when

firms are mobile. Industry sizes are determined either exogenously by capital endowments or endogenously through some zero-profit condition. In this paper I introduce wealth inequalities in the populations, and capital market imperfections that determine the degree of industrialization in each country. I also show how, when capital market imperfections are present, long-term effects from globalization may change some of the findings from the static version of the model.

2   The model

In this section I develop a two-country model of international trade, with mobile capital, immobile agents, and imperfect credit markets. First I discuss a static version of the model, which could be interpreted as a short-term view of the world. In this version credit constraints determine the number of firms in each country, which again determine trade and capital flows. Later I introduce some wealth dynamics over generations, and look at the long-term effects when the wealth distribution in the population is determined by intergenerational saving. This opens up the possibility that international market outcomes feed back into the credit constraints in both countries, thus affecting the total number of firms in the long-term equilibrium. In the static model lower trade costs will lead to a concentration of the manufacturing industry in the developed north. The number of firms is constant and agents are immobile, however, and southern-owned firms will repatriate their profits to their owners in the south. Welfare in the south is affected negatively by having to cover trade costs on goods that used to be produced in the south, but are now produced in the north. On the other hand, all goods that are imported from the northwill now have lower trade costs. In sum the latter effect will dominate the former in this version of the model, and welfare in the south is higher in the globalized equilibrium than in autarky for all levels of trade costs. In the long-term version of the model globalization may lead to a drop in the number of southern-owned firms, and under this formulation globalization may cause immiserizing deindustrialization in the south with a drop in both real wages and overall social welfare. Welfare effects of trade liberalization are thus similar to those in the work of Brander and Krugman (1983), where an initial trade integration from full autarky may reduce welfare as long as wasteful transportation costs dominate

the gains from increased variety. Further liberalization will, however, reduce the loss from transport costs, and for sufficiently free trade welfare will again be higher than in autarky.

2.1 Basic setup

There are two countries, north and south, where south is denoted by an asterisk. Each country has a population of measure one, $L = L^* = 1$, and each individual in the population is endowed with some initial wealth W_i, distributed according to some cumulative distribution function $G(W)$, and one unit of labor which he or she supplies inelastically in the market. Income is spent to maximize the following utility function, subject to their budget constraint:

$$U = \frac{1}{\alpha^\alpha \beta^\beta} C_M^\alpha C_A^\beta, \ C_M = \left(\int_{\epsilon n^w} c_i^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}, \ \alpha + \beta = 1 \ , 1 < \sigma \qquad (1)$$

Here C_A is consumption of a traditional good, C_M is a consumption-bundle of manufactured varieties, and σ is the constant elasticity of substitution between varieties. I assume the common feature of costless differentiation, so that each firm in the manufacturing sector produces a unique variety, and $n + n^* = n^W$ thus denotes both the total number of firms in the world and the number of varieties produced.

The consumers' budget is determined by their initial wealth and their earnings. Each individual may choose between three occupations; being a worker in the traditional sector, being a worker in the manufacturing sector, or becoming an entrepreneur and starting up their own business. Labor is homogeneous, and workers can move freely between the sectors, meaning that any surplus labor supply from the manufacturing sector will work producing the traditional good, and wages will be fixed at the level determined by prices and productivity in this sector. Each agent then chooses the occupation that maximizes income, given his credit constraint.

Becoming an entrepreneur implies undertaking an initial investment I measured in units of labor, and supplying the unit of labor in administration. The latter implies that each individual may start at most one firm. For this setup to

be interesting, income from being a firm-owner must be higher than regular wages, which I assume to be the case: $\pi - I > w$. Imperfections in the credit market, however, mean that not all agents can become entrepreneurs. I follow the standard approach in the literature, and assume that individuals can only borrow some multiple $\delta > 1$ of their initial wealth. This multiplier is a function of a number of factors in the country, such as the contractual climate, borrowers' ability to use assets as collateral for loans, rule of law, political risks, etc., and it can be used as an aggregate indicator of the sophistication of the financial system in the country. Assuming that the north is the more developed country, I thus assume that $\delta > \delta^*$, meaning that individuals in the north are able to borrow a larger multiple of their own wealth. In order to become an entrepreneur, an individual must have enough initial wealth and access to loans to be able to undertake the initial investment,

$$\delta W \geq wI,$$

which again defines the cutoff value of initial wealth needed to become an entrepreneur:

$$\bar{W} = \frac{wI}{\delta}.$$

This shows that the minimum initial wealth required to become an entrepreneur in the north will be lower than in the south, and since initial wealth is identically distributed in the two countries, there will be more unconstrained agents in the north. Agents are immobile between countries, and a potential entrepreneur is subject to the financial environment in his home country. As long as profits from being an entrepreneur are still higher than normal wages in both countries, there will be more entrepreneurs in the north than in the south; $n > n^*$. In other words, even though both countries are endowed with equal amounts of financial capital, the allocation is more efficient in the north, and the north will be more physical capital-rich than the south. I will thus classify the north as the capital-rich country for the rest of this paper.

Both countries can produce both the traditional and the manufacturing good. I assume that the parameter values are such that the traditional product is produced in both countries in equilibrium. This sector exhibits a constant returns to scale technology, using only labor as input, and is traded costlessly across the

borders under perfect competition. This somewhat unrealistic but very convenient assumption ensures that the price of this good is equal in both markets, and with identical technologies, also causes wages to be equal in both countries. I make the common assumption in the literature that units and productivity in the traditional sector are such that wages are equal to the price of the traditional good, and use this as the numeraire, hence $w = P_A = 1 = P_A^* = w^*$.

The manufactured goods sector exhibits traditional Dixit-Stiglitz monopolistic competition. Varieties of the manufactured good is produced with increasing returns to scale, requiring an initial investment wI to set up a headquarters, which must be located in the home country, and a production unit which may be located in either country. The production process uses only labor as input. Specifically the total costs for a firm in the manufacturing sector is $w\left(I + x_i\right)$, where x_i is produced quantity by firm i.

I will first present a static version of the model where the initial wealth distribution determines the number of firms in each country. This version resembles most closely other economic geography models, and illustrates some central mechanisms in an intuitive way. Later I introduce dynamics into the model, and show how this changes some important results from the static version of the model.

2.2 Autarky

The utility function permits the use of two-stage budgeting. The first stage determines the optimal shares spent on each type of good:

$$C_A = \frac{\beta Y}{P_A}$$
$$C_M = \frac{\alpha Y}{P_M},$$

where Y is total expendable income, P_A is the price of the traditional good, and P_M is the price index for the manufactured goods.

Consider the set of consumed manufactured varieties as an aggregate good

$$C_M = \left(\int_{\epsilon n^w} c_i^{\frac{\sigma-1}{\sigma}} \, di\right)^{\frac{\sigma}{\sigma-1}},$$

29

with the corresponding aggregate price

$$P_M = \left(\int_{\epsilon n^W} p_i^{1-\sigma} di \right)^{\frac{1}{1-\sigma}}.$$

With the traditional constant elasticity of substitution assumptions of Dixit-Stiglitz, all varieties are equally good substitutes for each other, and the consumers express love-of-variety preferences. In equilibrium, the aggregate demand for a given variety can be written

$$c_i = \frac{p_i^{-\sigma}}{\int_{\epsilon n^W} p_i^{1-\sigma} di} \alpha Y.$$

Demand for a given variety is thus always decreasing in its own price, and increasing in the prices of competing varieties. As is also normal in this kind of model, profit maximizing generates first-order conditions which imply that the equilibrium prices are constant markups over costs:

$$p_i = \frac{\sigma}{\sigma - 1} w > w. \tag{2}$$

Rearranging these first-order conditions, yields

$$(p_i - w) c_i = \frac{p_i c_i}{\sigma},$$

where the left-hand side obviously denotes operating profits. This means that we can express operating profits as some constant fraction $\frac{1}{\sigma}$ of total revenue.[2] All firms are identical and produce with the same costs, and prices are therefore the same for all varieties. In autarky only locally produced varieties are available, so the integral $\int_{\epsilon n^w} p_i^{1-\sigma} di = n p^{1-\sigma}$. Using the result that $w = 1$, operating profits can thus be written:

$$\pi^A = \frac{\alpha Y}{\sigma n}. \tag{3}$$

This shows, quite intuitively, that consumers' preferences for the manufactured good α, and the total budget in the country Y in sum defines the size of the national market, and a larger market means larger profits for firms. Profits are

---

[2]I assume that the second-order conditions for utility maximization are fulfilled for all values of n^W. This implies that a limitation for my parameters is that $\alpha \leq \frac{\sigma - 1}{\sigma}$.

naturally falling in the number of firms n, which simply expresses how many firms will have to share the market. The elasticity of substitution σ captures the competition aspect of this model. A higher value of σ means that consumers are better able to substitute one variety for another, and this limits the firms' ability to set the price above the marginal costs. More similar varieties will thus decrease firms' market power in their variety's segment, and will reduce profits.

The national income is determined by

$$Y = \int W \, dG\left(W\right) + \left(1 - n\right) + n\left(\pi - I\right).$$ (4)

This is simply the sum of all the individuals' initial wealth, plus the share of the population who are workers, and earning normal wages, plus the share of the population who are entrepreneurs, earning $\pi - I$. Equations (3) and (4) define a unique equilibrium, and simultaneously determine firms' operating profits and expendable income as functions of the number of firms. This solution can also be used to calculate the real wages in this economy:

$$\omega^A = \left(\frac{\sigma - 1}{\sigma}\right)^\alpha n^{\frac{\alpha}{\sigma - 1}}.$$

Social welfare can be expressed as

$$V^A = \frac{\sigma}{\sigma - \alpha} \left[\int W \, dG\left(W\right) + 1 - n\left(1 + I\right)\right] \left(\frac{\sigma - 1}{\sigma}\right)^\alpha n^{\frac{\alpha}{\sigma - 1}}.$$

This is increasing in the number of firms, implying that the country as a whole will benefit from improvements in the financial system that increase the degree of industrialization, which seems realistic given most indices of quality of life and degree of industrialization.

Behind these results lies the fact that firm owners may lose from competition from new firms, whereas the constrained agents in the economy always gain from an increased number of firms, thus creating an insider-outsider problem, where once on the inside, firm owners do not want anyone else to be able to start up a business, even though it would be for the benefit of the society as a whole. This could explain situations in countries where a privileged few may work against reforms that would

improve local credit markets in order to maintain their favorable position. Such issues are, however, beside the scope of this paper and will not be addressed.

2.3 Static model in a globalized world

With no fixed costs in exporting and costless differentiation, all firms sell in both markets. CES demand functions and standard Dixit-Stiglitz monopolistic competition imply that mill-pricing is optimal. This means that a firm producing in the north and selling its product in the north at a price p, will sell its product in the south at a price $p^* = \tau p$, where $\tau \geq 1$ denotes the tradition iceberg trading costs. Total operating profits for a firm producing in the north can thus be written

$$\pi = \left(\frac{Y}{\int_{\epsilon n W} p_i^{1-\sigma} di} + \frac{\tau^{1-\sigma} Y^*}{\int_{\epsilon n W} p_i^{*1-\sigma} di} \right) \frac{\alpha p_i^{1-\sigma}}{\sigma}.$$

Here $\int_{\epsilon n W} p_i^{1-\sigma} di$ can be seen as the degree of competition in the northern market. Since prices are constant mark-ups over marginal costs and marginal costs are equal for all firms I drop the subscripts, and competition in a market can be written

$$\int_{\epsilon n W} p_i^{1-\sigma} di = np^{1-\sigma} + n^* (\tau p)^{1-\sigma} = \left(n + \tau^{1-\sigma} n^* \right) p^{1-\sigma}.$$

Following common practice in the literature, I let $\phi = \tau^{1-\sigma}$ denote the freeness of trade. With $\sigma > 1 \implies 0 \leq \phi \leq 1$, meaning that $\phi = 1$ denotes completely costless trade, and $\phi = 0$ implies infinite trade costs.

Operating profits for a firm producing in the north can thus be written

$$\pi = \left(\frac{Y}{n + \phi n^*} + \frac{\phi Y^*}{\phi n + n^*} \right) \frac{\alpha}{\sigma}. \tag{5}$$

Conversely, a firm producing in the south earns operating profits of

$$\pi^* = \left(\frac{\phi Y}{n + \phi n^*} + \frac{Y^*}{\phi n + n^*} \right) \frac{\alpha}{\sigma}. \tag{6}$$

Total expenditure in the north will be

$$Y = \left[\int W dG\left(W\right) + 1 \right] - n\left(1 + I\right) + n\pi. \tag{7}$$

In the south total expenditure will be

$$Y^* = \left[\int W dG\left(W\right) + 1 \right] - n^*\left(1 + I\right) + n^*\pi^*. \tag{8}$$

Equations (5)-(8) form a system of equations that can be solved to express the equilibrium in the model in terms of n and n^*; the number of northern- and southern-owned firms, respectively.

When capital is mobile, unconstrained individuals may choose to produce in either country. With costless reallocation firms will naturally flow towards the market where profits are higher, until profits are equal in both markets or all firms are concentrated in one market. The difference between profits for firms operating in the north and profits in the south can be written

$$\pi - \pi^* = \left(\frac{Y}{n + \phi n^*} - \frac{Y^*}{\phi n + n^*} \right) \frac{\alpha\left(1 - \phi\right)}{\sigma}.$$

The sign of this expression is determined by the relative sizes of the markets over market competition in the two countries. A larger market is more attractive if the number of firms is equal in both countries, but if there is a sufficiently high number of firms in the large market, profits might be higher for firms in the smaller market. Note, however, that the owners are not mobile, and profits will be repatriated and used in consumption in the home country.

An individual firm's incentive to move takes into consideration that each firm is assumed to be infinitesimal, and will not individually affect the price index in any of the countries. However, in aggregate, the moving firms will affect market conditions, and this feedback must be incorporated into the equilibrium condition. Let $-n \leq m \leq n^*$ be the measure of firms relocating from the south to the north. The firms' profit expressions can then be rewritten

$$\pi = \left(\frac{Y}{n + \phi n^* + \left(1 - \phi\right) m} + \frac{\phi Y^*}{\phi n + n^* - \left(1 - \phi\right) m} \right) \frac{\alpha}{\sigma}, \tag{5'}$$

33

and

$$\pi^* = \left(\frac{\phi Y}{n + \phi n^* + (1 - \phi) m} + \frac{Y^*}{\phi n + n^* - (1 - \phi) m} \right) \frac{\alpha}{\sigma}. \tag{6'}$$

In equilibrium, as long as there is not full specialization, i.e. both countries still have manufacturing firms, profits must be equal for firms located in both markets, meaning that

$$\pi - \pi^* = \left[\frac{Y}{n + \phi n^* + (1 - \phi) m} - \frac{Y^*}{\phi n + n^* - (1 - \phi) m} \right] \frac{\alpha (1 - \phi)}{\sigma} = 0.$$

This can be rearranged to express the measure of firms moving from the south to the north:

$$m = \frac{(\phi n + n^*) Y - (n + \phi n^*) Y^*}{(1 - \phi) (Y + Y^*)}. \tag{9}$$

Equations (5'), (6'), and (7)-(9) complete a set of five equations that determines capital and trade flows in equilibrium as functions of n and n^*. First, simply from the condition that in equilibrium $\pi = \pi^*$, it can be shown that the north will always have more than a proportional share of manufacturing firms, due to the better functioning credit markets; $\frac{n}{n^*} > \frac{Y}{Y^*}$. To see this, note that with firms' profits being equal for northern- and southern-owned firms in equilibrium, dividing equation (7) over (8) yields

$$\frac{Y}{Y^*} = \frac{\left[\int W dG (W) + 1 \right] + n (\pi - 1 - I)}{\left[\int W dG (W) + 1 \right] + n^* (\pi - 1 - I)}.$$

If this is to be larger than the relative number of firms, it must be that $n^* > n$, which will never be the case as long as initial wealth is identical in the two countries, and the north has better functioning credit markets.

In equilibrium all firms' profits will be equal to

$$\pi^I = \frac{2 \left[\int W dG (W) + 1 \right] - (n + n^*) (1 + I)}{n + n^*} \frac{\alpha}{\sigma - \alpha}.$$

Note that this expression is independent of trade costs, ϕ. Since country income was only potentially affected by trade costs through firms' profits, income is also independent of ϕ. These facts greatly simplify the discussion on how trade inte-

gration affects capital flows, as m is only affected by ϕ directly. The change in the measure of northbound-moving firms from a marginal change in trade costs can then be expressed

$$\frac{\partial m}{\partial \phi} = \frac{(Y - Y^*)(n + n^*)}{(Y + Y^*)(1 - \phi)^2},$$

which is clearly positive. It is also easily shown that this effect is convex in ϕ. In other words, as the countries get more integrated, more firms will be located in the north, and this effect is accelerating as trade costs fall.

There are, however, other interesting features of the function for moving firms. If trade barriers approach infinity, the number of firms moving north can be expressed as $\frac{n^*Y - nY^*}{Y + Y^*}$. Using the fact that $\frac{n}{n^*} > \frac{Y}{Y^*}$ it is easy to see that this will always be negative, meaning that there will be a flow of firms moving from the north to the south. This result comes from the fact that as trade in manufactured goods goes towards zero, firms in the south are completely protected from competition from northern firms. Since the north initially has a more than proportional share of manufacturing firms, profits will be larger for southern firms in this protected state of the world, thus attracting northern firms to move production south. In this case the model replicates the predictions from the neoclassical models, where capital will flow from capital-rich to capital-poor countries. It also illustrates an example of capital movements as a substitute for trade, as famously argued by Mundell (1957). However, in my model trade flows and capital flows may be both complements and substitutes, and capital may flow both to and from the capital-poor south, all depending on the level of trade costs. To see this, consider a reduction in trade costs. This has two effects on firms' profits: on one hand it lowers the final price that firms charge in their foreign market, and thus makes them more competitive in this market. This increases the exports' contribution to total profits. On the other hand, foreign firms become more competitive in the local market, thus stealing from the firms' home market. Since there are always more firms in the north than in the south, it can be shown that for firms in the south the second effect dominates the first. Solving the system of equations it is easy to show that $\frac{\partial m}{\partial \phi} \geq 0$, which implies that a reduction in trade costs always makes locating in the northern market relatively more attractive. For subsequent reductions in trade costs this net effect will be even stronger since some firms have

now moved from the south to the north, and the competition effect for firms in the south will now be even more dominant than for the initial reduction in trade costs. This explains why trade integration does not only increase m, but does so exponentially; $\frac{\partial^2 m}{\partial \phi^2} \geq 0$. Further, it is also the case that at the limit, when $\phi \to 1$, the right hand side of (9) goes to infinity, meaning that there exists a level of trade costs such that the movement of firms switches, and that for freer trade than this level, southern firms will start locating in the north. With the measure of moving firms being bounded by $-n \leq m \leq n^*$ this also implies that for sufficiently free trade all manufacturing firms will want to locate in the north, and the south will be completely deindustrialized.

These results imply that there must be some value of trade integration where factor flows reverse. This point is defined as the value of $\phi = \phi^{CR}$ that yields $m = 0$, which can be shown to be

$$\phi^{CR} = \frac{nY^* - n^*Y}{nY - n^*Y^*}.$$

Following the arguments above it must be the case that $0 < \phi^{CR} < 1$.

There will be some value of ϕ that will lead to full agglomeration of manufacturing firms in the north. This occurs when the expression for m reaches its upper bound; $n^* = m$. Solving this gives the simple solution that full agglomeration occurs when

$$\phi^{CP} = \frac{Y^*}{Y}.$$

This is exactly the same result as in the standard 'Footloose Capital' model, i.e. $\phi^{CP} = \frac{Y^*}{Y} = \frac{1-s}{s}$, where s denotes the north's share of the joint (world) economy (see Baldwin et al., 2003). This means that $0 < \phi^{CR} < \phi^{CP} < 1$. In other words the above discussion shows that when trade is sufficiently costly, but real capital flows freely, firms will move from the north to the south, but that trade integration will always lead more firms to locate in the north, and for sufficiently free trade there will be FDI flows from southern owners in the north, while for further reductions in trade costs there will be a level of integration such that all existing firms will be located in the north. The measure of firms moving from the

south to the north, m, can be depicted as a function of ϕ as shown in Figure 1.
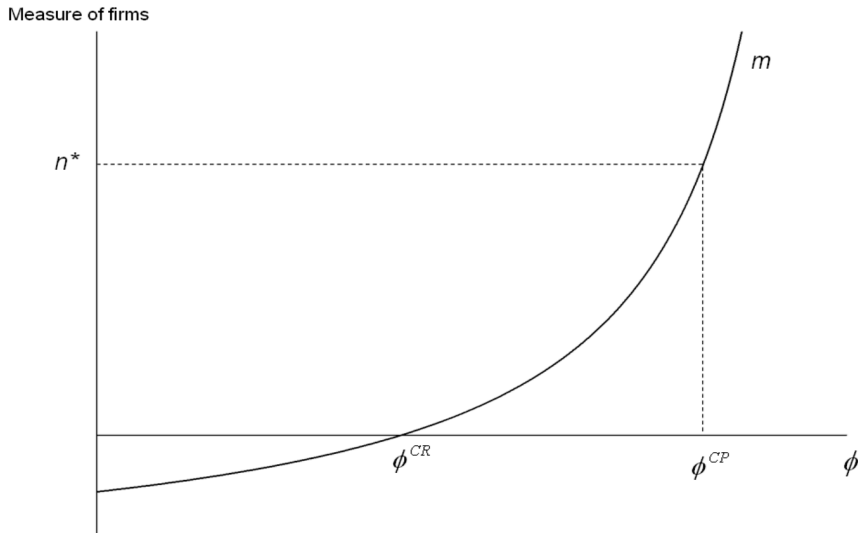
Measure of firms

Figure 1: Measure of firms moving north

The model permits analytical predictions for the welfare effects of trade liberalization. The real wage in the north in the international equilibrium is

$$\omega^I = \left(\frac{\sigma - 1}{\sigma}\right)^\alpha [n + \phi n^* + (1 - \phi) m]^{\frac{\alpha}{\sigma - 1}}$$

which is higher than in autarky if

$$\phi n^* + (1 - \phi) m > 0.$$

This will be the case when $\phi (n + n^*) Y + n^* Y - n Y^* > 0$, which is clearly not the case for sufficiently low levels of trade integration, since $\frac{n}{n^*} > \frac{Y}{Y^*}$. It is increasing in ϕ, however, and will hold for

$$\phi > \frac{nY^* - n^*Y}{(n + n^*)\,Y}.$$

It can be shown that $\frac{nY^*-n^*Y}{(n+n^*)Y} < \phi^{CR}$, which means that real wages in the north will be higher with globalization than in autarky even under some level of trade costs that will actually lead to a deindustrialization of the north. Further, the real wage in the north is always increasing in ϕ, so further trade integration will always increase the welfare of workers in the north. These results are similar to the ones in Brander and Krugman (1983) previously mentioned.

It is quite intuitive that the real wage in the north is lower in the international equilibrium than in autarky when trade is costly, and that it is increasing in the degree of freedom of trade. Since nominal wages and the price of the traditional good are constant, all changes follow directly from changes in the aggregate prices for the manufactured varieties. In the international equilibrium when trade costs are prohibitively high, I have shown that firms will move to the south. Since with trade costs at this level there will be no trade, the only difference between the international and the autarky equilibria is that there will be fewer firms in the north in the international equilibrium, the aggregate price of manufactured varieties will be higher, and the real wage will be lower. A trade liberalization, however, has two effects that decrease the aggregate price, and hence increase the real wage. First, as shown above, freer trade means that more firms will locate in the north, meaning that their products will be sold without trade costs, and secondly, the varieties that are still produced in the south and thus include transport costs in the final price will have lower trade costs. There will thus be a level of trade costs, where for trade costs lower than this the real wage in the north will always be higher in the international equilibrium than in autarky.

The real wage in the south in the static international equilibrium will be

$$\omega^{*I} = \left(\frac{\sigma - 1}{\sigma}\right)^{\alpha} \left[\phi n + n^* - (1 - \phi)\,m\right]^{\frac{\alpha}{\sigma - 1}},$$

which is higher than in autarky if

$$\phi n - (1 - \phi)\,m > 0$$

This is always the case, and contrary to the case for northern wages, wages in the south are always higher in the static international equilibrium than in autarky. The left hand side of this inequality is increasing in ϕ. Since $m < 0$ for $\phi = 0$ the above condition must hold for all non-negative values of ϕ, implying that real wages are always higher under international trade than in autarky in the south. The intuition here is not as clear as for the northern country. It is straightforward that the international equilibrium with prohibitively high trade costs means increased real wages in the south, as firms will move from the north to the south in this case. The two effects from a lowering of trade costs now have opposite effects, however: imported varieties from the north become cheaper, but more firms will locate in the north, and will thus include trade costs in their final price in the south. The first effect will, however, always dominate, and the real wage is higher in the international equilibrium than in autarky, and this difference increases as trade gets freer.

This means that the model predicts two very different scenarios: if trade is relatively costly, $\phi \epsilon \left(0, \frac{nY^*-n^*Y}{nY-n^*Y^*}\right)$, a move from autarky to the international equilibrium means that real wages will fall in the north, while they will increase in the south. When trade is freer than this, however, real wages will increase in both countries under internationalization.

3 Static model when trade costs are asymmetric

The above results; that a trade liberalization may lead to deindustrialization in the south is somewhat different from what Antràs and Caballero (2009) argue in their Heckscher-Ohlin based model with finance market imperfections. One reason for this is that in a Heckscher-Ohlin model of trade there is no intra-industry trade, so any reduction in trade costs in one sector will work as increased market access for one of the countries. In this subsection I will show that my model will generate similar predictions to those of Antràs and Caballero when I consider a unilateral trade liberalization instead of a symmetrical reduction in trade costs.

Let us now assume that market access to the foreign market is not necessarily identical for firms producing in the north and in the south. This can be thought of as some import tax (that is wasted), or import costs associated with custom

clearance, paperwork, etc., t, such that $\phi = (\tau + t)^{1-\sigma}$, and where it may be the case that $t \neq t^*$. Let ϕ denote the degree of access to the southern market for the northern firms, while ϕ^* indicates southern firms' access to the northern market. The operating profits for northern firms will then be

$$\pi = \left(\frac{Y}{n + \phi^* n^*} + \frac{\phi Y^*}{\phi n + n^*} \right) \frac{\alpha}{\sigma}.$$

Conversely, a firm producing in the south will earn operating profits of

$$\pi^* = \left(\frac{\phi^* Y}{n + \phi^* n^*} + \frac{Y^*}{\phi n + n^*} \right) \frac{\alpha}{\sigma}.$$

The measure of moving firms is now defined as the m that ensures that

$$\pi - \pi^* = \left(\frac{(1 - \phi^*) Y}{n + m + \phi^* (n^* - m)} - \frac{(1 - \phi) Y^*}{\phi (n + m) + n^* - m} \right) \frac{\alpha}{\sigma} = 0.$$

Solving this with respect to m yields the following expression:

$$m = \frac{Y (1 - \phi^*) (\phi n + n^*) - Y^* (1 - \phi) (n + \phi^* n^*)}{(1 - \phi^*) (1 - \phi) (Y + Y^*)}.$$

The national incomes are determined as above, which implies that a drop in trade costs for manufactured varieties produced in the south and sold in the north will affect the measure of firms moving north in the following way:

$$\frac{\partial m}{\partial \phi^*} = - \frac{(n + n^*) Y^*}{(1 - \phi^*)^2 (Y + Y^*)} < 0,$$

which tells a similar story to that of Antràs and Caballero (2009) where firms move to the south and export their goods back to the north. The intuition behind this is straightforward; the unilateral trade liberalization increases competition in the north, thus reducing the profits of firms producing in the north. At the same time it makes firms producing in the south more competitive in the northern market, thus increasing profits for firms producing in the south. Competition in the southern market is unaffected. This yields an increase in profits for firms in

the south, and a drop in profits for firms producing in the north, which will be compensated by firms moving from the north to the south, until equilibrium is restored.

This illustrates how different the effects of these different trade liberalizations are for the involved parts, and thus shows that one should be careful when discussing the effects of globalization on both capital movements and welfare levels. In the rest of the paper I will however, stick to the symmetrical reduction in trade costs.

4 Model with inter-generational savings

Up until now, I have looked at static version of the model. There are, however, some interesting effects when I introduce generations and savings through bequests into the model. In this section I will incorporate this, first into the autarky version of the model, and later into the model with firm mobility and international trade.

The representative consumer's utility function is now:

$$U = \frac{1}{\alpha^\alpha \beta^\beta \gamma^\gamma} C_M^\alpha C_A^\beta B^\gamma, \ 0 < \alpha, \beta, \gamma < 1, \ \alpha + \beta + \gamma = 1, \qquad (1')$$

where C_M and C_A are as before, and B represents bequests left to the next generation. This is a reduced-form altruism, where bequests leave the giver with a "warm, fuzzy feeling". The great advantage of this, is that individuals will leave a constant share γ of their expendable income as bequests for their offspring. A more realistic way of modelling altruism would be to let the offspring's utility enter directly into the utility-function of the giver. Such a utility function would lead to nonlinearities in the share of total income individuals will leave to their offspring for individuals who would be marginally too poor to leave their offspring unconstrained in my formulation. This complicates calculations severely, but does not change the conclusions of the model qualitatively, and I therefore prefer the simplified version for modeling altruism. For a more thorough discussion of the issues related to this simplification, see Chesnokova (2007). The wealth dynamics

41

for "family" i are as follows,

$$W_{i,t+1} = \begin{cases} \gamma\left(W_{i,t}+1\right), & \text{if } W_{i,t} < \bar{W} \\ \gamma\left(W_{i,t}+\pi_t-I\right), & \text{if } W_{i,t} \geq \bar{W} \end{cases}.$$

I call the W's initial wealth, as this is the wealth an individual has at the start of his life, which is the wealth that determines whether he is credit-constrained or not. This should not be confused with the individual's budget constraint, which will be the sum of this initial wealth, and earnings. This start-of-period initial wealth will through the above dynamics converge towards $W_u = \frac{\gamma}{1-\gamma}\left(\pi-I\right)$ for unconstrained agents, and $W_c = \frac{\gamma}{1-\gamma}$ for constrained agents. This can be depicted in Figure 2. The slope of the inter-generational wealth dynamics are equal for constrained and unconstrained agents, but the graph for the unconstrained agents will always lie above the one for the credit-constrained agents. Since $\gamma < 1$ these slopes are flatter than the 45° line where $W_{t+1} = W_t$, which ensures that they will cross this line once from above.
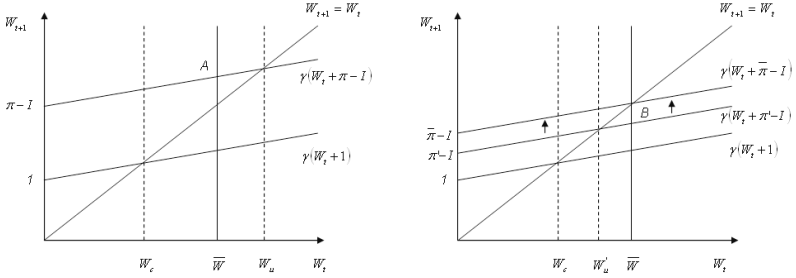
Figure 2: Inter-generational wealth dynamics

Credit-constrained agents with initial wealth below W_c will leave their offspring with more initial wealth than they had themselves. Conversely, credit-constrained agents with initial wealth above W_c will leave less in bequests than they started out with. This implies two important things: the share of the population that is credit-constrained does not diminish, and initial wealth for credit-constrained

agents will converge towards $W_c = \frac{\gamma}{1-\gamma}$.

The story for the unconstrained agents is somewhat different. While it is the case that agents with $W_i > W_u$ will leave less in bequests than they started with, and agents with $\bar{W} < W_i < W_u$ will leave more than they started with, the unconstrained share of the population may diminish if profits are too low. Figure 2 depicts the wealth dynamics for two situations: when profits are $\pi > \bar{\pi}$, and when they are $\pi' < \bar{\pi}$. Note that for the credit-constrained agents, wealth dynamics are unchanged in the two cases. In the diagram to the left profits are π and all unconstrained agents are able to leave their offspring unconstrained as well. To see this, note that the poorest unconstrained agent with wealth $W_i = \bar{W}$ will still be able to leave the next generation with sufficient funds to be unconstrained (point A). Since his initial wealth is below W_u he will also be able to leave his offspring with more initial wealth than he started out with himself. In this case the number of unconstrained agents, and thus also the number of firms, stays the same, profits are unchanged, and initial wealth among unconstrained agents converges towards $W_u = \frac{\gamma}{1-\gamma}(\pi - I)$. If, initially, profits are π' the story is different. In this case all unconstrained agents have initial wealth above W_u' and will thus leave their offspring with less initial wealth than they had. This means that the poorest unconstrained agent, with $W_i = \bar{W}$, (point B) will not be able to leave the next generation unconstrained. This means that in $t+1$ there will be fewer firms, which we from (3) clearly see will increase profits. This again shifts the wealth dynamic function upwards, meaning that the long-run equilibrium will *not* converge to W_u'. This process will rather repeat itself until profits have been pushed up to $\bar{\pi}$. At this point the poorest unconstrained agent will earn exactly enough to leave the next generation unconstrained, and the initial wealth of the unconstrained agents will converge towards $\frac{\gamma}{1-\gamma}(\bar{\pi} - I) = \bar{W}$. Since we know from before that $\bar{W} = \frac{I}{\delta}$ this profit level is determined by

$$\bar{\pi} = \frac{1 + (\delta - 1)\gamma}{\gamma \delta} I.$$

It is easy to see that the minimum operating profits that can sustain the population of firms is decreasing in the sophistication of the contractual environment, δ. This is intuitive, as when credit constraints are less binding, less wealthy potential

43

entrepreneurs can get access to financing to start up their businesses, and thus need to leave less bequests to their offspring for them to be financially unconstrained as well. The dynamics explained in Figure 2 determine the size of the unconstrained share of the population, and also the equilibrium wealth levels of both constrained and unconstrained agents. Over time all agents will converge to these wealth levels, and total expendable income in the country will converge towards

$$Y = \frac{1 + n\left[\pi - (1 + I)\right]}{1 - \gamma}.$$

If the initial number of firms can be sustained in the long-term equilibrium, inserting this income level into (3) will determine the operating profits as a function of the number of firms, where this number is again determined by the share of the population that is initially unconstrained, just as in the static version of the model. Equilibrium profits can thus generally be expressed

$$\pi = \max\left\{\frac{1 + (\delta - 1)\gamma}{\gamma\delta}I, \frac{\alpha}{\sigma(1 - \gamma) - \alpha}\left(\frac{1 - n}{n} - I\right)\right\}.$$

If profits initially are too low to sustain the number of firms, the unconstrained share of the population will shrink according to the mechanisms illustrated in Figure 2 until:

$$\frac{1 + (\delta - 1)\gamma}{\gamma\delta}I = \frac{\alpha}{\sigma(1 - \gamma) - \alpha}\left(\frac{1 - \bar{n}}{\bar{n}} - I\right).$$

The number of firms in the dynamic equilibrium is thus

$$n = \min\left\{1 - G\left(\frac{I}{\delta}\right), \frac{\alpha\gamma\delta}{\alpha\gamma\delta(I + 1) + [\sigma(1 - \gamma) - \alpha][1 + (\delta - 1)\gamma]I}\right\}.$$

Both of these expressions are increasing in the quality of the contractual climate, δ, meaning that the number of entrepreneurs will always be larger in the northern country when the countries are identical in all other aspects than in contract enforceability. The number of firms in the dynamic model is also equal to or lower than than in the static model. Since there is no leapfrogging in the income ranking, it is possible to define a wealth level $\tilde{W} \geq \bar{W}$ that is the minimum wealth level in

period $t = 0$ that will leave the agent's successors unconstrained in the long-run equilibrium. This wealth level is implicitly defined by

$$\int_{\bar{W}}^{\tilde{W}} dG\left(W\right) = 1 - G\left(\frac{I}{\delta}\right) - n.$$

If initial profits are sufficiently high to sustain the number of firms, the dynamic model does not generate any interesting changes from the static version, so in the rest of the paper I will consider the situations where this is not the case, and the equilibrium number of firms is determined endogenously by credit constraints and profits.

4.1   Globalization in the dynamic model

Since the equilibrium number of firms in the dynamic model is determined endogenously and, as I will show, is now affected by the international competition, I will in the following denote the number of firms in autarky and in the international equilibrium by n^A and n^I, respectively. If equilibrium profits are too low to sustain the initial number of firms, profits in autarky will converge towards

$$\bar{\pi} = \frac{1 + (\delta - 1)\gamma}{\gamma\delta} I \tag{10}$$

$$\bar{\pi}^* = \frac{1 + (\delta^* - 1)\gamma}{\gamma\delta^*} I. \tag{11}$$

This expression is clearly decreasing in δ, and since $\delta > \delta^*$, this implies that $\bar{\pi} < \bar{\pi}^*$, meaning that profits must be higher for unconstrained agents in the south to be able to leave their offspring unconstrained as well. This can be seen as an analogy to a situation where projects in the south must present a higher expected profitability in order to attract financing, which is reflected in the insurance costs of projects in developing countries compared to similar projects in the developed world.[3]

In the international equilibrium with firm mobility and international trade, profits will be the same for both northern and southern firms. There are three

---

[3]See for example price differences for investment guarantees at MIGA (http://www.miga.org/) and similar organisations.

possible scenarios when comparing short-term profits with the sustainable profit levels:

1. $\bar{\pi}^* > \bar{\pi} > \pi^I$

2. $\bar{\pi}^* > \pi^I \geq \bar{\pi}$

3. $\pi^I \geq \bar{\pi}^* > \bar{\pi}$

Of these, only case 3 is a stable equilibrium when taking into account the intergenerational wealth dynamics. In case 1 profits are too low in both countries to maintain the current number of firms. This means that over time fewer entrepreneurs will leave their offspring with sufficient bequests to start a business, total number of firms will fall, and profits of the remaining firms will increase, until $\pi^I = \bar{\pi}$. At this point the situation will be as described in case 2. At this point, the poorest entrepreneur in the north will be able to leave his offspring enough bequests for him to be unconstrained as well, and from this point in time the number of firms in the north will be stable. However, profits are still not high enough to sustain the unconstrained share of the population in the south. The number of southern firms will thus keep falling until $\pi^I = \bar{\pi}^*$. In this stable equilibrium operating profits will be $\pi^I = \frac{1+(\delta^*-1)\gamma}{\gamma\delta^*}I$. This long-run equilibrium condition from the credit constraints can be used to determine the equilibrium number of firms as in the autarky example. This level of operating profits means that expendable income in each country will converge towards

$$Y = \frac{[(1-\gamma)I - \gamma\delta^*]n^I + \gamma\delta^*}{\gamma\delta^*(1-\gamma)} \tag{12}$$

$$Y^* = \frac{[(1-\gamma)I - \gamma\delta^*]n^{*I} + \gamma\delta^*}{\gamma\delta^*(1-\gamma)}. \tag{13}$$

The measure of firms moving north can now be determined by

$$m = \frac{\phi[(1-\gamma)I - \gamma\delta^*](n^I + n^{*I}) - (1-\phi)\gamma\delta^*}{(1-\phi)[(1-\gamma)I - \gamma\delta^*](n^I + n^{*I}) + 2(1-\phi)\gamma\delta^*}(n^I - n^{*I}). \tag{14}$$

Firms will still move until $\pi = \pi^*$, so inserting (12), (13), and (14) into (5')

equilibrium profits can be expressed as a function of the total number of firms

$$\pi = \frac{\left[(1-\gamma)\,I - \gamma\delta^*\right]\left(n^I + n^{*I}\right)\alpha + 2\alpha\gamma\delta^*}{(1-\gamma)\,\gamma\delta^*\left(n^I + n^{*I}\right)\sigma}.\tag{15}$$

This profit expression is thus determined by total demand and the number of competing firms. Since I am only focusing on the situation where initial profits are too low to sustain the initial number of firms, the long-run equilibrium number of firms can be found by setting (11) equal to (15), and will be

$$n^I + n^{*I} = \frac{2\alpha\gamma\delta^*}{\alpha\gamma\delta^*\,(1+I) + \left[\sigma\,(1-\gamma) - \alpha\right]\left[1 + (\delta^* - 1)\,\gamma\right]I}.$$

It is easy to see that as long as both goods are produced in both countries, the total number of firms in the world economy only depends on the quality of the financial system in the south, δ^*. This happens because the north will always reach a state where the unconstrained share of the population is stable before the south does. After this point, the number of southern-owned firms will keep falling. A less developed financial system in the south means that the total number of firms in the long run will be lower. However, it also means that the number of northern-owned firms will be higher. This follows naturally from the fact that as the number of southern firms fall, profits for the remaining firms, both northern and southern, rise, thus making a lower share of the northern population capital constrained in the long run. Stricter credit constraints in the south will mean that more southern-owned firms go out of business each period. This again increases profits faster, so they will reach the critical level to sustain the number of northern-owned firms, $\bar{\pi}$, faster, thus leaving the north with a larger share of the manufacturing industry. The financial system in the north does not affect the total number of firms in equilibrium, but a high δ implies that $\bar{\pi}$ will be lower, and thus will be reached faster, implying that a better functioning financial system in the north leads to a crowding out effect in the south, and will increase the north's share of world manufacturing industry. In the long-run equilibrium, it is thus the financial system in the south that determines the total number of firms in the world, but it is the relative strengths of the financial systems that determine the relative number of entrepreneurs in the two countries.

In autarky, the number of southern firms was determined by

$$n^{*A} = \frac{\alpha\gamma\delta^*}{\alpha\gamma\delta^* \left(I + 1\right) + \left[\sigma\left(1 - \gamma\right) - \alpha\right]\left[1 + \left(\delta^* - 1\right)\gamma\right]I},$$

which is exactly one half of the total number of firms in equilibrium with international trade, thus $2n^{*A} = n^I + n^{*I}$. Further, with profits being equal for all firms, there can be no leapfrogging in wealth among unconstrained agents in either country, and there will always be more northern-owned firms than southern-owned firms. This again implies that the number of southern firms is always lower in the international equilibrium than in autarky, $n^{*A} \geq n^{*I}$. The reason for this is that with stricter credit constraints and fewer firms in the south, the firms located here will be more protected than their northern counterparts in autarky, which leads to higher profits in equilibrium. With international competition these profits are pushed down, and some firms must leave the market in order to raise profits to sustainable levels again. Since equilibrium profits for northern firms are higher in the international equilibrium than in autarky, the number of northern-owned firms may actually be higher in the globalized world than in autarky if the drop in the number of southern firms happens sufficiently fast; however, this depends on the distribution of initial wealth, and the rates at which the world converges to the long-run equilibrium.

Comparing the above results to the results from the static model we can immediately point out some important differences. With mobile firms, all firms earn the same profits. Since profits for southern firms are equal in autarky and in the international equilibrium, and further are higher than regular wages, this means that fewer firms means lower total income in the country. An effect from international trade and capital mobility for the southern country as a whole that is ignored in the static model is thus that when credit constraints are binding, there will be a crowding out effect with international trade. Competition from northern firms will drive some of the southern firms out of the market, and the deindustrialization will reduce the total country income. However, as I showed above, welfare levels may indeed increase from internationalization in spite of falling national income, as international trade means increased variety of manufactured goods and a drop in the price index. However, compared to the static model the dynamic equilibrium

predicts a worse outcome for the south for this effect for all levels of trade costs, as income and the total number of firms will be lower.[4] Full agglomeration occurs when trade integration levels are sufficiently high, so that $m = n^{*I}$, which can be shown to be

$$\phi^{CP} = \frac{\gamma\delta^* + [(1-\gamma)I - \gamma\delta^*]n^{*I}}{\gamma\delta^* + [(1-\gamma)I - \gamma\delta^*]n^I}.$$

Similarly to in the static model, this is equal to $\frac{Y^*}{Y}$, and since $n^I > n^{*I}$ it is also the case that $0 < \phi^{CP} < 1$.

The model with intergenerational savings thus reproduces the main predictions about trade and capital flows from the static model. However, I have also shown that when the number of firms is determined by the long-term credit constraints, the competition effect from international trade reduces the number of southern firms compared to the autarky outcome. This reduction in the number of southern-owned firms will for some levels of trade costs reduce the number of firms moving north, but this reduction is never enough to compensate for the direct reduction in the number of firms producing in the south, and it can be shown that $\frac{\partial(n^*-m)}{\partial n^*} > 0 \forall \phi$. This means that for any level of trade costs the number of firms producing in the south will be lower, compared to the static model. This means that both the real wage and social utility in the south are lower in the dynamic model than in the static model. Looking specifically at the real wage in the south, it can be shown that in the dynamic model the difference between the autarky and the international equilibria is

$$\omega^{*I} - \omega^{*A} = \left(\frac{\sigma-1}{\sigma}\right)^\alpha \left\{\left[\phi n^I + n^{*I} - (1-\phi)m\right]^{\frac{\alpha}{\sigma-1}} - \left(n^{*A}\right)^{\frac{\alpha}{\sigma-1}}\right\}.$$

Since wages are normalized to 1, the only effect on real wages in the model is through changes in the price index, and real wages in the international equilibrium are thus higher than in autarky if the price index is lower, which is equivalent to the condition:

$$\phi n^I + n^{*I} - (1-\phi)m > n^{*A}. \tag{16}$$

---

[4]This follows from my choice to only focus on the situation where credit constraints are so severe that the initial number of firms is not sustainable. However, as there will never be more firms in the dynamic equilibrium compared to the static equilibrium, welfare levels will never be higher in the dynamic model.

In stark contrast to the static equilibrium, this is no longer always the case. Inserting for m it is quite straightforward to show that the left-hand side of the above expression is monotonically increasing in ϕ. Further, inserting $\phi = 0$ and rearranging yields that $\omega^{*I} > \omega^{*A}$ if and only if:

$$[(1-\gamma)I - \gamma\delta^*]\left(n^I + n^{*I}\right)\left(n^{*I} - n^{*A}\right) > 0,$$

which only holds for $n^{*I} > n^{*A}$, something I have already shown will never be the case. The real wage is thus lower in the international equilibrium when trade costs are prohibitively high. The other extreme is when $\phi = 1$. Condition (16) then simplifies to

$$n^I + n^{*I} > n^{*A}.$$

Since I have already shown that $n^I + n^{*I} = 2n^{*A}$, the above condition must hold, and the real wage is higher in the international equilibrium when there is perfect trade integration between the countries. Since (16) in monotonically increasing in ϕ there must also exist a level of trade costs such that the real wage in the south is higher in the international equilibrium for all values of ϕ above this. For the unconstrained agents in the south, the effects are even more severe. All families that started out as unconstrained will experience a drop in income. Some, due to the fact that the long-run equilibrium profits in the international equilibrium will be lower than in autarky, and some because they will become credit-constrained, and start earning normal wages. For sufficiently free trade this drop in income may be compensated for by an increased variety of goods, and a lower price index, but the required trade integration for this is even higher than the one that makes $\omega^{*I} > \omega^{*A}$. This implies that for high levels of trade costs, overall welfare in the south will be lower in the international equilibrium than in autarky.

Although both the static and the dynamic versions of the model predict that real wages will be maximized under completely free trade, there are some considerable and important differences. The most important of these is that in the dynamic version of the model, the international competition will push down profits for southern-owned firms, thus strengthening credit constraints, and reducing the unconstrained share of the population and the number of firms. In this version of the model globalization leads to a deindustrialization of the least developed

country, and if trade is sufficiently costly, to a drop both in real wages and also in overall welfare levels.

This shows that when taking into consideration the effects of finance market imperfections, internationalization may lead to deindustrialization and a drop in real wages in the less developed country. One must however be cautious in interpreting this as a defense for import substitution and protectionism. In my analysis I have looked at globalization as the mobility of both goods and firms. The potential losses from globalization in my model come mainly from deindustrialization in the south due to competition from northern firms, but also from southern firms moving closer to the larger market in the north. If the government in the southern country has limited possibilities in preventing inhabitants from investing abroad, or its firms from moving abroad, limiting trade will thus only increase welfare losses.

5   Conclusion

Foreign direct investment and trade flows are both important aspects of the process of globalization. In this paper I have shown how the two mutually affect each other, and thus in sum determine real wages and social welfare in a globalizing world. This interdependence shows the danger in discussing globalization in light of only trade flows, or only capital flow. Symmetrical reductions in trade costs will cause firms to move towards the larger market.

In addition, I have also shown how imperfections in the financial markets may generate the heterogeneity between countries that may lead to an industrialized centre and a deindustrialized periphery. In the long-run equilibrium such imperfections may also cause international competition to squeeze southern firms out of the market, and thus increase the difference in economic size between the developed north and the developing south. If trade is sufficiently costly, but capital mobile, this may lead to a drop in utility for both firm owners and workers in the south. This result stands in contrast to the simple static model where the initial number of firms is constant. In this case real wages in the south are always higher in the international equilibrium than in autarky. This highlights the difficulties of predicting welfare effects of proposed trade reforms. The different predictions

from the simple static model and the more complex model with inter-generational savings are general to any initial wealth distribution, and to all permitted parameter values. National welfare for the two countries is purposely left out of the discussion in this paper, as this discussion would require certain such assumptions. Generally, one can say that social welfare would be lower in the international equilibrium with high trade costs than in autarky in the north in both versions of the model, and in the south in the long-term version of the model. However, whether this would also be the case in the short-term model, and whether social welfare would turn positive for sufficiently free trade, would depend on the shape of $G(W)$ and the relative sizes of α and σ.

From a development point of view, the important conclusions are that international competition may crowd out southern firms, and lead to a deindustrialization that could possibly hurt both social welfare and real wage levels in the country with the least developed financial system.

References

Antràs, P., & Caballero, R. 2009. Trade and Capital Flows: A Financial Frictions Perspective. *Journal of Political Economy*, **117**(4), 701–744.

Baldwin, R., & Okubo, T. 2006. Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting. *Journal of Economic Geography*, **6**, 323–346.

Baldwin, R., Forslid, R., Martin, P., Ottaviano, G., & Robert-Nicoud, F. 2003. *Economic Geography and Public Policy*. Princeton: Princeton University Press.

Banerjee, A., & Newman, A. 1993. Occupational Choice and the Process of Development. *Journal of Political Economy*, **101**(2), 274–298.

Brander, J., & Krugman, P. 1983. A 'Reciprocal Dumping' Model of International Trade. *Journal of International Economics*, **15**, 313–321.

Chesnokova, T. 2007. Immiserizing Deindustrialization: A Dynamic Trade Model with Credit Constraints. *Journal of International Economics*, **73**, 407–420.

Ciccone, A., & Hall, R. 1996. Productivity and the Density of Economic Activity. *American Economic Review*, **86**(1), 54–70.

Das, S. 2006. Trade, Skill Acquisition and Distribution. *Journal of Development Economics*, **81**, 118–141.

Dupont, V., & Martin, P. 2006. Subsidies to Poor Regions and Inequalities: Some Unpleasant Arithmetic. *Journal of Economic Geography*, **6**, 223–240.

Duranton, G., & Overman, H. 2005. Testing for Localization Using Micro-Geographic Data. *Review of Economic Studies*, **72**(4), 1077–1106.

Janeba, E., & Schjelderup, G. 2003. *The Future of Globalization: Tax Competition and Trade Liberalization*. World Development Report. World Bank.

Krugman, P. 1991. Increasing Returns and Economic Geography. *Journal of Political Economy*, **99**(3), 483–499.

Krugman, P., & Venables, A. 1995. Globalization and the Inequality of Nations. *Quarterly Journal of Economics*, **110**(4), 857–880.

Martin, P., & Rogers, C. 1995. Industrial Location and Public Infrastructure. *Journal of International Economics*, **39**, 335–351.

Mundell, R. 1957. International Trade and Factor Mobility. *American Economic Review*, **47**(3), 321–335.

Prebisch, R. 1950. *The Economic Development of Latin America and Its Principal Problems.* New York: United Nations.

Puga, D. 2010. The Magnitude and Causes of Agglomeration Economies. *Journal of Regional Science*, **50**(1), 203–219.

Ranjan, P. 2001. Dynamic Evolution of Income Distribution and Credit-Constrained Human Capital Investment in Open Economies. *Journal of International Economics*, **55**, 329–358.

UNCTAD. 2008. *World Investment Report 2008: Transnational Corporations and the Infrastructure Challenge.* World Investment Report Series. UNCTAD.

Appendix A1: The number of firms producing in the south

Proof of $\frac{\partial\left(n^{*I}-m\right)}{\partial n^{*I}} \geq 0$

When the level of trade costs are such that there is full agglomeration, $m = n^{*I}$, and it follows naturally that $\frac{\partial\left(n^{*I}-m\right)}{\partial n^{*I}} = 0$. When both countries have a manufacturing sector, the differential will always be positive. For notational simplicity, let $D := [(1-\gamma)I - \gamma\delta^*] > 0$. The differential can then be written

$$\frac{\partial\left(n^* - m\right)}{\partial n^*} = \frac{2\gamma\delta^*\left(1-\phi\right)\left(n^I D + \gamma\delta^*\right) + 4D\gamma\delta^* n^{*I} + \left(n^I + n^{*I}\right)^2 D^2}{\left(1-\phi\right)\left[D\left(n^I + n^{*I}\right) + 2\gamma\delta^*\right]^2}.$$

With $D > 0$ and $1 > \phi$ all the elements of both the numerator and the denominator of this fraction are strictly positive, thus making the whole expression strictly positive, and for $\frac{\partial\left(n^{*I}-m\right)}{\partial n^{*I}} \geq 0$ to hold, it is thus sufficient to show that $D > 0$. I show this using the fact that $\frac{\partial D}{\partial I} > 0$, and that credit constraints are assumed to be binding, which means that $I > \frac{\gamma}{1-\gamma}\delta^*$. The minimum value I can take is thus $\frac{\gamma\delta^*}{1-\gamma} + \varepsilon$, with ε marginally larger than zero. Inserting this minimum value of I into D, we get:

$$(1-\gamma)\left(\frac{\gamma\delta^*}{1-\gamma} + \varepsilon\right) - \gamma\delta^* = (1-\gamma)\varepsilon > 0$$

QED.

Appendix A2: Potential welfare gain in the south from internationalization

I want to prove the following statement:

$$V^{*I} - V^{*A} > 0 \Leftrightarrow Y^{*I}\left[\phi n^I + n^{*I} - (1-\phi)m\right]^{\frac{\alpha}{\sigma-1}} - Y^{*A}\left(n^{*A}\right)^{\frac{\alpha}{\sigma-1}} > 0$$

where

$$Y^{*I} = \frac{[(1-\gamma)I - \gamma\delta^*]n^{*I} + \gamma\delta^*}{\gamma\delta^*(1-\gamma)}$$

$$Y^{*A} = \frac{[(1-\gamma)I - \gamma\delta^*]n^{*A} + \gamma\delta^*}{\gamma\delta^*(1-\gamma)}$$

$$m = \frac{\phi[(1-\gamma)I - \gamma\delta^*](n^I + n^{*I}) - (1-\phi)\gamma\delta^*}{(1-\phi)[(1-\gamma)I - \gamma\delta^*](n^I + n^{*I}) + 2(1-\phi)\gamma\delta^*}(n^I - n^{*I})$$

To prove that this may be the case, it is sufficient to insert a set of permitted parameter values such that $V^{*I} - V^{*A} > 0$. This can be done by inserting a set of permitted parameter values into the expression for n^{*A}. One can for example choose:

$$\gamma = 0.1$$
$$\delta^* = 2$$
$$I = 0.5$$
$$\alpha = 0.6$$
$$\sigma = 1.4$$

This implies that $n^{*A} = 0.179$. I have shown that $n^I + n^{*I} = 2n^{*A}$ which means that we have $n^I + n^{*I} = 0.358$. Exactly what the number of firms will be in each country in the international equilibrium will depend on δ and the initial income distributions in the countries. Without assuming any specific values and shapes for these, I merely define some values:

$$n_1^I = 0.2 \text{ and } n_1^{*I} = 0.158$$
$$n_2^I = 0.33 \text{ and } n_2^{*I} = 0.028$$

to illustrate one case with relatively similar financial markets in the two countries, and one where they are very different. These values imply that $\phi_1^{CP} = 0.958$ and $\phi_2^{CP} = 0.732\,74$. The two cases are shown below through the graphical expressions of the welfare gain from going from autarky to the international equilibrium $V_j^{*I} - V_j^{*A}$, $j = 1, 2$.
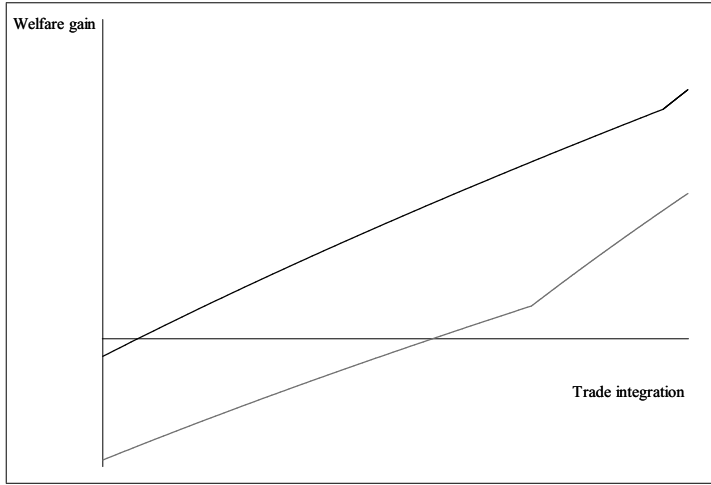
As the graphs show, in both cases social welfare in the south will be lower in the international case when trade costs are high (ϕ is low), but that it will be higher for sufficiently high trade integration. The discontinuities show the levels at which there will be complete agglomeration, and all firms will be located in the north. From this point and onwards the difference in welfare in the international equilibrium and in autarky increases more rapidly with the lowering of trade costs, as this now only makes imported goods cheaper, without causing more firms to move (as there are no firms left in the south). The higher (black) line shows the case of relatively similar countries, while the lower (red) line depicts the situation where credit constraints are much more severe in the south than in the north. These clearly show that the international equilibrium yields higher social welfare than does autarky for a much wider range of ϕ-values when the countries are more similar.

Chapter 3

Democracy and Expropriations

III

Democracy and Expropriations

Jonas Gade Christensen[*]

University of Bergen

December 31, 2010

Abstract

In this paper I develop a voting model that shows the different effects of democratic competition and political constraints on the probability of expropriations of foreign investments. I show that these two aspects of liberal democracy might have very different effects on expropriation risks. Particularly interesting is the prediction that for low to intermediate levels of political competition for executive power, increased competition will lead to higher risk of expropriation.

Testing this and other predictions on panel data for actual expropriations in 27 developing countries, I find support for the predictions from the model.

59

1 Introduction

Expropriations of foreign capital have been largely ignored by the academic literature for long spells of time. Kobrin (1984) argued that the expropriations during the 1960s and 1970s were an attempt by national authorities to control multinational enterprises (MNEs), but that they by the end of the 1970s had regulative tools that could complete this function. However, recent government expropriations in Bolivia, Venezuela, and Russia have showed that the conclusion that expropriations are a phenomenon from the past might be premature. The existence of international insurance companies that insure goods and investments against political risk implies that the mere possibility of expropriations in itself increases frictions on international trade and capital flows, and that this is still an important question today. The question is whether it is possible to find a systematic explanation of expropriations. Are they a result of rational behavior, or merely an outcome of a burst of national populism? In this paper I show both theoretically and empirically that expropriations can be the result of rational utility maximization, and that it to a certain degree can be explained by the political institutions that define the framework within which the politicians act. Specifically I propose that democratic competition for power may shorten executives' time perspective, thus making expropriations a more attractive policy option.

Democracies are very heterogeneous, and their different characteristics have different implications on economic performance. A political system with a strong, charismatic president is likely to be different in many aspects from a parliamentary system where party identities are more important than their respective candidates. In the first, the executive may focus more on his own career, whereas in the second he will have to consider the legacy he leaves behind to his party. Going into detail with all the variations of democracy is well beyond the scope of this paper, but I will focus on two aspects that I believe are central for studying the mechanisms at work between democracy and state expropriations: the competition for office, and checks on executive power. Both of these are essential for a political system to be characterized as a liberal democracy, but I will argue that they can potentially affect investment and growth in very different ways. The likelihood of losing office in elections, or due to term limits may lead executives in democracies

to have shorter time horizons than a lifetime dictator would, thus making long-term investments more insecure. Bó and Rossi (2008) study Argentinian legislators and conclude that longer terms lead to better legislative performance. Titiunik (2008) finds similar results for US Senators in Texas and Arkansas. On the other hand, the checks and balances that are often associated with democracies may ensure more predictability about the policies that will be enacted, and as such reduce uncertainties and improve the investment climate.

1.1 Previous literature

Understanding expropriations is important for individual firms that invest in countries with high political risk, but also for policy makers who work toward better growth conditions in capital-poor countries. There is a large literature studying directly how political institutions affect economic growth.[1] Generally this literature argues that the important aspects of political institutions are the ones that affect the investment climate in the country, and another branch of the literature studies directly how democracy affects FDI flows. Lucas (1990) criticizes the simple neoclassical models that predict that investment will go to countries with little capital relative to labor, and argues that one of the main reasons MNEs do not invest more in capital-poor countries is that these investments are subject to significant political risk. Jensen (2003) and Busse (2004) among others find that more democratic regimes attract more FDI, while Li and Resnick (2003) argue that democracy increases incoming FDI only through stronger property right protection, but that when these are controlled for, democracy as such actually reduces FDI inflows. Busse and Hefeker (2007), using an extensive data set of developing countries, find that government stability and law and order greatly increase foreign investments, whereas democratic accountability does this only to a lesser degree. These studies highlight the important difference between electoral competition and political constraints, but their dependent variable, incoming FDI, is

---

[1] There is a rich literature looking for the causality of institutions on economic performance (see for example Knack and Keefer, 1995; Mauro, 1995; La Porta et al., 1998; Hall and Jones, 1999; and Rodrik, 1999). Acemoglu et al. (2001) challenge endogeneity issues in the literature by instrumenting institutions with settler mortality rates in former colonies. Rodrik et al. (2004) point out the importance of institutions within a democracy. Besley and Kudamatsu (2007) also focus on institutions, and show that autocracies with good institutions perform quite well.

not a direct measure of political risk. There might also be systematic differences between countries that are rich in natural resources, and those that are not. Natural resources may attract large foreign investments, but they may also affect the political climate in the country.[2]

When it comes to explaining political risk theoretically, many authors consider different formulations of the hold-up problem to explain various factors that might discipline host governments. Eaton and Gersovitz (1984) assume that foreign investments exhibit some intangible asset, for example managerial skill, that the foreign investor can pull out in case of expropriation. The cost of expropriation for the local government is thus a drop in productivity. Knowing this effect, the foreign investor only invests as long as these costs marginally surpass the gains from expropriation. In equilibrium there is thus no expropriation, but investment levels are suboptimal compared to a situation with enforceable contracts. Thomas and Worrall (1994) follow the same reasoning in an infinite-horizon model where the cost of expropriating is the loss of future investments. The optimal self-enforcing contract is one where investments are gradually built up and transfers to the host country increase over time. Along these same lines Schnitzer (1999) shows how an investor's control rights over some essential production factor can protect the investment. Like in Eaton and Gersovitz (1984), the threat of expropriation causes under-investment if these control rights are not sufficiently strong. Konrad and Lommerud (2001) point to asymmetric information and how joint ventures with local investors can be a possible solution to the hold-up problem. Profit shifting through intra-firm trade can protect some of the profit from creeping expropriation by the host government if the information about opportunity costs is private to the foreign investor. Selling shares in the affiliate to local investors increases this mechanism, as this increases information rent through profit shifting, and also reduces the local government's incentives to expropriate. Another potential solution to the risk of expropriation is presented by Engel and Fischer (2008), who propose contracts that give the authorities a larger share of windfall revenue when prices of the produced goods are high. Common to all these models is that a host country will expropriate whenever short-term gains from expropriating outweigh

---

[2]For a good overview of the literature on the 'natural resource curse', see Sachs and Warner (2001).

the long-term costs, but also that they all view the host country as one agent.

I introduce "new" domestic agents in a simple model where foreign investments are reduced by fear of expropriation, but where also the political institutions affect the risk of expropriation. The government's actions determine its reputation, which again determines future investments. I attempt to capture the trade-off between short-term gains from expropriation and the long-term reputation costs, but also how these effects are increased or reduced by different political institutions, and democratic imperfections. The argument follows the same line of reasoning as the "stationary bandit" story discussed for example in Olson (1993), where a stationary bandit can earn a bigger profit by promoting growth and "stealing" through taxation, instead of stealing all he can get in a one-time raid. Like in some of the above mentioned models, loss of future investment is one cost of expropriation, but in addition to this the personal utility from being in power might discipline an executive that would otherwise expropriate. However, as the future costs of expropriation might be the problem of a future president, frequent elections and the likelihood of being replaced at the end of the period might shift the focus of the incumbent towards the short-term gains of expropriation. The model is related to the above discussed literature, as well as a body of literature discussing expropriations and stochastic alteration of power (see for example Azzimonti and Sarte, 2007). Further it draws upon the probabilistic voting models from the political economy literature where welfare shifting towards the executive's own group may gain him support in spite of poor political performance. An example of this is Padró I Miquel (2007), who shows how an ethnic group may prefer a kleptocratic ruler to the risk of getting an equally corrupt ruler, from another ethnic group.

Although following in the footsteps of several strands of literature, my paper distinguishes itself in several important ways. To the best of my knowledge, the model is the first to predict that democratic competition for political power may increase the likelihood of expropriations. I show that elections have two separate effects on the executive's expropriation incentives: they reduce the likelihood of remaining in power, thus making the executive more short-sighted, but also provide a mechanism through which the likelihood of remaining in power is conditional on previous performance. In order for democratic competition to have a disciplinary

effect, the political competition must be above a critical level of responsiveness. Below this level, increased political competition will rather increase expropriation risk.

The paper also contributes to the empirical literature on expropriations. I test the above predictions on data for actual expropriations in seven mineral sectors in 27 developing countries between 1960 and 2002. Extractive sectors should be "ideal" for expropriations. Most of the investment costs are sunk at the establishment of the mine or the well, the key resource is perfectly immobile, and the technology needed for running an already operating mine should be attainable for most developing countries, once the exploration, drilling, and operation routines are in place. In addition, the final product is rather homogenous, and not protected by trademark property rights, meaning that if expropriated, the product should sell at a similar price in the international market as the original owner would have been able to achieve. The data seem to a large degree to support the theoretical predictions from the model.

The rest of this paper is organized as follows. Section 2 presents the theoretical model of the expropriation decisions a president makes, first in a simple model where voters cast their ballots strictly according to a social conflict type of election model, and thereafter in a more realistic version where voters take into consideration the incumbent's actions when deciding their vote. In section 3 I present the data and test the predictions from the theoretical model. Section 4 concludes.

2   The model

In this section I develop a simple Bayesian game voting model where I show how the limitations on executive power and degree of political competition for office affect the executive's incentives to expropriate a foreign investment. The model is probably most closely related to Besley and Kudamatsu (2007), but departs from this model in that I include foreign investors as actors in the game. I also allow for more variations in the level of democracy than in their model.

The political economy literature has long given great importance to social conflict between the ins and the outs, the empowered and the disempowered. When there are different payoffs for the social group with *de facto* political power and

other social groups, policies pursued by a country might very well be suboptimal for the country as a whole. This differs greatly from models where the agent is a country maximizing some social welfare function, and I show how the degree of stability in the allocation of *de facto* power has the potential to greatly affect the political risks for foreign investments in a country.

The model runs over two periods, and in each period the sitting president decides whether to expropriate foreign investments or not, and then divides the budget consisting of gains from foreign investments (expropriated profits or tax revenue) and other government revenue between voters through transfers directed at specific groups in society. The population consists of two groups, and these groups could be thought of as ethnic, religious, social, or any other division criteria that could identify one group of the population with a certain political party or candidate. For simplicity I assume that the groups are of the same size, and that transfers are equally divided between the members of each group. Between the two periods there might be democratic elections between the incumbent and a candidate from the opposing group where the winner is decided through a first-past-the-post ballot.

Voters have two-dimensional policy preferences. The expected utility-function for a representative voter is assumed to be

$$U_i^V = \sum\nolimits_{t=1}^{2} \left(y_{it} + \alpha_{it} \right).$$

(1)

I assume a utility from consumption that is simply equal to the amount of money spent on consumption goods. This, together with the assumption that the groups are of equal size and that transfers are equally distributed within the groups, means that I can simply write an individual's utility as a function of the total transfers to his group. The last component α_i denotes some personal preference for the sitting president relative to the opposition candidate, independent of the individual's social group. Voters thus take both their material payoff and their subjective preference for the incumbent into consideration when they vote. This could be thought of as cross-cutting cleavages in the Rokkan and Lipset terminology of voter alignments.

Political candidates get utility from winning office, and thus directing resources

towards their own social group. More specifically I assume that a candidate from group s has an expected utility-function

$$U^{P,s} = \sum_{t=1}^{2} \phi_t Y_t^s, \ s = 1, 2 \tag{2}$$

where $Y^s = \sum_{i \in s} y_i$ is the total value transferred to group s, and $\phi_t = \{0, 1\}$ indicates whether the candidate is president in period t or not. There is no discounting between periods for presidential candidates or voters.

In each period there exist two non-divisible investment opportunities in the host country (HC) that cannot be exploited by domestic investors due to lack of means or technical know-how. Investment costs are sunk at the beginning of the period, and output is produced at the end of the period, at which point the investment is undone. An identical investment opportunity will however be available in the next period. The output is equal for each investment, but the investment costs differ. The price per unit of this output is determined in an international market that is unaffected by the host country's supply, and without loss of generality I chose units of output and the expected price so that the operating profits of each investment at the end of the period can be normalized to 1. The realized price, however, might vary, and I denote by P a multiplier relative to the expected price so that $P > 1$ means that the realized price is higher than expected, and vice versa. If there is no risk of expropriation in HC a foreign investor will thus undertake the investment if profits net of taxes are larger than the initial investment, $(1 - T) \geq I_j \ j = 1, 2$, where T is a flat tax rate.[3]

However, after the investment is realized, the local authorities might decide to expropriate the produced output, in which case the country receives the full value of the produced output. If the host country does not expropriate it receives a share T of this revenue through taxation. The investor will take this into account when deciding whether to invest or not. Let there be two kinds of presidents: an

---

[3]Naturally, the tax rate is also a choice variable for the president in this type of model, and can be used for creeping expropriation. However, when presidents are assumed to be either of an expropriating type or investor-friendly, such creeping expropriation would also reveal the president's type. As such, if trying to expropriate through taxation, the president would rationally set $T = 1$, which would yield identical results to direct expropriation. I will therefore ignore the possibility of expropriation through taxation, and take the tax rate as given in this paper.

opportunistic, that will expropriate foreign investments whenever this maximizes his expected utility, and an investor-friendly who will never expropriate. A president's type is unobservable to both investors and voters. Denote by $\Pr(o) = \pi$ the share of potential presidential candidates that are of the opportunistic type, and hence the prior probability that any sitting president is of this type. When the sitting president is newly elected, the foreign investor will thus invest in the second period if $(1-T)[1-(1-\mu)\pi] \geq I_j$, where $\mu \epsilon [0,1]$ captures constraints on the executive, as the host country may have institutions that will overturn the executive's decision to expropriate. These constraints can be thought of as entities with veto power over the executive's decisions, for example a congress, senate or supreme court. The nature of veto power means that the probability that an expropriation will be stopped is equal to the probability that at least one of the constraining institutions is controlled by individuals or groups that will lose from the expropriation. For simplicity this probability is seen as exogenous throughout the paper.

I assume that investment costs are such that

$$(1-T) > I_2 > (1-T)[1-(1-\mu)\pi] > I_1 > (1-T)\mu, \tag{3}$$

which implies that whenever the type of the executive is unknown (with the proxy π determining the expected type), investment 1 will be profitable in expectation, and will thus be realized. Investment 2, however, is more costly to undertake, and will only be realized if the investor has received some further information indicating that the executive is of the investor-friendly type, i.e. has obtained some adjusted $\tilde{\pi} < \pi$. When the executive has shown himself to be of the expropriating kind, no investments are profitable for the investor, and none will be realized. In addition to income from foreign investments, the executive can also distribute domestically generated state revenues, G.

The executive could be a president, a king, a high priest, or any other form of government, but for the rest of the paper I will denote the executive as president. As the case with an investor-friendly president is quite uninteresting, as they will never expropriate, I will only focus on the expropriation incentives for an opportunistic president.

The timing of the game is as follows:

1. [Start of period 1] Investor invests

2. President decides on expropriation

3. State revenue is divided among voters and consumed

4. Elections might be held

5. [Start of period 2] Investor decides whether to invest

6. President decides on expropriation

7. State revenue is divided between voters and consumed

The model is solved through backward induction. The last steps are trivial. At stage seven, in the final period, the president transfers as much to his own group as is possible. Being the last period, and since I am only considering opportunistic presidents, there are no arguments against expropriating, so if the investor invested, the revenue from the investment will be expropriated in stage 6. Knowing this, the investor will only invest in stage 5 if the sitting president did not expropriate at stage 2, or if the sitting president was newly elected at stage 4. The interesting stage in the model is thus stage 2. Figure 1 shows the game tree for stages 2-5 of the above described game. I denote by U_{1-4} the second-period expected utility in the outcomes that are possible in the model. The assumption (3) ensures that U_1 is the only realized outcome if an expropriating president remains in power in period 2. Similarly, it ensures that $U_2 = 0$ is the only relevant outcome if a new president is elected at stage 4. If a non-expropriating president stays in power in period 2 the investors will always invest in the second period, but whether only I_1 or $I_1 + I_2$ will be realized depends on parameter values. The value and likelihood of U_{1-4} determine expected utility for the president in the possible outcomes of the game, and thus also his incentives for expropriating in period 1. The three relevant expected utility-expressions, conditional on the president's

President [Stage 2]

Expropriate   Not expropriate

Elections [Stage 4]

Re-election   New president   Re-election

Investor [Stage 5]

Invest 0   Invest 0   Invest 0
Invest I_1 Invest I_1 Invest I_1
Invest $I_1 + I_2$ Invest $I_1 + I_2$ Invest $I_1 + I_2$

U_1 U_2 U_3 U_4
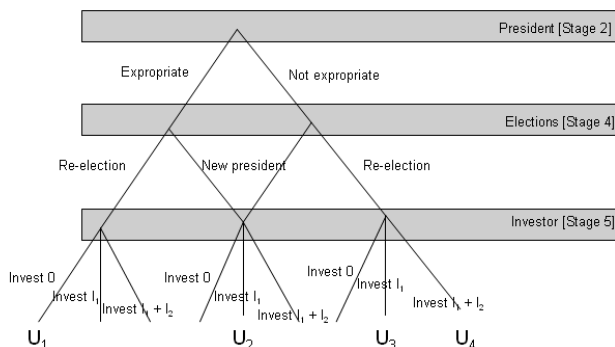
Figure 1: Game tree

actions at stage 2, will be:[4]

If expropriating at stage 2:

$$U_x^P = \text{Utility first period with expr} + \Pr\left(\text{reelection} \mid \text{expr}\right) U_1$$

If not expropriation at stage 2, I_1 invested in the second period:

$$U_{\bar{x},1}^P = \text{Utility first period without expr} + \Pr\left(\text{reelection} \mid \text{no expr}\right) U_3$$

If not expropriation at stage 2, $I_1 + I_2$ invested in the second period:

$$U_{\bar{x},2}^P = \text{Utility first period without expr} + \Pr\left(\text{reelection} \mid \text{no expr}\right) U_4$$

It is important to note that the different probabilities above depend on the president's action, but also on parameter values that determine how much information the realized action provides.

Specifically, the expected utility for the president if he expropriated in the first

---

[4]All expressions also include the probability of being ousted through elections, multiplied by the resulting utility U_2, which is equal to zero, and thus disappears in all cases.

period will be:

$$U_x^P = \{[1 - (1 - T)\,\mu]\,P + G\}\,\sigma + [1 - (1 - \gamma_x)\,\theta]\,G\sigma.$$

The first part of the expression, $\{[1 - (1 - T)\,\mu]\,P + G\}\,\sigma$, is the expected utility from transferring money to his group in period one. The rest of the expression is the expected utility for period two after expropriating in the first. I let θ denote the probability of facing elections between the periods, which I will interpret as the level of political competition, and γ_x denotes the likelihood of winning the elections after expropriating in the first period.[5] The combination of these, $[1 - (1 - \gamma_x)\,\theta]$, is thus the probability of remaining in office in the second period. The second part of the expression is then the probability of remaining in office, multiplied by the utility this will generate. Note that the transfers will now only be σG since the investor will not invest in the second period. If the president waited for the investor to invest in the second period and then expropriated he would get utility depending on whether second period investments are I_1 or $I_1 + I_2$ respectively:

$$U_{\bar{x},1}^P = (TP + G)\,\sigma + [1 - (1 - \gamma_{\bar{x}})\,\theta]\,([1 - (1 - T)\,\mu] + G)\,\sigma$$

$$U_{\bar{x},2}^P = (TP + G)\,\sigma + [1 - (1 - \gamma_{\bar{x}})\,\theta]\,([1 - (1 - T)\,\mu]\,2 + G)\,\sigma.$$

The expected utilities are all functions of the level of political competition, θ. Various levels of this variable define three types of equilibria in the model; for low levels of θ incentives to expropriate will be so low that no president will ever expropriate, and we have a pooling equilibrium where presidents of both types act identically. For sufficiently high levels of θ opportunistic presidents will always expropriate, and we have a separating equilibrium, where the two types of presidents always act differently. Intermediate levels of θ will cause opportunistic presidents to play mixed strategies, and there will be a semi-separating equilibrium. In the following I will solve the different equilibria.

Initially I let the voters give their support purely out of group identity, and

---

[5]Political competition does unarguably consist of more than just the regular occurrence of elections, but the risk of facing a real challenge by an opposing candidate constitutes an important aspect of this. More than probability of elections, θ could represent the frequency of elections.

they will never change their vote according to the president's behavior, hence $\gamma_{\bar{x}} = \gamma_x = \gamma$. In section 2.2 I relax this assumption. An opportunistic president will expropriate in the first period if the expected utility of doing so is larger than the expected utility of waiting and expropriating in the second period. Knowing this, the foreign investor will condition his investments in the second period, on the actions of the president in the first period. For the investor to undertake both investments in the second period, it must be the case that the president had incentives to expropriate in the first period and forego twice the incoming investments in the second period, and did not do so. Mathematically, this incentive constraint can be derived as below.

Taking the difference between the expected utilities gives the net incentives for the president to expropriate.

$$U_x^P - U_{\bar{x},1}^P = (1-T)(1-\mu)P - [1-(1-\gamma)\theta][1-(1-T)\mu] \tag{4}$$

$$U_x^P - U_{\bar{x},2}^P = (1-T)(1-\mu)P - 2[1-(1-\gamma)\theta][1-(1-T)\mu]. \tag{5}$$

Both of these expressions are clearly increasing in θ. Denote by $\check{\theta}$ the value of θ that makes the president indifferent between expropriating or not in equation (4), and similarly $\hat{\theta}$ for (5).

$$\check{\theta} = \frac{1-\mu(1-T)-(1-T)(1-\mu)P}{(1-\gamma)[1-(1-T)\mu]} \tag{6}$$

$$\hat{\theta} = \frac{2[1-\mu(1-T)]-(1-T)(1-\mu)P}{2(1-\gamma)[1-\mu(1-T)]}. \tag{7}$$

For $\theta < \check{\theta}$ expected utility from expropriating is always lower than from not doing so, meaning that $\theta \epsilon [0, \check{\theta}]$ defines the pooling equilibrium. Similarly, for $\theta > \hat{\theta}$ opportunistic presidents will always expropriate, defining the separating equilibrium as the outcome when $\theta \epsilon (\hat{\theta}, 1]$. It is straightforward to show that $\hat{\theta} > \check{\theta}$, meaning that there exists an interval, $\theta \epsilon (\check{\theta}, \hat{\theta}]$, where there will be a semi-separating equilibrium.

For intermediate levels of democratic competition, θ, the president will have to weigh two opposing factors against each other. On the one hand political competition represents a risk of foregoing a potential gain by not expropriating

in the first period should the president be replaced between the periods. On the other hand, by not expropriating the president sends a signal to potential investors that he might be of the investor-friendly type, increasing the probability that both investments will be realized in the second period, and thus increasing the potential gains of expropriating in the second period. In equilibrium it must be the case that both the president and the investor randomizes between their strategies in such a way that they both leave the other indifferent between the two possible actions. Denote the probability that the president will expropriate by p^P. The investor will then estimate a posterior risk that the president is of the opportunistic type based on the observation that no expropriations occurred in the first period:

$$\Pr\left(o \mid \bar{x}\right) = \frac{\left(1 - p^P\right)\pi}{1 - p^P\pi}.$$

This means that the expected profits from the investment in the second period after observing that the sitting president did not expropriate in the first period will be

$$\Pi^E = (1 - T)\left(1 - (1 - \mu)\frac{\left(1 - p^P\right)\pi}{1 - p^P\pi}\right).$$

If there were no expropriations in the first period, the investor will decide to invest in either one or both of the projects. The p^P must be chosen so that he is indifferent between the two options:

$$p^P = \frac{I_2 - (1 - T)(1 - (1 - \mu)\pi)}{(I_2 - (1 - T)\mu)\pi}. \tag{8}$$

Further, for this to be an equilibrium, it must be the case that when the investor then decides to invest $I_1 + I_2$ with probability p^I in the second period, the president is indifferent between expropriating or not in the first period.

$$U_x^P = p^I U_{\bar{x},2}^P + \left(1 - p^I\right)U_{\bar{x},1}^P$$
$$= U_{\bar{x},1}^P + p^I\left(U_{\bar{x},2}^P - U_{\bar{x},1}^P\right).$$

$U_{\bar{x},2}^P - U_{\bar{x},1}^P$ is always positive, which implies that for $\theta < \check{\theta}$, p_I must be negative for the above to hold with equality, and there will be a pooling equilibrium where

neither investor-friendly nor opportunistic presidents will expropriate. If p^I is at its maximum value one, the condition collapses to

$$U_x^P = U_{\bar{x},2}^P,$$

which is the case when $\theta = \hat{\theta}$. This implies that for $\theta > \hat{\theta}$, p^I must be larger than one for there to be a separating equilibrium, which is impossible. Hence I have shown that for levels of θ that are below $\breve{\theta}$ we get the pooling equilibrium, from this cutoff, and up until $\hat{\theta}$ the semi-separating equilibrium will be the outcome, whereas for higher levels of political competition there will be a separating equilibrium. In this equilibrium we have that the investor will undertake both investment opportunities in the second period with probability

$$
\begin{aligned}
p^I &= \frac{U_x^P - U_{\bar{x},1}^P}{U_{\bar{x},2}^P - U_{\bar{x},1}^P} \\
&= \frac{(1-T)(1-\mu)P - [1-(1-\gamma)\theta][1-(1-T)\mu]}{[1-(1-\gamma)\theta][1-(1-T)\mu]\sigma},
\end{aligned}
\tag{9}
$$

which is an increasing function of θ, and the president will expropriate in the first period with probability p^P defined in (8).

2.1 Comparative statics

From this most simple version of the model, I can derive at least three interesting, testable predictions concerning the president's incentives to expropriate in the first period. If the realized price is as expected, both $\breve{\theta}$ and $\hat{\theta}$ are positive, but they will decrease for higher values of P. It is also quite straightforward to see that incentives to expropriate are higher for higher price levels of the output. This leads to the following predictions:

Prediction 1: *Expropriations are more likely when prices are high.*

The intuition behind this is simply that if the value of a foreign investment today increases, relatively to the expected value of the same investment in the future, potential gains from expropriation increase, and *ceteris paribus* incentives to expropriate increase.

If chances of being challenged in a democratic election are zero, $\theta = 0$ incentives to expropriate are always negative unless $P \geq 2\frac{(1-\mu(1-T))}{(1-T)(1-\mu)} > 2$, i.e. the price of the output is more than twice the expected level. This also implies that under the given assumptions, the long-term welfare of the host country is maximized by not expropriating in the first period.

Prediction 2: *Expropriations are increasing in the level of democracy.*

Low levels of θ yield a pooling equilibrium where expropriations never occur in the first period. For higher levels, in the semi-separating equilibrium, expropriations occur with probability πp^P, whereas for the highest values of θ, in the separating equilibrium, expropriations occur with probability π. The mechanism at work here is simply that as the likelihood of an election increases, the likelihood of remaining in office, $[1 - (1 - \gamma)\theta]$, decreases, thus working as an increased discounting factor.

Prediction 3: *Expropriations are less likely when political constraints are high.*

Looking at $U_x^P - U_{\bar{x},i}^P$, $i = 1,2$ it can be shown that the expressions are decreasing in μ for any relevant values of the model's parameters.[6]

In countries with low levels of political competition investor-friendly presidents have no way of credibly showing investors their true type since no information is transmitted through their decision to not expropriate. In the high-competition countries, where the separating equilibrium will be the outcome, the absence of expropriations will always lead to both investment projects being realized unless the president is removed in elections between the periods. In the semi-separating equilibrium no expropriations in the first period will lead to expected investments in the second period of $1 + p^I \epsilon(1,2)$. Further, p^I is an increasing function of θ, meaning that conditional on no expropriations occurring in the first period, expected investments in the second period are increasing in the degree of political

---

[6]The proof that this is the case for $U_x^P - U_{\bar{x},1}^P$ is trivial, but for $U_x^P - U_{\bar{x},2}^P$ it is a bit, but not much, more complex. Taking the difference of the expression with respect to μ yields $(1 - T)(1 - 2\theta(1 - \gamma))$ which is negative for $\theta > \frac{1}{2(1-\gamma)}$. Inserting $\theta = \frac{1}{2(1-\gamma)}$ into the incentive expression yields expropriation incentives equal to $-T < 0$, implying that for such low values of democracy the president will never intend to expropriate anyway. This means that for any level of democracy where the president might want to expropriate, increasing levels of political constraints reduce the incentives to expropriate.

competition. This means that expected investments in the second period, conditional on the absence of expropriations in the first period, will be increasing in the level of political competition. This explains the somewhat counter-intuitive story that foreign investments may be higher in countries with a high level of political competition at the same time that I argue that presidents in these countries have stronger incentives to expropriate.

2.2 Endogenous election outcomes

One aim of this paper is to model expropriations in a social conflict framework. For one, changing regimes is one source of uncertainty that makes investment decisions difficult for foreign investors. Further, it seems like a rather strong assumption that voters' preferences are not affected by whether the president expropriates or not. In the following I will thus show how the model changes when voters may change their vote as a response to expropriations. Voters are rational, forward-looking individuals who cast their vote to maximize expected future utility. A changed vote is therefore not a direct punishment for past expropriations, but merely a response to new information about the incumbent president's type.

Voters share the same information as the foreign investor about the incentives to expropriate for the president, and can determine the level of second period investments and probability of expropriations in the second period. If the incumbent expropriated in the first period and remains in office, investments in the second period will be zero. If expropriations did not occur in the first period and the president is reelected, voters will find themselves in one of the following three situations.[7]

1. The democracy level in the country is below $\check{\theta}$, expected investments in the second period are I_1, and the risk of expropriation is π.

2. The democracy level in the country is between $\check{\theta}$ and $\hat{\theta}$. Investments in the second period will be $I_1 + I_2$ with probability p^I, and only I_1 with probability $\left(1 - p^I\right)$, and will be expropriated with probability $\frac{\left(1 - p^P\right)}{1 - \pi p^P}\pi < \pi$.

---

[7]Note that the cutoff points, $\check{\theta}$ and $\hat{\theta}$, are now different from the simple case above, as they are endogenously determined by the responsiveness of the voters to the president's actions.

3. The democracy level in the country is above $\hat{\theta}$. Investments in the second period will be $I_1 + I_2$, and no expropriations will occur.

In all situations election victory by the opposition candidate will mean that second period investments will be I_1, which will be expropriated with probability π.

This means that in the first case voters have received no information from observing outcomes in the first period, and have to choose between two seemingly identical candidates from different social groups. They will thus vote according to social identity and their exogenous preference shock α_i. In this case the model collapses to the simplified version above where $\gamma_{\bar{x}} = \gamma_x = \gamma$.

In case number 2 a voter from the incumbent's own group will have expected utility in period 2 if the incumbent wins of:

$$\left[\frac{(1 - p^P)}{1 - \pi p^P} \pi \left[p^I 2 \left(1 - \mu\right) \left(1 - T\right) + \left(1 - p^I\right) \left(1 - \mu\right) \left(1 - T\right) \right] + T + G \right] \sigma + \alpha_i,$$

and conversely, if the opposition candidate wins:

$$\left[\pi \left(1 - \mu\right) \left(1 - T\right) + T + G \right] \left(1 - \sigma\right).$$

This means that a utility-maximizing voter will support a non-expropriating candidate from his own group if:

$$\alpha_i > \pi \left(1 - \mu\right) \left(1 - T\right) \left(1 - \frac{1 - \pi p^P + \left(1 - p^P\right) \left(p^I + 1\right)}{1 - \pi p^P} \sigma \right) - \left(T + G\right) \left(2\sigma - 1\right).$$

If for similar levels of democracy the president did expropriate a voter from his own group will support the incumbent if

$$\alpha_i > \pi \left(1 - \mu\right) \left(1 - T\right) \left(1 - \sigma\right) - G \left(2\sigma - 1\right) - T \left(\sigma - 1\right).$$

Let α_i be distributed according to a continuous and twice differentiable probability distribution function f, with the corresponding cumulative distribution function F, independently of group identity. The expected change in support from his own

group for an expropriating incumbent compared to not expropriating can then be expressed as

$$\Pr\left(s, s \mid x\right) - \Pr\left(s, s \mid \bar{x}\right)$$

$$= \int_{\pi(1-\mu)(1-T)(1-\sigma)-G(2\sigma-1)-T(\sigma-1)}^{\pi(1-\mu)(1-T)\left(1-\frac{1-\pi p^P+\left(1-p^P\right)\left(p^I+1\right)}{1-\pi p^P}\sigma\right)-(T+G)(2\sigma-1)} dF < 0.$$

This can be shown to be negative when $-\frac{(1-T-I_2)\left(1+p^I\right)\sigma}{(1-\mu)(1-T)\pi} - T\sigma < 0$, which is always the case.

Voters from the opposition group solve a similar problem, and the change in support for the incumbent from opposition group voters if he expropriates is thus

$$\Pr\left(s, \bar{s} \mid x\right) - \Pr\left(s, \bar{s} \mid \bar{x}\right)$$

$$= \int_{\pi(1-\mu)(1-T)\sigma+G(2\sigma-1)+T\sigma}^{\pi(1-\mu)(1-T)\frac{\sigma\left(1-\pi p^P\right)-\left(1-p^P\right)\left(p^I+1\right)(1-\sigma)}{1-\pi p^P}+(T+G)(2\sigma-1)} dF < 0,$$

which is also negative since $-\frac{(1-T-I_2)\left(1+p^I\right)(1-\sigma)}{(1-\mu)(1-T)\pi} - (1-\sigma)T < 0$. Since p^I is increasing in θ it is plain to see that the responsiveness of the voters is increasing in θ as well. This means that in countries with stronger political competition the disciplining effect of elections is stronger.

In the third case, when $\theta \geq \hat{\theta}$ the voters' problems are

$$\Pr\left(s, s \mid x\right) - \Pr\left(s, s \mid \bar{x}\right)$$

$$= \int_{\pi(1-\mu)(1-T)(1-\sigma)-G(2\sigma-1)-T(\sigma-1)}^{\pi(1-\mu)(1-T)(1-\sigma)-G(2\sigma-1)-(3\sigma-1)T} dF < 0,$$

since $-2\sigma T < 0$, and

$$\Pr\left(s, \bar{s} \mid x\right) - \Pr\left(s, \bar{s} \mid \bar{x}\right)$$

$$= \int_{\pi(1-\mu)(1-T)\sigma+G(2\sigma-1)+T\sigma}^{\pi(1-\mu)(1-T)\sigma+G(2\sigma-1)+T(3\sigma-2)} dF < 0,$$

since $-2\left(1-\sigma\right)T < 0$. This shows that when voters respond to new information

about the president's type, the likelihood of winning an election is larger when not expropriating than when expropriating, i.e. $\gamma_{\bar{x}} > \gamma_x$. It can also be shown that this effect is stronger in more democratic countries. Implementing this into equations (6) and (7) yields some interesting results. The net incentives to expropriate are now

$$U_x^c - U_{\bar{x},1}^c = (1-T)(1-\mu)P - (\gamma_{\bar{x}} - \gamma_x)\theta G - [1 - (1-\gamma_{\bar{x}})\theta][1 - (1-T)\mu]$$
(4')

$$U_x^c - U_{\bar{x},2}^c = (1-T)(1-\mu)P - (\gamma_{\bar{x}} - \gamma_x)\theta G - 2[1 - (1-\gamma_{\bar{x}})\theta][1 - (1-T)\mu].$$
(5')

Solving these as in the previous sub-chapter, yields the analogue cut-off points that determine the areas for the pooling-, separating-, and semi-separating equilibria:

$$\check{\theta} = \frac{1 - \mu(1-T) - (1-T)(1-\mu)P}{(1-\gamma_{\bar{x}})[1 - \mu(1-T)] - (\gamma_{\bar{x}} - \gamma_x)G}$$
(6')

$$\hat{\theta} = \frac{2(1 - \mu(1-T)) - (1-T)(1-\mu)P}{2(1-\gamma_{\bar{x}})[1 - \mu(1-T)] - (\gamma_{\bar{x}} - \gamma_x)G}.$$
(7')

Comparing (6) to (6') it is clear that the cutoff point that determines the end of the pooling equilibrium, $\check{\theta}$, is higher when voters are responsive to the president's actions. Similarly it is easy to show that the cutoff point described in (7') is larger than the one in (7). In sum, this shows that overall, expropriations are less likely when voters are responsive to the president's actions than when they vote strictly according to group identity. Starting from the situation where $\gamma_{\bar{x}} = \gamma_x = \gamma$, and gradually letting the former increase while the latter decreases, it is trivial to show that incentives to expropriate in (4') and (5') fall, and the cutoff-points $\check{\theta}$ and $\hat{\theta}$ in (6') and (7') go towards infinity. This implies that for sufficiently responsive voters, there will never be expropriations in the first period. This is due to the fact that as voters get more responsive to the president's actions, incentives for opportunistic presidents to pretend to be of the investor-friendly kind increase. This is in itself interesting, and similar to the results in Padró I Miquel (2007). However, comparative statics are not very interesting in the case where $\check{\theta} > 1$ and expropriations in the first period never occur. I thus focus on parameter values

that assure that the three kinds of equilibria all still exist in the following.

Looking at the differentials with respect to P and μ we see that they follow the pattern from the simple model, and predictions 1 and 3 do therefore still hold in the simple two-period framework. However, the effect of higher levels of political competition is now no longer linear.

$$
\frac{\partial \left(U_x^P - U_{\bar{x},1}^P \right)}{\partial \theta}
$$
$$
= \left[1 - \mu \left(1 - T \right) \right] \left[\left(1 - \gamma_{\bar{x}} \right) + \theta \frac{\partial \left(1 - \gamma_{\bar{x}} \right)}{\partial \theta} \right] - \left(\gamma_{\bar{x}} - \gamma_x \right) G - \theta G \frac{\partial \left(\gamma_{\bar{x}} - \gamma_x \right)}{\partial \theta}.
$$

Here the additional terms compared to the simpler version are both negative, with absolute values increasing in θ, meaning that for sufficiently high levels of θ, a further increase in θ will reduce incentives to expropriate. This implies that if the shape of f is such that voters are very responsive to the president's actions, and θ is sufficiently large, the disciplining effect of elections will dominate the discounting effect of the same for lower values of democracy. In terms of the model, this can be explained as follows. Assume a θ-value below $\check{\theta}$, so that we are in the pooling equilibrium where expropriations never occur in the first period. An increase in political competition will now have two effects; it will move the president's incentive constraint towards the ex-ante cutoff point $\check{\theta}$, at which point expropriations will become a viable option. Simultaneously, however, the increase in θ will also define a new value for $\check{\theta}$, that is higher than before, thus moving the incentives relatively further away from the cutoff point. When voters are sufficiently responsive, the increase in $\check{\theta}$ will be larger than the increase in θ, thus in total moving the president's incentive constraint down the scale, relative to the cutoff points. For values of θ just above $\check{\theta}$, an increase in θ may thus cause the new situation to move from the semi-separating equilibrium to the pooling equilibrium, thus reducing the risk of expropriation.

Prediction 4: *In countries with sufficiently high levels of political competition, and sufficiently responsive voters, expropriations are less likely.*

In countries with low levels of democracy increases in the likelihood of elections will increase the incentives to expropriate, but this effect decreases as voters

79

become more responsive to the incumbent's performance. The discounting effect of elections is linear in the level of political competition. The disciplining effect through voters' responsiveness, however, is zero for $\theta = 0$, but negative and concave as this variable increases. The total effect of an increase in political competition will thus start out as increasing the likelihood of expropriations, but this effect will decrease as levels of competition increase, and may under certain parameter values turn negative for sufficiently high levels of political competition.

Note also that the effect of an increase in θ is lower for higher values of G, the size of other state revenue, relative to revenue from foreign investments. This means that:

Prediction 5: *When revenues from foreign direct investments are high relative to other national income, the risk of expropriations is more responsive to increases in political competition.*

The mechanism behind this result is that expropriating in the first period has two sources of potential future costs for the president: the lack of investment in the second period, and the increased possibility of being removed from office in an election. This second cost is obviously increasing in the size of the available funds that the president would have if he remained in office, and thus the disciplining effect of elections is increasing in the relative importance of other income relative to foreign investments.

The model, with its implications, thus extends the existing literature in several important ways. First of all, democracy, social divisions, and uncertainty of continuing as president may introduce a gap between the president's and the voters' interests, where a president may rationally chose to expropriate a foreign investment although this generates welfare losses for the country as a whole. Further, the model predicts that, at least below a certain level of political competition, an increased likelihood of elections will increase expropriation risks, whereas political constraints on the executive will always reduce these risks. In traditional principal-agent models this is typically not an issue, since the principal is a country as a whole that acts in its own self-interest as if governed by a total welfare-maximizing social planner. Thus, when decisions made by the executive may affect some groups of the population adversely, the possibility for these groups to veto the president's

decisions is important.

3   Data and econometric model

In this section I test the theoretical predictions derived from the model on a data set from Duncan (2006) on government expropriations in developing countries in seven major mineral sectors between 1960 and 2002. The data are collected at the sector level, and also contain information on countries producing lead, silver, copper, nickel, zinc, bauxite, and tin. In each sector the eight largest developing country exporters, based on average exports during the period 1965-75, are chosen. Expropriations are defined as "any seizure of assets, demand for equity stake or increase in taxes by the host government that was not a condition of the original contract" (Duncan, 2006). In the whole sample there are only 50 cases of expropriations, of which 20 were seizures of assets, 12 were uncontracted increases in equity, and the remaining 18 were uncontracted increases in taxes on the mining companies. Tax increases that were economy-wide, and not specifically aimed at the investment were not classified as expropriations. There was at least one expropriation in 18 of the 27 countries in the data set, and mines in all seven mineral sectors were expropriated. In addition to expropriations, the price variable is collected at the sector level, using the international price for each mineral in the sample as the relevant price for each sector, normalized so that the average price over the entire period, over all sectors, is equal to 1. All other control variables are at the country level.

For the political competition variable I use Vanhanen's (2000) political competition index, which is simply the relative size of the opposition measured by support in elections. I chose this variable among the many similar alternatives because it measures more directly the risk of being ousted from office through elections, compared with other indices that use a broader concept of democracy that often includes and overlaps partly with my political constraints variable. Testing the robustness of the results by replacing this variable with alternative specifications of political competition I also use Vanhanen's index of democracy, which is constructed by multiplying the political competition index variable by the degree of participation in the elections, as well as the Polity2 index from the Polity IV

database (Marshall et al., 2006), which ranges from -10 (strongly autocratic) to 10 (strongly democratic). Following a common practice in the literature (Persson and Tabellini, 2006; Jones and Olken, 2005; and others), I also construct a dummy taking the value one when the polity variable is positive, and zero otherwise. Different cut-off values for this dummy generate similar results, so these results are not reported here. Finally, I use a dummy for whether the country chooses its executive through elections (Marshall et al., 2006).

To capture the political constraints on the executive, I use the index constructed in Henisz (2000), which uses the number of veto players and their distribution of preferences as an indication of limits to the executive's power. In addition to being one of the most complete indexes over political constraints, this variable also enjoys the advantage of representing a similar specification of constraints on the executive to the one used in the theoretical model in the previous section. I also test the model proxying political constraints using a dummy that takes the value one when the executive answers to a legislative body, and zero otherwise. My goal in this section is to test the separate effects of political competition, and political constraints, so my preferred variables ex ante are the political competition index from Vanhanen, and the political constraint index from Henisz.

The institutional constraints on the executive pose a potential endogeneity problem. Expropriations may affect the political climate in several directions. It may coincide with or trigger a wave of nationalism, or it might even cause a strengthening of private property protection in order to attract new investments after old ones have been expropriated. To avoid any problems from this I follow Acemoglu et al. (2001) and use mortality rates of early colonists as an instrument for institutions. Put briefly, the intuition behind this is that in colonies where Europeans faced high mortality rates, they did not settle to the same degree as in colonies where the climate and the disease environment were favorable. Where they did not settle, they were more likely to set up extractive institutions, while where they settled they set up institutions with stronger property right protection. The institutions set up around 1900 have been shown to be quite persistent, and early settler mortality can thus be used as an instrument for institutions today. For a more thorough discussion of the suitability of this instrument, see Acemoglu et al. (2001; 2006) and Albouy (2008).

Country-specific control variables are taken from the IMF's International Financial Statistics database (IMF, 2007). Descriptive statistics are shown in Table 1, separately for the observations where expropriations took place, and in the ones where it did not.

Expropriation	Political competition	Political constraints	Colonists mortality rates	Price	Past expropriations	GDP per capita
No	26.98	0.21	4.64	0.99	0.06	2239.07
Yes	23.91	0.19	4.78	1.38	0.42	1972.17
Total	26.91	0.21	4.65	1.00	0.07	2231.07

Expropriation	Real GDP growth	International reserves per cap.	Balance of payments	Newly independent	Openness	N
No	3.65	125.67	-3.31	0.11	53.08	2,358
Yes	3.72	49.69	-1.34	0.14	70.44	50
Total	3.65	124.11	-3.26	0.11	53.47	2,408

Table 1: Descriptive statistics

The descriptive statistics show that the cases in which there were expropriations differ from the observations where no expropriations took place in several ways. Looking at the main variables of interest, political competition and political constraints, both are lower in the expropriation cases, although the difference does not seem striking. There are larger differences when we look at the current price of the mineral in question. In years where mines where expropriated, the price of the mineral that was expropriated was, on average, 38% above the sample price average. Past expropriations, defined as the occurrence of expropriation(s) in any of the past three years, are also much greater in the expropriation part of the sample. This variable is generated at the country level, and not the sector level, as the intuition goes that past expropriations damage the reputation of the government, independently of which sector it expropriated. Maybe somewhat surprisingly, real GDP growth, balance of payments, and the degree of openness are all higher in the expropriating cases, whereas GDP per capita and international reserves per capita are both higher in the non-expropriating cases. I will return to the discussion around the importance and intuition behind these variables below.

The empirical analysis studies the determinants of host country expropriations of foreign investments in various mining sectors. I test this by estimating variations over the following equation

$$\Pr\left(EXPR_{ijt} = 1|X_{ijt}, Z_{it}\right) = \alpha + X'_{ijt}\beta + \gamma'Z_{it} + \varepsilon_{ijt}$$
$$= \Phi\left(\alpha + X'_{ijt}\beta + Z'_{it}\gamma\right), \qquad (10)$$

where $EXPR_{ijt} = 1$ means that an expropriation took place in country i, in sector j in year t, and zero otherwise. $\Phi\left(\cdot\right)$ is the standard normal distribution function, α is a constant, and Z_{it} is a vector of controls, including region dummies and trend adjusting variables. These controls are discussed in more detail below. X_{ijt} is the vector of my variables of interest, namely political competition, political constraints, and price.

3.1 Results

Initially I test predictions 1 through 3, that expropriations are more likely when prices are high, when political competition is strong, and when political constraints are weak.

Regression (1) in Table 2 shows the most direct estimation of the theoretical model with only region dummies as controls. All variables are of the expected sign, and all but the political constraint variable are significantly different from zero. It can be argued that there might be reverse causality from expropriations to the price of the expropriated mineral. Simple regressions of the effect of expropriations on the price support this. Both the expropriation variable and the lagged expropriation variable have positive and significant effects on the price in simple regressions, also when it is controlled for time trends and world growth. I therefore use the lagged price variable in all further regressions. In regression (2) we see that the variables keep their sign, but that the lagged price variable is not significant. Duncan (2006) uses a "boom" dummy which takes the value one whenever the lagged price is more than 50% above the sample average, and found this to be a strong predictor of expropriations. Constructing the boom dummy from the lagged price, however, again yields insignificant results. There thus seem to be

some indications that above normal prices may drive expropriations, but without any good instruments to control for the possible reverse causality it is difficult to draw any strong conclusions on this question. As this is not the main focus of this paper, I leave this discussion to others, and use the lagged price variable in the following.

As argued before, there might also be endogeneity issues with the political constraint variable, which I aim to solve by instrumenting political constraints by settler mortality rates. The results can be seen in regression (2) in Table 2. The political constraint variable is still negative, but now also statistically significant. The validity of settler mortality as an instrument for institutions that secure property rights has been argued above. The first-step regression shows that settler mortality has strong explanatory power for the degree of political constraints. A case could be made, however, that it could also be used as an instrument for democratic rights, including political competition, protecting the citizens' right to influence the government. To check for this I try using settler mortality as an instrument for political competition, but the first-step regression now shows that settler mortality has no significant effect on this variable, and I take this to show that my instrument captures the right aspects of democracy. I also run a rare effect logit as suggested by King and Zeng (2001), that should adjust for the fact that the dependent variable only takes the value one in about 5% of the observations. The results from this regression are not significantly different from the regular probit regression, but complicates the use of instrumental variables, so I choose to use the regular probit model in the rest of the tests.

Previous literature and intuition suggest that there may be factors affecting the probability of expropriations other than political competition, political constraints, and prices. I thus control for a series of other variables, individually and simultaneously. Regression (3) shows the results from the full model. In the theoretical model no investments take place if the sitting president expropriated in the previous period, as expropriations would be certain to happen in the second period in this case. In real life there is uncertainty along other a number of dimensions, and some investments might occur even after a country has expropriated in one period. In this situation the executive does not have much reputation to lose by expropriating again however, and may be more likely to do so again. To control

for this I use past expropriations, as defined above. The sign of this variable is positive as expected, and statistically significant.

	(1) Probit	(2) IV-Probit	(3) IV-Probit	(4) IV-Probit
Political constraints	-0.210 (-0.78)	-3.455*** (-4.71)	-4.094*** (-10.26)	-6.521*** (-3.60)
Political competition	0.0102*** (3.30)	0.0227*** (8.65)	0.0170*** (5.59)	0.0272*** (4.17)
Price	0.487*** (3.53)			
Lagged price		-0.0295 (-0.26)	-0.00143 (-0.01)	0.0711 (0.43)
Past expropriations			0.308*** (2.65)	-0.0307 (-0.22)
Real GDP growth			0.00742 (0.73)	-0.0261 (-1.64)
GDP per capita			0.0000383 (0.68)	0.00176*** (5.73)
International reserves per capita			0.00601*** (3.33)	-0.0135*** (-3.56)
Balance of payment			0.0141 (1.19)	0.0553*** (3.13)
Newly independent			1.333*** (3.46)	1.383*** (3.32)
Government expenditure			0.0164*** (2.00)	0.00920 (0.80)
Openness			-0.00968** (-2.41)	-0.0107** (-2.48)
Country dummies	No	No	No	Yes
Region dummies	Yes	Yes	Yes	No
Pseudo R²	0.160			
N	1518	1277	956	601

t statistics in parentheses
$^* p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$

Table 2: Test of Predictions 1-3

Previous empirical studies have found that the expropriation risk is low when GDP growth is high (see for example Jones, 1984 and Picht and Stüven, 1991). I therefore include real GDP growth as a control variable, but this turns out to be insignificant. It is also possible that expropriations could be caused by the need

for foreign currency, whereas conversely it could also be argued that since many expropriations include some form of compensation, more international reserves make expropriations relatively less costly. The sign of this variable turns out to be positive, and statistically significant, indicating that the compensation story might have more hold in reality.[8]

The level of economic development should also be controlled for. It could affect the risk of expropriation directly, but it could also be important to control for it merely because it is expected to be correlated both with the level of political competition and political constraints, although the direction of the causality in these relationships is a topic that is beside the focus of this paper. I include GDP per capita as a proxy for economic development, and this turns out to be positive, but not significant. Further, it does not seem to affect the coefficient or significance of the political constraint, or the political competition variables.

In all regressions where I include the control variables individually and also in the full model, the variables of interest seem largely unaffected by the inclusion of various control variables, and remain of the expected sign and highly significant in all specifications. The conclusion that political competition increases the risk of expropriations when political constraints are controlled for seems to be robust for different specifications of the model. Since my data set is compiled from different and only partly overlapping sources, each included control variable causes me to lose observations. As the number of observations is already limited I therefore drop all control variables that do not have any effect on the coefficients of interest in the rest of the analysis.

Controlling for unobserved fixed effects would be natural in these regressions; however, consistent fixed effect estimators in probit-regressions is problematic. Some of the effect should be captured by the regional dummies, but for a better control I include country dummies. The results are shown in regression (4). Again, the main findings are unchanged, except that even more observations are dropped when I include country dummies. This is because countries that have never experienced expropriations are dropped since their corresponding dummy perfectly predicts the dependent variable. Since these observations do contain some rele-

---

[8]As shown below, though, this result seems to be driven by a few observations, and should be interpreted with caution.

vant information about the risk of expropriation, and since the main results do not seem to change drastically, I choose to use only the region dummies to control for unobserved fixed effects in the following.

To make sure that the results are not driven by the construction of the competition and constraint variables, I also test the results against different measures of political constraints and political competition. Table 3 shows the results for various combinations of alternative specifications of the variables of interest. In all regressions I include the control variables from regression (3).

	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Political constraints	-3.828*** (-9.97)	-4.065*** (-11.21)	-3.170*** (-4.84)	-3.638*** (-7.52)	-3.958*** (-10.26)					
Legislative body						-2.755*** (-21.56)	-2.928*** (-33.44)	-2.585*** (-23.60)	-2.585*** (-47.00)	-2.528*** (-41.77)
Index of democracy	0.0396*** (4.06)					0.0997*** (10.07)				
Political competition		0.0179*** (6.20)					0.0336*** (14.11)			
PolityIV			0.0928*** (6.80)					0.0962*** (8.57)		
PolityIV dummy				1.191*** (7.76)					1.001*** (8.68)	
Elections					1.147*** (5.84)					0.426*** (4.00)
N	993	993	1077	1077	993	1008	1008	1077	1077	993

t statistics in parentheses
* *p* < 0.10, ** *p* < 0.05, *** *p* < 0.01

Table 3: Alternative political indicators

In regressions (5) to (9) I still use Henisz's political constraint index, but replace Vanhanen's political competition index with alternative variables in the following order: First, the same author's index of democracy. Secondly I replace it with the commonly used Polity2 index, and the dummy variable constructed from the same index. Finally, I use a dummy for whether elections are held or not as an indicator of political competition. As the results show, all specifications of political competition significantly increase the probability of expropriations. Further, the political constraint variable remains negative and significant for all the alternative

specifications of competition.

In the second part of Table 3 I repeat the above exercise, but this time with political constraints measured with a dummy that takes the value one if the executive answers to a legislature, according to the ACLP database (Alvarez et al., 1996). Again all measures of political competition are positive and all are highly significant. Further, the political constraint measure is also negative and significant through all the different specifications of competition. This shows that in addition to being robust to other control variables, my findings seem robust to alternative measures of political constraints and -competition.

3.2 The disciplining effect of elections

Predictions 1 to 3 could be derived from the most simple version of the theoretical model. When voters were assumed to be responsive to the president's expropriation decisions, however, the effect of democracy became more complex. In this section I therefore test Predictions 4 and 5, that the effect of an increase in the political competition could reduce the probability of expropriations if the level of democracy is sufficiently high, or if the relative value of inward FDI is low.

It is hard to get good measures of voters' responsiveness, so my best test of Prediction 4 will be to study the non-linearity of the index of democracy on the probability of expropriations. I do this by including a squared term of the variable in a regression with the same specifications as regression (3).

Table 4 shows the results. The first regression (15) reports the results when testing Prediction 4 on the full sample. The signs of the political competition variable and its squar are both as expected, but only the linear variable is statistically significant. Investigating the relationship further, I find that there is a negative effect on the probability of expropriation when increasing the political competition from the minimum level of zero. Regression (16) thus shows the result when I exclude all observations in the sample with political competition equal to zero. The absolute values of the coefficients are now larger, and also strongly significant. This regression partly supports Prediction 4, that for a sufficiently high level of political competition, increased competition will work in a disciplinary manner, and reduce the probability of expropriations. For countries with a political competition

score of about 51 or more, further increases in political competition reduces the risk of expropriation. This is the case for about 20% of the observations in my sample. It should be pointed out, however, that all the countries in the sample are developing countries, and that the measured effects may not be representative for all countries. The difference between regressions (15) and (16) shows that there is also an effect at work in the countries with the weakest political competition that is not captured by my model.

	(15)	(16)	(17) Low FDI	(18) High FDI
Political constraints	-4.137***	-4.643***	-11.66***	-5.143***
	(-10.44)	(-11.64)	(-4.30)	(-27.70)
Political competition	0.019**	0.083***	0.00849	0.0221***
	(2.07)	(6.07)	(0.88)	(7.80)
Political competition squared	-0.0000349	-0.0007991***		
	(-0.21)	(-4.17)		
Lagged price	-0.0031	-0.1165	-0.129	-0.0234
	(-0.02)	(-0.97)	(-0.48)	(-0.20)
GDP per capita	0.000033	0.000204***	0.000865***	0.000925***
	(0.58)	(2.77)	(2.96)	(10.51)
International reserves per capita	0.00613***	0.00498**	-0.00104	-0.00254
	(3.44)	(2.03)	(-0.09)	(-1.61)
Openness	-0.01008**	-0.00556	0.0713***	-0.0132***
	(-2.56)	(-1.45)	(3.75)	(-5.92)
Past expropriations	0.305***	0.270**		
	(2.64)	(2.35)		
Real GDP growth	0.0078	-0.0087		
	(0.75)	(-0.55)		
Balance of payments	0.0136	0.0206		
	(1.18)	(1.49)		
Newly independent	1.316***	1.364***		
	(3.48)	(4.22)		
Government expenditure	0.0164**	0.0403***		
	(2.05)	(5.06)		
N	956	615	480	551

f statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Test of Predictions 4-5

The last two regressions in Table 4 show the results from splitting the sample between countries with high and low levels of inward stock of FDI relative to GDP, respectively. The coefficient for political competition is 2.6 times larger for

the high FDI sub-sample than for the low FDI sub-sample, in accordance with the theoretical predictions. However, a test to see if the coefficients are significantly different only yields a chi squared of 1.83, and the zero hypothesis of no statistical difference cannot be dismissed.

4   Conclusions

I have developed a theoretical model that explicitly shows three channels through which two important concepts of liberal democracy affect the likelihood of expropriations of foreign property. Whereas checks and balances that constrain the executive power always reduce the probability of expropriations, competition for political power has a more complex effect on such risks. For one, the risk of losing power through elections may work as an increased discount factor, making the president favor short-term gains over long-term costs, and thus increasing the likelihood of expropriations. On the other hand, if the president depends on the support of the voters in order to stay in office, he may be forced to act more in the long-term interest of the country as a whole, and expropriations might thus become less likely when competition for power is stronger. In my theoretical framework the former effect is dominant for low levels of democratic competition, but the net effect may be reversed for sufficiently high levels. Similarly high levels of incoming FDI relative to national revenue make short term gains relatively higher, compared to the potential loss of utility from being ousted in elections. In countries where national production is higher, relative to incoming FDI, the disciplining effect of elections will dominate the discounting effect, and more political competition may reduce the risk of expropriations.

Testing these predictions on a data set of actual expropriations in important mineral industries in developing countries, I find support for the main predictions of the model. Since the predicted effect of democratic competition is non-linear in both the importance of inward FDI and in the general level of democracy itself, it would be interesting to test if these results hold when including developed countries, and in general on a larger data set of expropriations that might allow for more complex robustness tests of the results. Even so, these findings highlight the importance of specifying which aspect of democracy one is using when measuring

its effect on different measures of economic performance. Many indices of democracy include both checks and balances and competition for power as parts of the measure, meaning that countries that according to this paper are very different with respect to their predicted respect for private property might still end up with the same democracy score, and thus influence the predicted effect of democracy on economic performance.

References

Acemoglu, D., Johnson, S., & Robinson, J. 2001. Reply to the Revised (May 2006) version of David Albouy's "The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data". *American Economic Review*, **91**, 1369–1401.

Acemoglu, D., Johnson, S., & Robinson, J. 2006. *The Colonial Origins of Comparative Development: An Empirical Investigation*. mimeo. MIT.

Albouy, D. 2008. *The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data*. NBER Working Paper Series 14130. National Bureau of Economic Research.

Alvarez, M., Cheibub, J., Limongi, F., & Przeworski, A. 1996. Classifying Political Regimes. *Studies in Comparative International Development*, **31**(2), 1–37.

Azzimonti, M., & Sarte, P. 2007. Barriers to Foreign Direct Investment Under Political Instability. *Economic Quarterly*, **93**(3), 287–315.

Besley, T., & Kudamatsu, M. 2007. *Making Autocracy Work*. mimeo. London School of Economics.

Busse, M. 2004. Transnational Corporations and Repression of Political Rights and Civil Liberties: An Empirical Analysis. *Kyklos*, **57**(1), 45–66.

Busse, M., & Hefeker, C. 2007. Political risk, institutions and foreign direct investment. *European Journal of Political Economy*, **23**, 397–415.

Bó, E., & Rossi, M. 2008. *Term Length and Political Performance*. NBER Working Paper Series 14511. National Bureau of Economic Research.

Duncan, R. 2006. Price or politics? An investigation of the causes of expropriation. *The Australian Journal of Agricultural and Resource Economics*, **50**, 85–101.

Eaton, J., & Gersovitz, M. 1984. A Theory of Expropriation and Deviations from Perfect Capital Mobility. *The Economic Journal*, **94**(373), 16–40.

Engel, E., & Fischer, R. 2008. *Optimal Resource Extraction Contracts under Threat of Expropriation*. NBER Working Paper Series 13742. National Bureau of Economic Research.

Hall, R., & Jones, C. 1999. Why Do Some Countries Produce So Much More Output Per Worker than Others? *Quarterly Journal of Economics*, **114**(1), 83–116.

Henisz, W. 2000. The Institutional Environment for Economic Growth. *Economics and Politics*, **12**(1), 1–31.

IMF. 2006. *International Financial Statistics*. International Monetary Fund.

Jensen, N. 2005. *Democratic Governance and Multinational Corporations: Political Regimes and Inflows of Foreign Direct Investment*. mimeo. Washington University.

Jones, B., & Olken, B. 2005. Do Leaders Matter? National Leadership and Growth Since World War II. *Quarterly Journal of Economics*, **120**(3), 835–864.

Jones, R. 1984. Empirical Models of Political Risks in U.S. Oil Production Operations in Venezuela. *Journal of International Business Studies*, **15**(1), 81–95.

King, G., & Zeng, L. 2001. Logistic Regression in Rare Events Data. *Political Analysis*, **9**(2), 137–163.

Knack, S., & Keefer, P. 1995. Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures. *Economics and Politics*, **7**(3), 207–227.

Kobrin, S. 1984. Expropriation as an Attempt to Control Foreign Firms in LDCs: Trends from 1960-1979. *International Studies Quarterly*, **28**(3), 329–348.

Konrad, K., & Lommerud, K. 2001. Foreign Direct Investment, Intra-Firm Trade and Ownership Structure. *European Economic Review*, **45**, 475–494.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., & Vishny, R. 1998. Law and Finance. *Journal of Political Economy*, **106**(6), 1113–1155.

Li, Q., & Resnick, A. 2003. Reversal of Fortunes: Democratic Institutions and Foreign Direct Investment Inflows to Developing Countries. *International Organization*, **57**, 175–211.

Lucas, R. 1990. Why Doesn't Capital Flow from Rich to Poor Countries? *American Economic Review*, **80**(2), 92–96.

Marshall, M., Jagger, K., & Gurr, T. 2006. *Polity IV Project, Political Regime Characteristics and Transitions, 1800-2004*. Maryland: Center for International Development and Conflict Management, University of Maryland, College Park.

Mauro, P. 1995. Corruption and Growth. *Quarterly Journal of Economics*, **110**(3), 681–712.

Olson, M. 1993. Dictatorship, Democracy, and Development. *The American Political Science Review*, **87**(3), 567–576.

Padró i Miquel, G. 2007. The Control of Politicians in Divided Societies: The Politics of Fear. *Review of Economic Studies*, **74**, 1259–1274.

Persson, T., & Tabellini, G. 2006. *Democratic Capital: The Nexus of Political and Economic Change*. CESifo Working Paper 12175. National Bureau of Economic Research.

Picht, H., & Stüven, V. 1991. Expropriation of Foreign Direct Investments: Empirical Evidence and Implications for the Debt Crisis. *Public Choice*, **69**, 19–38.

Rodrik, D. 1999. Where Did All the Growth Go? *Journal of Economic Growth*, **4**(4), 385–412.

Rodrik, D., Subramanian, A., & Trebbi, F. 2004. Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, **9**, 131–165.

Sachs, J., & Warner, A. 2001. The Curse of Natural Resources. *European Economic Review*, **45**, 827–838.

Schnitzer, M. 1999. Expropriation and control rights: A dynamic model of foreign direct investment. *International Journal of Industrial Organization*, **17**, 1113–1137.

Thomas, J., & Worrall, T. 1994. Foreign Direct Investment and the Risk of Expropriation. *Review of Economic Studies*, **61**, 81–108.

Titiunik, R. 2008. *Drawing Your Senator from a Jar: Term Length and Legislative Behavior*. Mimeo. UC Berkeley.

Vanhanen, T. 2000. A New Dataset for Measuring Democracy, 1810-1998. *Journal of Peace Research*, **37**(2), 251–265.

Chapter 4

Productivity, Size and the Disintegration of Industrial Production

IV

Productivity, Size, and the Disintegration of Industrial Production

Jonas Gade Christensen*

University of Bergen

Abstract

I develop a theoretical model of firms' sourcing decisions along the productivity dimension as in Antrás and Helpman (2004), while also incorporating task trade as in Grossman and Rossi-Hansberg (2008). The combination of these two effects permits a framework for sourcing strategies along two dimensions, which generates results where firms spread the production process of the final good over several different sources simultaneously. While reproducing the results from the aforementioned models, my model contributes refined and more detailed predictions. Testing these on firm-level data for Spanish manufacturing firms, I find strong empirical support for the model's predictions.

1 Introduction

Over the last decades there has been a disintegration of industrial production, in the sense that the production chain increasingly has been split up and undertaken at separate locations outside of the firm's own production plants in the home country. This disintegration can take the form of domestic outsourcing where firms buy inputs or services from other companies in the same country, or they may acquire these from subsidiaries or from unrelated companies abroad. All of these forms of external sourcing seem to be increasing (Feenstra, 1998; Hummels et al., 2001; Antràs and Helpman, 2004, 2008 and others). There are several explanations of these trends, but global integration seems to be a common denominator for most of them. Lower transportation costs have made it more profitable to produce intermediate inputs away from the assembly plant, and the technological advances has made it possible to undertake certain accounting, engineering, and programming services anywhere in the world. Further, cheaper and easier communication has facilitated a closer contact between headquarters and production plants when it comes to sending plans, giving instructions, etc.

The international trade literature has attempted to explain this development. Departing from the explanations of Melitz (2003) and Helpman et al. (2004) of how firms sort into different forms of accessing foreign markets for final goods consumption according to productivity levels, Antràs and Helpman (2004) develop a similar framework for explaining which kind of sourcing firms will choose, depending on their productivity levels. In a model where headquarter-services are produced in-house, but manufacturing is undertaken elsewhere, they show how incomplete contracts between the final-good producers and the producers of intermediate inputs may distort the quantity and quality of inputs away from optimal levels, thus affecting variable costs of production. These distortions may be reduced with vertically integrated plants, but the fixed costs of establishing such plants are higher than the costs associated with arm's-length outsourcing. More productive firms with larger production will be able to spread these fixed costs over more units, and will thus choose to vertically integrate production.

Grossman and Helpman (2002) argue that firms make 'make or buy' decisions, based on a trade-off between transaction costs of outsourcing associated

with searching and incomplete contracts, against governance costs of vertical integration. Grossman and Rossi-Hansberg (2008) extend this reasoning, while approaching the topic from a slightly different angle. They propose a theory that views the production of final goods as a series of tasks that have to be done to complete the product. They assume that tasks can be ordered according the their offshoring costs. These costs may be due to transportation, moral hazard problems, the importance of personal delivery of the tasks, or a combination of these or other reasons. The important assumption is that offshoring costs differ between tasks, and that these can be ordered in a non-decreasing manner. As with Antràs and Helpman, wages abroad are lower than in the home country, and firms will take advantage of this for all the tasks that can be undertaken cheaper abroad than at home. In their discussion, Grossman and Rossi-Hansberg focus on differences in skill intensities between industries, and the wage effects in the home country, rather than firm heterogeneity and selection into offshoring.

In this paper I combine the strength of both approaches in order to present a theoretical model that represents the actual sourcing strategies of firms more accurately than the previous literature has done. Starting from a model similar to that one of Antràs and Helpman (2004), I introduce the task-dimension from Grossman and Rossi-Hansberg (2008). This allows me to map firms' sourcing strategies along both the firms' productivity and the tasks' outsourceability, which generates predictions where firms use a combination of the available sourcing options available, both domestically and abroad. Apart from being a much more realistic representation of firms' real sourcing strategies, this mapping generates more detailed predictions than the previously mentioned literature. The model reproduces the predictions from Antràs and Helpman (2004) that only the most productive firms will source inputs from vertically integrated plants abroad through foreign direct investment (FDI), the somewhat less productive will source through arm's-length contracts with firms abroad. Even less productive firms will integrate domestically, and the least productive firms will buy inputs through arm's-length dealings with other domestic firms. However, I also show that firms will use several sourcing options simultaneously. Even the most productive firms may buy some inputs through arm's-length dealings domestically. Instead of sorting firms into four types depending on which sourcing option they use, my model predicts the intensities

with which each sourcing option is used as functions of output levels, or productivity.

Firm-level data from extensive surveys among Spanish manufacturing firms permits detailed testing of the theoretical predictions. With specific questions about the intensities in the use of each of the possible sourcing options, the data contain much more detailed information on firms' sourcing strategies than other, similar data sources. The empirical testing shows that the model to a large degree seems to describe the actual relationship between output levels, productivity, and sourcing strategies.

2   The model

In this section I develop a theoretical model in an attempt to explain why different firms choose different sourcing strategies. It is important to note that I use *sourcing option* to identify the source which a firm uses to undertake a specific task, whereas by *sourcing strategy* I consider the entire mix of different sources that the firm uses in the entire production. The empirical evidence in section 3 shows that about half of the firms in my sample use more than one sourcing option, meaning that a framework that allows for combinations of several different sourcing options is necessary for a realistic discussion about firm organization.

The economic environment in the model is common in the international trade literature, and I will not dwell with equilibrium conditions in the final goods market. Products are differentiated along the lines of Dixit-Stiglitz (1977), and heterogeneous firms enter and exit as in Melitz (2003).

Consumers have demand functions,

$$U = \sum_{j=1}^{J} \mu_j \ln \left[\int_0^{N_j} q_j \left(i \right)^{\rho_j} di \right]^{\frac{1}{\rho_j}}, \tag{1}$$

where μ_j denotes the share of total spending the consumer uses on varieties from industry j, $q_j \left(i \right)$ is consumed quantum of variety i in industry j, and $\rho_j \epsilon \left(0, 1 \right)$ is the degree of product differentiation between varieties in the industry. The constant elasticity of substitution in industry j can thus be denoted $\sigma_j = \frac{1}{1-\rho_j} > 1$. This

familiar setup yields inverse demand functions,

$$p_j(i) = A_j^{1-\rho_j} q_j(i)^{\rho_j - 1}.$$

From the individual firm's point of view, A_j is taken as constant, and expresses total expenditure on varieties from industry j, over the price index for the industry.

$$A_j = \frac{\mu_j E}{\int_0^{N_j} p_j(i)^{-\frac{\rho_j}{1-\rho_j}} di},$$

where E is total expendable income, and $\int_0^{N_j} p_j(i)^{-\frac{\rho_j}{1-\rho_j}} di$ is the price index over all varieties of good j, weighted by their share in consumption from industry j. This means that revenue for the firm can be denoted

$$r_j(i) = A_j^{1-\rho_j} q_j(i)^{\rho_j}.$$

In the following I drop subscripts for industries and individual firms as it is the heterogeneity between firms within the same industry that is the main focus of this paper.

Potential entrants to the industry may enter by sinking a fixed cost of entry f_e. This permits the firm to draw its productivity level θ from a known distribution over $(0, \infty)$. The productivity level can be thought of as a total factor productivity (TFP), meaning that it works as a multiplier of the production process to determine total output. The production process is a Leontief-type technology, where a measure of different tasks all have to be undertaken in fixed amounts, here normalized to one. There are no possibilities of substitution between tasks. In the following I will use x to denote the number of times the entire production process is undertaken, which means that total output will be this production intensity, multiplied by the TFP

$$q = \theta x.$$

I assume that all tasks are undertaken outside the limits of the headquarters, but may either take place in vertically integrated plants or be bought through arm's-length outsourcing. Both sourcing options are available in the home country

101

N, or abroad in country S. In sum there are thus four different sourcing options for the firm:

Domestic Integrated (NI)	Domestic Outsourcing (NO)
Foreign Integrated, FDI (SI)	Foreign Outsourcing (SO)

Since all of the tasks have to be completed once in order to run the production process one time variable costs of production can be written

$$xc\left(s\right) = x \int_0^1 c_k^l\left(\omega\right) d\omega, \ k = \{O, I\}, \ l = \{N, S\}.$$

Each sourcing option implies some fixed costs, as well as the variable costs of production. In the case of vertically integrated plants, these fixed costs are naturally related to the investment costs of building the plant and buying the necessary machinery. For arm's-length outsourcing, the fixed costs could be searching costs to find an appropriate provider, training of workers, and testing to ensure that tasks are undertaken at an acceptable quality level. Further, I assume that any such fixed costs associated with either vertically integrated plants or outsourcing, are higher when done abroad instead of domestically. This can be explained through lack of knowledge about legal systems, local markets, language barriers, etc. In sum, I follow most of the relevant literature and assume that the fixed costs associated with sourcing strategies, f_k^l, can be ordered in the following way:[1]

$$f_I^S > f_O^S > f_I^N > f_O^N. \tag{A.1}$$

The fixed costs have to be paid for each individual task that is done within or outside the limits of the firm. This differentiates my model from Antràs and Helpman (2004), where once the fixed costs for a sourcing option are paid, the entire production process can be undertaken there. This is indeed also what happens, since their model never gives firms incentives to split the production process, as they will always choose the sourcing option with the lowest marginal costs in production, given that their production justifies the fixed costs associated with this option. There will never be incentives to split part of the production from another

---

[1] See for example Antrás and Helpman (2004), Helpman (2006), and others.

source with higher marginal costs, and in addition have to pay another fixed cost. My motivation for assuming that there are fixed costs associated with each task in the production process is both theoretically and empirically motivated. Firms that choose to outsource tasks will seldom find a provider that can deliver the best offer for all tasks, and may thus have to pay search costs for providers for each task. This is definitely the case when firms buy inputs from different countries, as is the case with the production of a typical "American" car, which is simultaneously produced in the United States (37%), Korea (30%), Japan (17.5%), Germany (7.5%), Taiwan (4%), Singapore (4%), the United Kingdom (2.5%), Ireland (1.5%), and Barbados (1.5%) (Antràs and Helpman, 2004). Similarly, Grossman and Rossi-Hansberg (2008) report that the production of the Boeing 787 involves 43 suppliers, producing at 135 sites worldwide.

Following the above assumptions the firms' profit maximizing problems can be written:

$$\max_{\{x,s\}} \left[A^{1-\rho} \left(\theta x \right)^\rho - x c \left(s \right) - f \left(x \right) \right]. \tag{2}$$

Here, the first term is the revenue term derived above, the second term expresses the variable costs, and the last term denotes the fixed cost associated with all the tasks that have to be performed. It will become clear later that the optimal sourcing strategy s is a function of x, which means that both the variable costs, and the "fixed" costs depend on the equilibrium production intensity. The first order conditions wrt x can thus be written:

$$\rho A^{1-\rho} \theta^\rho x^{*\rho-1} - \left. \frac{\partial c \left(x \right)}{\partial x} \right|_{x=x^*} x - c \left(x^* \right) - \left. \frac{\partial f \left(x \right)}{\partial x} \right|_{x=x^*} = 0,$$

where $\frac{\partial c(x)}{\partial x} = \frac{\partial c(s)}{\partial s} \frac{\partial s}{\partial x}$, which relation will become clear at a later stage. This condition yields a specific expression for neither production intensity x^*, nor output θx^*. It does, however, implicitly define these identities, and I assume that firms are able to derive their optimal production levels from this condition. It can also be shown that as long as the second-order conditions for profit maximization hold, more productive firms will produce at a higher intensity, $\frac{\partial x}{\partial \theta} > 0$, and thus also $\frac{\partial q}{\partial \theta} > 0$. The proof of this is relegated to Appendix A2. This shows that although the cost structure in the model is different from the Melitz model, the

key results come through. It also means that all the qualitative relations between production intensity x and sourcing strategies that I discuss below will also hold for productivity and sourcing strategies.

Knowing its optimal production intensity, the firm will chose a sourcing strategy in order to maximize profits. In the following I will make the simplifying assumption that the impact of an individual task on the optimal production is negligible, so that firms will disregard the output effect from switching from one sourcing option to another for a given task.

As discussed above, all tasks can potentially be undertaken either in the north or in the south. The variable costs of production in each place will naturally depend on wages in the respective countries. I assume that wages are lower in the south than in the north, $w^N > w^S$, as otherwise no firm will ever choose to have any tasks done abroad. Further, I assume that this difference is sufficiently large so that $w^N \left(f_I^S - f_O^S \right) > w^S \left(f_I^N - f_O^N \right)$ also always holds[2].

All of a measure of tasks must be undertaken in order produce final products. These tasks can be ordered according to their degree of outsourceability.[3] The intuition behind this is that tasks can be ordered according to how standardized, or 'codifiable,' they are. Some tasks are easier to define in writing, thus making it easier both to give clear instructions to workers in spite the lack of physical proximity, and lowering the possibility of moral hazard-related problems due to contract incompleteness.[4] I denote the *ad valorem* costs of producing outside the limits of the firm by $t(\omega)$. The total costs associated with each sourcing strategy

---

[2]There are many factors that could affect this inequality, few of which will be adressed here. Some that could be argued to support this condition are, for example, that a large part of the fixed costs of establishing a connection abroad is considered to be red tape and bureaucracy. This should normally affect the costs of establishing a producing unit much more than importing products from the country in question, thus increasing $f_I^S - f_O^S$ compared to $f_I^N - f_O^N$.

[3]Note that this differs from Grossman and Rossi-Hansberg (2008), who assume that tasks differ in terms of offshoreability instead of outsourceability.

[4]I will not specify the bargaining problem from incomplete contracts in this paper. Rather, I just assume that tasks can be ordered according to their degree of outsourceability. For a more specific discussion on the form of moral hazard and bargaining in the outsourcing literature, see for example Antràs (2003; 2005), Antràs and Helpman (2004; 2007), and Acemoglu et al. (2007).

for a given task ω can be written:

$$
\begin{aligned}
C_O^N &= t(\omega)w^N x + f_O^N \\
C_I^N &= w^N x + f_I^N \\
C_O^S &= t(\omega)w^S x + f_O^S \\
C_I^S &= w^S x + f_I^S
\end{aligned}
$$

The optimal sourcing strategy will thus be defined by

$$
s^*(\omega;x^*) = \arg\min\left\{c_k^l x^* + f_k^l\right\}, \tag{3}
$$

which states that for any level of outsourceability ω and optimal production intensity, the firm will choose the cheapest available sourcing option. Since there are no externalities in production between the tasks, the firm will naturally chose the sourcing strategy for each task that minimizes total costs for that specific task individually, without taking sourcing decisions for other tasks into consideration. I can thus solve for the optimal sourcing strategy for a firm with total production $x^*(\theta)$ through pairwise comparisons between all possible sourcing options for all tasks.

When production is close to zero, domestic outsourcing will always be the preferred source for all tasks, since this is the option with the lowest fixed costs. However, as production increases, sourcing options with lower variable costs may justify paying higher fixed costs. These differences in variable costs means that for each pair of sourcing options, there will be a cutoff value for $\omega = \omega'$ below which one sourcing option will be preferred over the other for all tasks $\omega \epsilon [0,\omega')$, and contrary for all tasks, ω, above this level.[5] Compare for example, the costs of outsourcing domestically (NO) with those of vertically integrating domestically (NI). With fixed costs ranked as assumed above, it is easy to see that for low values of $t(\omega)$ NO is preferable to NI. As the costs of outsourcing increase, however, this order will be reversed. Since $t(\omega)$ is monotonically increasing in ω, there will be

---

[5]The one exception to this is the comparison between domestic and foreign vertical integration, as these are both independent of outsourceability. There are still differences in marginal costs between these, but the preferred option will solely be determined by the production quantity.

a value $\omega = \omega_{NONI}$ for which the firm will be indifferent between choosing NO or NI.[6] This value is implicitly defined by:

$$t(\omega_{NONI}) w^N x + f_O^N = w^N x + f_I^N$$
$$t(\omega_{NONI}) = 1 + \frac{f_I^N - f_O^N}{w^N x}$$

Similarly, cutoff-values for ω for all pairwise rankings can be derived to be:[7]

$$t(\omega_{NONI}) = 1 + \frac{f_I^N - f_O^N}{w^N x}$$
$$t(\omega_{SONI}) = \frac{w^N}{w^S} - \frac{f_O^S - f_I^N}{w^S x}$$
$$t(\omega_{SOSI}) = 1 + \frac{f_I^S - f_O^S}{w^S x}$$
$$t(\omega_{NOSI}) = \frac{w^S}{w^N} + \frac{f_I^S - f_O^N}{w^N x}$$
$$t(\omega_{NOSO}) = \frac{f_O^S - f_O^N}{(w^N - w^S) x}$$

From these conditions one can also see that the cutoff outsourceability levels are dependent on x. This means that the optimal sourcing strategy for larger firms will be different from that of smaller firms. This is quite intuitive, as larger firms will benefit more from reductions in variable costs, as there are more units over which the fixed costs can be spread. The story here is analogous to the exporting models in Melitz (2003) and Helpman et al. (2004), where large firms can spread fixed costs over more units, and as such will opt for strategies with lower variable costs than smaller firms will. While the aforementioned papers study market access strategies for final goods, the story is similar for trade and investments in tasks, or intermediate inputs.

So far I have only determined conditions for pairwise comparisons of sourcing options. In order to get a complete mapping for which sourcing strategies a firm will opt for given production x, I need to determine simultaneous preference order-

---

[6]I denote these indifference loci of ω such that for $\omega < \omega_{ab} \rightarrow a \succ b$.

[7]The ranking of domestic integration (NI) versus foreign integration (SI) is naturally independent of ω, as ω indicate the outsourceability of a given task.

ings between all sourcing options. Since all the cutoff-values of ω are functions of x, it is possible to determine for which production levels a given cutoff ω is larger than another cutoff, i.e. when it is the case that for example $\omega_{NONI} > \omega_{NOSI}$.

$$\omega_{NONI} \geq \omega_{NOSI} \Leftrightarrow t\left(\omega_{NONI}\right) \geq t\left(\omega_{NOSI}\right) \Leftrightarrow 1 + \frac{f_I^N - f_O^N}{w^N x} \geq \frac{w^S}{w^N} + \frac{f_I^S - f_O^N}{w^N x}.$$

There exists a value for x that ensures that the above holds with equality:

$$\omega_{NONI} \geq \omega_{NOSI} \Leftrightarrow x \geq \frac{f_I^S - f_I^N}{w^N - w^S} = x_{NOSINONI}.$$

This means that for firms with optimal production intensity $x^* > x_{NOSINONI}$, the cutoff point ω_{NONI} comes for a higher value of ω than for ω_{NOSI}.[8] Similarly all pairwise comparisons of these cutoff-values can be expressed as functions of x. These sets of pairwise comparisons let me construct a complete mapping of sourcing strategies for firms. In the following I derive the determinants of the sourcing strategies for small firms as an illustrative example. A more complete presentation of this process can be found in Appendix A3.

It turns out that many of the cutoff values of x coincide, and as a result, there are five categories of firms, according to size. In each category the ranking of cutoff levels for ω defines which sourcing option will be preferred for different intervals over ω. Take for example the case of the smallest firms. For these firms it will be the case that

$$\omega_{SONI} < \omega_{NONI} < \omega_{SOSI} < \omega_{NOSI} < \omega_{NOSO}.$$

This means that for the tasks that are cheapest to outsource, $\omega \epsilon \left[0, \omega_{NISO}\right)$, the result of all pairwise comparisons of sourcing options will be

$$SO \succ NI, \ NO \succ NI, \ SO \succ SI, \ NO \succ SI, \ NO \succ SO,$$

which unambiguously shows that NO, outsourcing domestically, is preferred to any other sourcing option. Doing this for all the intervals of ω it turns out that this option is the Condorcet-winner until $\omega = \omega_{NONI}$. This comes as no surprise,

---

[8]Similarly to the ω-loci, I denote the cutoff production intensities such that $x < x_{ab} \rightarrow \omega_a > \omega_b$.

as this is exactly the cutoff-value of ω where NI becomes cheaper than NO. For this firm-size category NI stays the cheapest sourcing option for all tasks with $\omega \geq \omega_{NONI}$. The sourcing strategy for the smallest firms will thus be

$$s\left(x\right) = \begin{cases} \begin{cases} NO \text{ for } \omega\epsilon[0,\omega_{NONI}) \\ NI \text{ for } \omega\epsilon[\omega_{NONI},1] \end{cases} & \text{if} & \omega_{NONI}\epsilon\left[0,1\right] \\ NI \text{ for } \omega\epsilon\left[0,1\right] & \text{if} & \omega_{NONI} < 0 \\ NO \text{ for } \omega\epsilon\left[0,1\right] & \text{if} & \omega_{NONI} > 1 \end{cases}.$$

Note again, however, that ω_{NONI} is a function of x, meaning that the relative intensity between the two sourcing options will differ among firms within the category. Since $t\left(\omega_{NONI}\right) = 1 + \frac{f_I^N - f_O^N}{w^N x}$ is falling in x, larger firms within the category will undertake relatively more tasks in vertically integrated plants than the smaller firms will. The two last sourcing strategies in the expression above are corner solutions where firms will choose one sourcing option for all values of x^*.[9] In the rest of this section I will focus mainly on the internal solutions to simplify notation. Where relevant, corner solutions will be discussed in Appendix A1.

Repeating this exercise for all other categories of firm sizes I can define the cutoff-sizes of firms that divide firms in size-categories in increasing order as follows:

Category	Production intensity
1	$x_{NOSONOSI} > x^*$
2	$x_{NOSOSONI} > x^* > x_{NOSONOSI}$
3	$x_{SONINOSI} > x^* > x_{NOSOSONI}$
4	$x_{NISI} > x^* > x_{SONINOSI}$
5	$x^* > x_{NISI}$

Within each such category the ranking of the relevant cutoff points for ω are clearly determined. This permits a mapping of sourcing strategies along the dimensions of optimal production intensity, $x^*\left(\theta\right)$, and the cost of outsourcing, $t\left(\omega\right)$, which is depicted in Figure 1. In the figure the categories are shown, separated by the

---

[9]This is only the case for sufficiently low levels of x. This will become apparent later in this section.

vertical dotted lines.
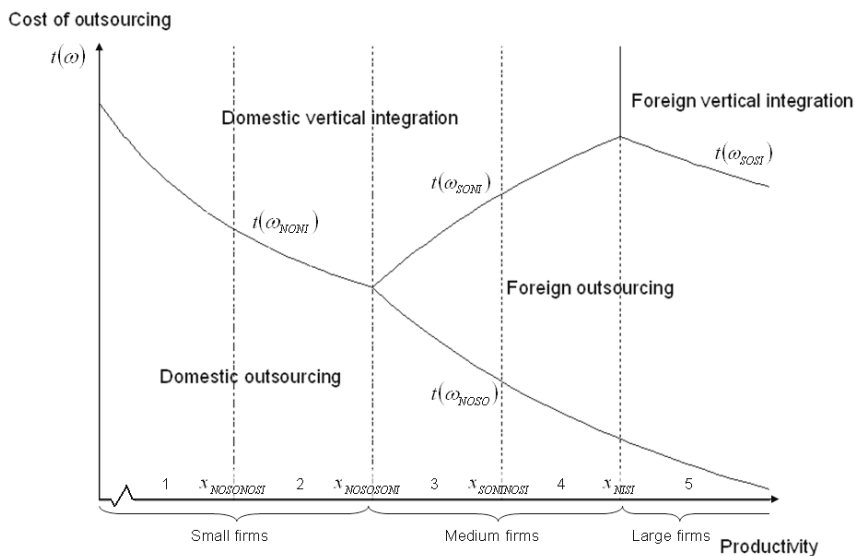
Figure 1: Sourcing strategies and productivity

It turns out that the two categories with the smallest firms (1 and 2), and the two categories with intermediate firms (3 and 4) are qualitatively the same, and these can be merged, so that I end up with three size categories of firms; small, medium, and large. The figure is thus completely determined by the following loci: $t\left(\omega_{NONI}\right)$, $t\left(\omega_{NOSO}\right)$, $t\left(\omega_{SONI}\right)$, $t\left(\omega_{SOSI}\right)$, $x_{NOSOSONI}$, and x_{NISI}. The critical size, $x_{NOSOSONI}$, determines which firms engage in international sourcing and which do not. Below this value, $t\left(\omega_{NONI}\right)$ determines the share of tasks that are undertaken in vertically integrated plants, and which are outsourced domestically. For firm sizes between $x_{NOSOSONI}$ and x_{NISI} the locus $t\left(\omega_{NOSO}\right)$ determines which tasks are outsourced domestically, which will be the tasks with $\omega < \omega_{NOSO}$. For these same firms, tasks with $\omega_{NOSO} < \omega < \omega_{SONI}$ are outsourced in the foreign country, and the tasks $\omega_{SONI} < \omega$ are undertaken in vertically integrated plants domestically. The largest firms, with production $x_{NISI} < x$, will again outsource

109

tasks $\omega < \omega_{NOSO}$ domestically and tasks $\omega_{NOSO} < \omega < \omega_{SONI}$ from the foreign country. These firms, however, will undertake the tasks $\omega_{SOSI} < \omega$ in vertically integrated plants abroad through FDI.

The above discussion means that within these categories the firms will choose their sourcing strategies from the following options:

Category	Sourcing strategies
Small	*NO, NI*
Medium	*NO, SO, NI*
Large	*NO, SO, SI*

Qualitatively there are two demarcation criteria that distinguish the categories. The first one is that small firms do not engage in international sourcing; they undertake all tasks domestically. The second one is that among the firms that do engage in international sourcing, only the largest choose to produce in integrated subsidiaries through FDI, whereas if the medium firms choose vertical integration, they will do so domestically. In other words, the model predicts that no firms will simultaneously undertake tasks in vertically integrated plants domestically and internationally. The reason for this is that I have assumed no inefficiencies in contractibility etc. for vertically integrated plants, so if a firm has a sufficiently large production and vertical integration abroad is cheaper than vertical integration domestically for one task, this will hold for all tasks. This assumption could be softened by introducing some distance costs associated with foreign production $\tau(\omega)$, as in Grossman and Rossi-Hansberg (2008), such that $c_O^S = t(\omega)\tau(\omega)w^s$ and $c_I^S = \tau(\omega)w^s$. This would require either an assumption that the ordering of tasks is such that both $t(\omega)$ and $\tau(\omega)$ are monotonically increasing, or introducing a third dimension along the $\tau(\omega)$-axis. For simplicity, and in order to be able to map sourcing strategies in a two-dimensional figure, however, I choose to stick to the assumption that $\tau = 1 \perp \omega$.

The variable costs of production for small, medium, and large firms, respec-

tively, can be expressed as $\frac{c(x^*)}{\theta}$, and is determined as follows:[10]

$$
\begin{array}{ll}
c^s(x^*) = & \int_0^{\omega_{NONI}} t(\omega) w^N d\omega + \int_{\omega_{NONI}}^1 w^N d\omega \\
c^m(x^*) = & \int_0^{\omega_{NOSO}} t(\omega) w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SONI}} t(\omega) w^S d\omega + \int_{\omega_{SONI}}^1 w^N d\omega \\
c^l(x^*) = & \int_0^{\omega_{NOSO}} t(\omega) w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SOSI}} t(\omega) w^S d\omega + \int_{\omega_{SOSI}}^1 w^S d\omega
\end{array}
$$

From the definitions of the $t(\omega)$-loci it is clear that $c^l \leq c^m \leq c^s$. (Proof is in Appendix A1.) This implies that there are two effects that make more productive firms larger than less productive firms. First, there is the direct effect that more productive firms will produce each unit of output at a lower price, and thus be more competitive and sell more units. They will also undertake the production process more times than will less productive firms. This second effect ensures that the more productive firms are able to choose cheaper sourcing strategies, which again lowers their costs per run of the production process, and thus also per unit costs, and further increases the size of more productive firms. The sum of these two effects would lead to a skewed distribution of firm sizes, with a longer right tail in the distribution than in the initial productivity distribution, even if the *ex ante* productivity distribution should be uniform.

The model predicts that the least productive firms will only engage in domestic sourcing, whereas only the most productive firms will do FDI and source from vertically integrated firms abroad. Firms with intermediate productivity levels will outsource to the foreign country, but not do FDI. This replicates the results from Antràs and Helpman (2004). However, the main contribution of this model is in the details. Whereas it reproduces the findings of previous models, it also allows firms to choose sourcing strategies involving several different sourcing options. This again generates predictions on the relative intensities of each sourcing option as functions of firm productivity. Looking at each sourcing option individually, the following predictions can be derived from the model:

Prediction 1: *Outsourcing in the north (NO) is a decreasing function of productivity.*

This is easily seen from the fact that both $\frac{\partial t(\omega_{NONI})}{\partial x} = -\frac{f_I^N - f_O^N}{w^N x^2} < 0$ and

---
[10]For $0 < \omega_{NONI},\ \omega_{NOSO},\ \omega_{SONI},\ \omega_{SOSI} < 1$.

$\frac{\partial t(\omega_{NOSO})}{\partial x} = -\frac{f_O^S - f_O^N}{(w^N - w^s)x^2} < 0$, which means that also $\frac{\partial t(\omega_{NONI})}{\partial \theta} < 0$ and $\frac{\partial t(\omega_{NOSO})}{\partial \theta} <$ 0, and NO is decreasing in θ for all $\theta \in (0, \infty)$.

Prediction 2: *Vertical integration in the north (NI) is initially increasing, and then decreasing in productivity.*

For $x^* < x_{NOSOSONI}$ vertical integration in the north is equal to $1 - \omega_{NONI}$. Since $\frac{\partial t(\omega_{NONI})}{\partial \theta} < 0$, it must be that $\frac{\partial(1 - \omega_{NONI})}{\partial \theta} > 0$, and NI is increasing in productivity in this area. For $x_{NOSOSONI} < x^* < x_{NISI}$ NI is determined by $1 - \omega_{SONI}$. $\frac{\partial t(\omega_{SONI})}{\partial x} = \frac{f_O^S - f_I^N}{w^S x^2} > 0$, which implies that $\frac{\partial t(\omega_{SONI})}{\partial \theta} > 0$, and NI is decreasing in this interval. For higher levels of productivity there will be no use of NI.

Prediction 3: *Outsourcing in the south (SO) is (initially) increasing in productivity.*

For productivities that give the interval $x_{NOSOSONI} < x^* < x_{NISI}$ the share of foreign outsourcing in a firm is equal to $\omega_{SONI} - \omega_{NOSO}$. We have already seen that $\frac{\partial t(\omega_{SONI})}{\partial \theta} > 0$ and $\frac{\partial t(\omega_{NOSO})}{\partial \theta} < 0$, implying that $\frac{\partial(\omega_{SONI} - \omega_{NOSO})}{\partial \theta} > 0$, but this would be irrelevant if $\omega_{SONI} - \omega_{NOSO} < 0$. It follows from the monotonicity of $t(\omega)$ that

$$\omega_{SONI} - \omega_{NOSO} > 0 \Leftrightarrow t(\omega_{SONI}) - t(\omega_{NOSO}) > 0.$$

Inserting the minimum x^*, $x_{NOSOSONI} = \frac{w^N(f_O^S - f_I^N) + w^S(f_I^N - f_O^N)}{w^N(w^N - w^S)}$ in this interval, yields $t(\omega_{SONI}) - t(\omega_{NOSO}) = 0$, meaning that the share of tasks that are offshored to the foreign country starts at 0 and then increases monotonically in the interval up to $x^* = x_{NISI}$. From this point on, the effect on SO is ambiguous. Both $\frac{\partial t(\omega_{NOSO})}{\partial \theta} < 0$ and $\frac{\partial t(\omega_{SOSI})}{\partial \theta} < 0$, but whether they fall at an equal pace, or whether one falls more rapidly than the other, depends on the shape of $t(\cdot)$. This implies that from this point and out the use of SO may be either increasing or decreasing in productivity.

Prediction 4: *Vertical integration in the south (SI) is increasing in productivity.*

This follows simply from the fact that for firms with productivity such that $x^* > x_{NISI}$, which are the only firms that will engage in SI, the share of tasks undertaken

in vertically integrated plants abroad will be $1-\omega_{SOSI}$. Since $\frac{\partial t(\omega_{SOSI})}{\partial x} = -\frac{f_I^S - f_O^S}{w^S x^2} < 0 \implies \frac{\partial t(\omega_{SOSI})}{\partial \theta} < 0$, it must be that $\frac{\partial(1-\omega_{SOSI})}{\partial \theta} > 0$ for these firms.

The empirical literature has often focused on offshoring, as this has been easier to obtain data for, than for each sourcing option separately. Grouping SO and SI together shows that offshoring should also be clearly increasing in both production intensity and productivity of firms. One could also group NI and SI to study whether more productive firms or firms with higher output levels would use more or less vertically integrated plants to produce intermediate inputs. Here the model has no clear predictions, as this relationship seems to be highly non-linear, but with no apparent dominating trend. If the distribution function from which the firms draw productivity is somewhat Pareto-shaped, as has been documented by empirical studies (Del Gatto et al., 2007), the small and medium sized firms would likely dominate the total effect on the use of vertical integration.[11] Taking this into consideration, the use of vertical integration as a function of production intensity or productivity should resemble the relationship between vertical integration in the north and the same independent variables.

These new and more detailed predictions, compared to previous models, will be the main focus of my empirical investigation in section 3.

2.1   Comparative statics

This mapping of sourcing strategies in a $(x^*(\theta), t(\omega))$-diagram allows for some comparative static analysis. In the following I will show what the model predicts for changes in wages in the north, w^N.

There is a growing literature on the effects of offshoring on domestic wages. The earliest arguments were that in developed, capital- and skilled labor-rich countries firms would offshore tasks that are intensive in the use of (low-skilled) labor, thus reducing demand for (low-skilled) labor, and hence also wages (see for example Feenstra and Hanson, 1996; 1999).[12] Others argued that the increased wage gap

---

[11]A sneak peek at the data reveals that around 85% of the firms in the sample are categorized as small or medium when using the demarcation criteria from the model, indicating that these firms should indeed dominate the overall effect.

[12]Offshoring is here defined as undertaking tasks abroad, and thus includes both outsourcing and vertical FDI in the south in my model.

between high- and low-skilled labor was driven by competition from low-wage countries, as well as technological advances (Leamer, 1996; Autor et al. 1998). Later contributions have shown how such offshoring may actually increase wages, through either productivity gains (Girma and Görg, 2004; Grossman and Rossi-Hansberg, 2008; and others), or through increased bargaining power for labor unions (Lommerud et al., 2009). The net effect on wages from offshoring thus depends heavily on which factors dominate. In this paper I will not go into this debate, but only show how a change in wages in the north will affect sourcing strategies for firms. This will then enable me to comment on possible feedback-effects if offshoring indeed changes the home country wages. If changes in offshoring lead to changing wages, this should over time affect the firms' optimal sourcing strategies. If it is indeed the case that increased offshoring will lead to increased wages, whereas increasing wages leads to increased offshoring, then a small initial change could potentially lead to large total effects.

As discussed above, the choice of sourcing strategies as a function of firm size is completely determined by $t(\omega_{NONI})$, $t(\omega_{NOSO})$, $t(\omega_{SONI})$, $t(\omega_{SOSI})$, $x_{NOSOSONI}$, and x_{NISI}. The effects of a change in northern wages on these are

$$
\begin{aligned}
\frac{\partial t(\omega_{NONI})}{\partial w^N} &= -\frac{f_I^N - f_O^N}{(w^N)^2 x} < 0 \\
\frac{\partial t(\omega_{NOSO})}{\partial w^N} &= -\frac{f_O^S - f_O^N}{(w^N - w^S)^2 x} < 0 \\
\frac{\partial t(\omega_{SONI})}{\partial w^N} &= \frac{1}{w^S} > 0 \\
\frac{\partial t(\omega_{SOSI})}{\partial w^N} &= 0 \\
\frac{\partial x_{NOSOSONI}}{\partial w^N} &= \frac{(w^N - w^S)^2 (f_I^N - f_O^N) - (w^N)^2 (f_O^S - f_O^N)}{(w^N - w^S)^2 (w^N)^2} < 0 \\
\frac{\partial x_{NISI}}{\partial w^N} &= -\frac{f_I^S - f_I^N}{(w^N - w^S)^2} < 0.
\end{aligned}
$$

The changes are depicted graphically in Figure 2.

This clearly shows that an increase in wages in the north will lead to an un-ambiguous increase in foreign sourcing if the distribution of firm sizes remains

constant. This is not at all surprising, but it illustrates an important point: if it is the case that offshoring indeed increases wage levels at home, this should lead to a circular effect where firms get increased incentives to seek foreign sourcing options. If this self-reinforcing effect is strong enough, a small increase in northern wages could ultimately lead to all tasks being offshored. This does not sound like a likely outcome, however, and could possibly indicate that if offshoring leads to increased wages, this may only be for certain degrees of offshoring. This is exactly what is predicted in Kohler and Wrona (2010).
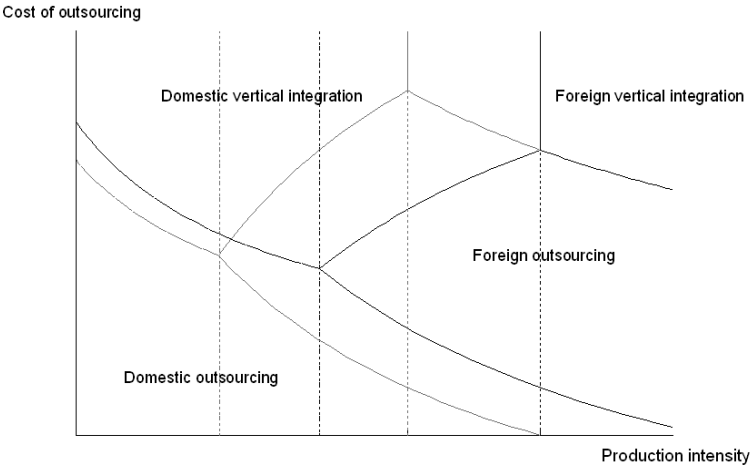
Figure 2: Increase in the wages in the north

3   Data and empirical strategy

In this section of the paper I will test some of the predictions generated in the theoretical model. Due to the nature of the data, not all predictions can be tested, and the ones I do test should not be interpreted as causal mechanisms. It is however a necessary, although not sufficient, condition for the validity of the theoretical model that the testable predictions be reflected in the data. Should

this be the case, then it would be an indication that the model cannot readily be seen as falsified by the empirical testing, and would as such indicate some support of the model.

The relationship between productivity and sourcing strategies has been tested in the empirical literature before. Nunn and Trefler (2008) test the predictions from Antràs (2003) and Antràs and Helpman (2004) on intra-firm trade, using data on U.S. firms' within-industry imports from foreign affiliates, and find some support for their predictions that intra-firm trade should be higher among firms that are intensive in their use of headquarter services. The same is also the case for firms that are skill- and capital-intensive. Corcos et al. (2008) refine this research using French firm-level data, and include TFP measures as an explanatory variable. Their study shows that also more productive firms are more likely to source inputs through intra-firm international imports. Both of these studies rely on import data to construct their sourcing variables, which may be an imprecise measure of sourcing, as it is hard to distinguish which imports are actually intermediate inputs in production. One study that avoids this problem is Tomiura (2007). Using survey data on Japanese firms, he has information directly from firms on "whether they contract out manufacturing or processing tasks to other firms overseas," a direct dummy for SO in my model: foreign outsourcing. His findings indicate that firms that outsource to foreign countries are less productive than firms that do FDI, but more productive than domestic firms. These findings compare directly to the "predicted firm sizes" from my model, and also show the same productivity ordering that my model predicts. Federico (2010) uses Italian survey data that includes information on all four sourcing options that I use in my model. Measuring the productivity premium for firms that use the different sourcing options, he concludes that firms that choose foreign integration are the most productive, and the ones choosing domestic outsourcing are the least productive. However, contrary to the predictions in my model, he claims that domestic integration is chosen by medium-high productivity firms, whereas medium-low productivity firms prefer foreign outsourcing. Kohler and Smolka (2009) use the same Spanish survey data as I use in this paper to study the connection between firm productivity and sourcing behavior. As I will come back to later in this chapter, these data contain information about the intensity with which all four sourcing

options are used, as well as data that distinguishes firms that are headquarters from the ones that themselves are subsidiaries of foreign firms. Interestingly, their study concludes that the unrestricted sample, when "non-headquarter" firms are included, gives results that are in line with those in Federico (2010). However, when studying only firms that are true headquarters, and as such assumed to have complete discretion over sourcing strategies, the productivity ordering between firms that outsource to the foreign country and those that integrate domestically is reversed and in accordance with the predictions from my model.

The data I use are acquired from the annual business survey from Fundación SEPI.[13] The survey covers about 2,000 firms with more than 10 employees annually, and report data for individual firms, and not corporate groups. All firms with more than 200 employees are invited to participate, whereas a random sample of about 5% of firms with 10-200 employees are asked.[14] This means that large firms are somewhat overrepresented in the sample, something I try to control for when possible and necessary.

The main advantages of these data, compared to other firm-level data, is the detail of the information that it contains, both of the main variables of interest and of some important control variables. As discussed above, many empirical investigations of offshoring use imports in the same SIC category as a proxy for offshoring. In the data used here the firms answer direct questions about the percentage of intermediate inputs they buy from other, related or unrelated firms, domestically and abroad. These four variables thus correspond directly to the four sourcing options described in my theoretical model. Unfortunately these questions have only been included in the survey for the years 2006-2008, and with little variation in sourcing strategies over such a short time period, the data do not permit the use of panel data techniques in order to perform better tests for causality, as well as selection- and learning-effects of different sourcing strategies.

---

[13]The survey "Encuesta sobre estrategias empresariales" is conducted by Fundación Sociedad Estatal de Participaciones Industriales. See http://www.funep.es/esee/en/einfo_que_es.asp for more information about the foundation and the survey.

[14]Starting from the initial sample from 1990, SEPI has included all newly incorporated firms with more than 200 employees, and a randomly selected sample of about 5% of the newly incorporated firms with 10-200 employees. As such the large firms are overrepresented in the data, but within each group careful measures are taken to ensure the representativeness of the data. Average response rate for 1990-2008 is an impressive 91.97%.

With the continued collection of these variables though, such investigations will become possible in the near future. The data do however permit the use of lagged values for productivity, as measured productivity could potentially be affected by sourcing behavior. The results do not seem to change with either one- or two-year lags, and only the results from running the unlagged variables are reported here. In total my data set is a relatively balanced panel that includes information on 4,629 firms from 1999 to 2008. As I only have information on my main variables of interest since 2006, the results reported in this section are from cross-sectional analysis for individual years between 2006-2008. The results reported here are from 2007, but all three years show similar results. The longer time-series are used in estimating total factor productivity for the individual firms. For a more thorough discussion of the data, see Kohler and Smolka (2009).

Making dummies for the use of each sourcing option, I get four not mutually exclusive categories. The summary statistics show (with large firms excluded in parentheses) that 4.2% (4.8%) of the firms in the survey buy no intermediate inputs through any of the sourcing options, 92.4% (93.0%) outsource domestically, 43.1% (34.6%) outsource from other countries, and 15.4% (8.4%) and 10.4% (3.7%) buy inputs from vertically integrated plants in Spain and other countries, respectively. Around 47.8% (37.4%) of the firms in the sample source from more than one of the four options every year. These values show that almost all firms buy some intermediate inputs from other firms, and also shows models that predict that firms will source all their input from the same provider are in discordance with the empirical observations.

I will let the empirical part of this study follow the theoretical model as closely as possible. The descriptive statistics in Table 1 are thus divided by predicted size according to the size definitions from the model. If a firm only buys intermediate inputs from domestic sources, it is classified as small. Firms that do buy inputs through arm's-length dealings with independent firms abroad (SO), but do not undertake FDI and produce in vertically integrated plants abroad, are categorized as medium, and finally firms that have vertically integrated plants abroad are classified as large. The summary statistics for each of these categories are shown

below.

Predicted size	Output	Productivity	Export status	Similar product	Capital intensity
Small	39.31	3.84	0.45	0.03	18.02
Medium	71.20	3.91	0.81	0.08	24.14
Large	339.72	4.16	0.97	0.26	33.65
Total	81.59	3.90	0.62	0.07	21.75

Predicted size	R&D intensity	Capacity utilization	Employees	N
Small	0.31	83.20	127	1,006
Medium	0.60	82.78	222	691
Large	1.86	82.47	691	209
Total	0.57	83.08	218	2,006

Table 1: Descriptive statistics

Table 1 summarizes the main variables I will focus on in this section, for each size group and for the sample average. The first thing we notice is that OUT-PUT, defined as the sum of sales plus the variation in inventory, in millions of Euros, corresponds well to the predicted size categories from the model, with the 'medium' firms producing a little less than the average output in the sample, while the 'small' firms produce around half, and the 'large' firms produce more than four times the average output. As predicted by the model, the PRODUCTIVITY is also increasing in the size categories. This variable is generated with the Olley and Pakes (1996) method, which has become the favored method of estimating total factor productivity in the economic literature, thanks to its correction for endogeneity issues related to productivity shocks and selection in the exit decisions. Among the important control variables when studying sourcing strategies and productivity I have also included EXPORT STATUS, which is a dummy for whether the firm is an exporter or not. This variable shows a very similar pattern to that of productivity, which is the familiar "Melitz-result" (2003), that more productive

119

firms will become exporters whereas less productive firms will produce only for the domestic market. The next variable, SIMILAR PRODUCT, is a dummy for whether the main foreign affiliate of the firm (if it has one) produces a similar product, or in other words if the group has undertaken horizontal FDI. The rationale for including this variable is that if a firm has several production plants producing the same product, the total production will be larger than that reported only by the domestic firm, and according to the forces at work in the theoretical model, it will thus have more units over which to spread the fixed costs, and stronger incentives to choose a sourcing option with lower marginal costs. In total only 6% of the firms produce a similar product to their main affiliate, but in the large firm category, 25% of the firms do, indicating that this is a prominent feature among the largest firms. The next two variables show that the larger firms both have higher CAPITAL INTENSITY and R&D INTENSITY than the smaller firms. I have also included the firms' degree of CAPACITY UTILIZATION. This does not seem to vary systematically across the different size categories, but could have some explanatory power over sourcing options. Specifically one could expect that firms would increase their domestic outsourcing when capacity utilization is very high, as this could be the cheapest short-term solution to cover for example a temporary demand shock. Finally, Table 1 shows that the predicted sizes correspond to the actual sizes when measured in the number of EMPLOYEES.

3.1 Results

The fit of predicted size categories according to the model, and real size categories, measured in total production, is reported in Table 2.

<div style="text-align:center">Actual company size</div>

P		1	2	3	4	5	Total
r	Small	346	275	229	149	107	1 106
e		31.28	24.86	20.71	13.47	9.67	100.00
d							
i	Medium	54	122	157	194	164	691
c		7.81	17.66	22.72	28.08	23.37	100.00
t							
e							
d	Large	0	3	15	60	131	209
		0.00	1.44	7.18	28.71	62.68	100.00
s							
i	Total	400	400	401	403	402	2 006
z		19.94	19.94	19.99	20.09	20.04	100.00
e							

<div style="text-align:center">Table 2: Predicted size versus actual firm size</div>

The actual sizes are divided into quintiles when firms are sorted according to total value of production. The second row for each predicted size category shows the percentage of the firms in this category that fits in each of the actual size categories. Reading these shares horizontally, we see that the firms that are predicted to be 'small' according to the theoretical model are predominantly among the small firms when sorted by actual production as well. In fact, 67% of the firms that are predicted to be small according to the model fall into one of the two groups of firms with the lowest actual output. The 'medium' category is quite evenly spread out over all levels of total output, whereas the 'large' category is predominant among the firms with large productions, with 74% of the firms in this group falling into the two top categories of actual output. A simple, univariate regression shows that more than 23% of the variation in the actual company size distribution can be explained by the predicted size categories. This gives a first indication that

the theoretical model may capture some of the underlying mechanisms in the real world: that the smallest, least productive firms source domestically, whereas only the most productive firms choose FDI. To test the model more stringently, I thus turn to the more specific predictions of the model, which to the best of my knowledge, has not been predicted by theoretical contributions in this literature before.

Predictions 1-4 in section 2 stated that the intensity of each individual sourcing option would be functions of the productivity of firms. More specifically, my model predicted that outsourcing in the north would be decreasing in firms' productivity and integration in the north would be strongly concave, whereas both outsourcing and integration in the south would be increasing in the firms' productivity. I test the mentioned predictions by regressing the share of intermediate inputs that the firms acquire through each sourcing option on productivity measures estimated by the Olley-Pakes method (Olley and Pakes, 1996) and the control variables discussed above, separately. These shares are bounded from below at 0% for all sourcing options, and also from above at 100% in the case of domestic and foreign outsourcing. I therefore use the Tobit model and censored regression to adjust for this. The results are shown in Table 3. It may be that establishing partnerships for outsourcing, or building integrated plants domestically or abroad, may take some time. Further, it has also been argued that firms that are controlled by foreigners will lack some domestic knowledge, while having superior knowledge about their own country, and that this will affect their sourcing strategies. I thus follow Kohler and Smolka (2009) and restrict the sample to firms that have existed for five years or more, and also exclude foreign firms, by which I mean firms where more than 33% of the shares are controlled by foreign shareholders.[15] Since the predicted effects of productivity on sourcing behavior are likely to be non-linear for most forms of $t(\omega)$ I include the squared term of productivity to allow for some non-linearity of a second-degree kind. As before, the results are reported for 2007, since these results are closest to the average results over the three years, for which I have the relevant data, but the results for 2006 and 2008 are again practically

---

[15]Among these "foreign firms," 96.51% are 50% or more controlled by foreign shareholders, and 83.95% have at least 98% of their shares controlled by foreign shareholders. 77.11% are 100% foreign-owned. Not very surprisingly, the results do not change much when I use 50% or 100% foreign shareholding as cutoff-levels for defining foreign firms.

identical. Regressions (1)-(4) show the tests of the predicted relationship between PRODUCTIVITY and the respective sourcing options. The relationships do seem to be quite non-linear. All coefficients that are predicted by predictions 1-4 are of the expected sign, and most of them are statistically significant. The effect of PRODUCTIVITY on foreign vertical integration seems somewhat less statistically significant than the others, but this might come from the fact that only about 200 of the firms in the sample source any intermediate input through this channel. The economic importance of these effects does also seem to be important. From the average level of productivity, a 10% productivity increase would predict a -1.75 percentage points change in domestic outsourcing, whereas domestic vertical integration, foreign outsourcing and foreign vertical integration would be expected to increase by 4.12, 4.77, and 5.26 percentage points, respectively. The main conclusions from studying regressions (1)-(4) however, is that none of the predictions 1-4 can be rejected. The results also show that sourcing behavior varies quite a lot between firms.

Compared to the model, there are some differences in the data. One example is that as many as 40% of the firms in this sample report sourcing percentages that add up to less than 100%. This is not a big problem, as in the real world some firms would produce some intermediate inputs themselves, in contrast to only producing headquarter services as assumed in the model. In-house production should best be thought of as a form a vertical integration, and the closest equivalent to the model would thus be to proxy $NI = 100 - (NO + SO + SI)$. Doing this does not change the results significantly, and I conclude that this deviation between the model and the data is not a problem. Another issue that might be of more concern, is that the dependent variables are connected through the substitutability between them. This could mean that a seemingly unrelated regressions approach would yield more efficient coefficients; however, the bias from a linear estimation on the truncated data would be of more concern, and I thus prefer to run with the tobit analysis. The tobit-model on the other hand hinges on the rather strict assumption that the error terms are normally distributed. A test for normality after tobit estimation derived by Skeels and Vella (1999) is thus run on regressions (2) and (4), and null

hypothesis of normality cannot be rejected in any of these cases.[16]

	(1) NO	(2) NI	(3) SO	(4) SI
Productivity	-17.54**	39.28**	66.50**	102.2*
	(-2.44)	(2.33)	(2.35)	(1.94)
Productivity squared	1.688**	-3.710**	-7.028**	-11.51*
	(2.26)	(-1.99)	(-2.34)	(-1.88)
Export status	-13.66***	14.16***	109.7***	38.73***
	(-5.10)	(3.03)	(9.94)	(3.25)
Similar products	-11.81**	16.17**	24.38	57.43***
	(-2.06)	(2.11)	(1.27)	(4.96)
Capital intensity	-0.0880**	0.159***	0.129	0.0197
	(-2.01)	(2.80)	(0.82)	(0.20)
R&D intensity	-0.235	2.975***	5.944**	0.411
	(-0.26)	(2.63)	(1.96)	(0.27)
Capacity utilization	0.222***	-0.279*	-0.100	0.207
	(2.60)	(-1.94)	(-0.32)	(0.68)
Pseudo R^2	0.00873	0.0652	0.0390	0.118
N	1663	1663	1663	1663

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Productivity and sourcing behavior

For tractability and simplicity, none of the control variables in Table 3 were incorporated in the theoretical model. Nonetheless, the analysis shows that some of these variables have important and interesting effects on the firms' sourcing strategies. Equation (2) showed that output is a convexly increasing function of productivity. The relationship between productivity and EXPORT STATUS has been widely argued in the international trade literature (see for example Clerides et al., 1998; Bernard and Jensen, 1999; Bernard et al., 2003; and many others), and strong support has been found for a selection effect that more productive firms self-select into exporting. More mixed results are found for the learning effect of exporting, i.e. whether exporting firms increase their productivity after becoming exporters. Table 3 shows that *ceteris paribus,* exporters use domestic outsourcing

---

[16]The tobcm test cannot be run on regressions (1) and (3) as this test can only be used on left-censored regressions with zero as the censoring point.

options less intensively than non-exporters do, and use all other sourcing options more intensively. These results may indicate that the costs of sourcing abroad may be lower if the firm already has some sales network operating abroad. Caution has to be taken, however, as this effect may just as well be in the other direction: that foreign outsourcing makes establishing an export network cheaper, and thus makes exporting more likely for firms that are already sourcing abroad.[17] Further, it may be the sourcing from sources with lower variable costs that contribute to the export status through lowering unit costs, and thus the competitiveness of the good; however, using the lagged variable of export status does not change the results.[18]

My proxy for horizontal FDI (SIMILAR PRODUCT) yields the expected results, as it is positive and statistically significant for the use of vertical integration, especially foreign vertical integration, which I interpret as an indication that these corporate groups have established plants producing intermediate inputs, which it delivers to several final producers around the world, and as such captures a volume effect that is not captured in the theoretical model, nor by the productivity effect in the empirical analysis. The effect on domestic outsourcing is negative and statistically significant, and for foreign outsourcing it is positive, but not significantly different from zero. This could also be coherent with the hypothesis that international conglomerates centralize their sourcing, both in vertically integrated plants and the inputs they contract at arm's-length. The effects of CAPITAL INTENSITY and R&D INTENSITY on firms' sourcing behavior is a bit harder to interpret, but there seem to be some effects going on, and they should be used to control for whatever mechanisms this might be. In the theoretical model, firms will always invest for, and produce the optimal quantity of goods, and therefore

---

[17] A random effects probit regression shows that firms with higher use of foreign outsourcing (4.29) and foreign vertical integration (3.27) are significantly more likely to be exporters, while domestic outsourcing (-3.20) and vertical integration (-3.05) reduces this likelihood (z-values are shown in parentheses). A fixed effect logit, however, generates coefficients of the same signs, but these are not statistically significant. This is not very surprising, since I only have information for these variables for the period 2006-08, and the low variance in export status over this period reduces the sample to a mere 442 observations.

[18] Other concerns could be that there is collinearity between productivity and export status, as argued theoretically by Melitz (2003) and Helpman et al. (2004) and others. However, the correlation between the two is a mere 0.04, so collinearity should not be a problem in the econometric analysis.

always produce at full capacity. The uncertainties of the real world are naturally not compatible with this, as shown in the coefficient for the variable CAPACITY UTILIZATION. This variable has a positive and significant effect on the use of domestic outsourcing (NO), and a negative effect on the use of domestic vertical integration (NI). My interpretation of this is that outsourcing domestically is the sourcing option with the lowest fixed cost, and is therefore a natural choice of sourcing in order to cover a temporal increase in demand. The negative effect on vertical integration could reflect its lack of flexibility, and thus the counter-cyclical intensity in the use of this sourcing option.

Table 4 shows the results for using two alternative measures of productivity: output per worker and value added per worker. These results are shown in Table 4.

Regressions (5)-(9) use output per worker as the productivity measure, whereas regressions (10)-(13) use value added per worker. The results are qualitatively very similar to the results from regressions (1)-(4), using the Olley-Pakes productivity measures as the main explanatory variable. The coefficients for the alternative productivity measures are also statistically significant in most cases. As the productivity estimates in Table 4 are not very sophisticated, while the Olley-Pakes productivity estimates are arguably the best estimates methodologically, and since they are also the most commonly used in the literature, I will use these estimates as my productivity variable in the following.

	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		Output per worker				Value added per worker		
	NO	NI	SO	SI	NO	NI	SO	SI
Productivity	-21.89***	42.75***	134.5***	32.75**	-0.311*	1.130***	2.074***	0.676
	(-4.55)	(6.26)	(7.46)	(2.43)	(-1.69)	(3.82)	(3.15)	(1.29)
Productivity squared	4.148***	-6.459***	-20.51***	-4.701*	0.00122	-0.00640***	-0.00695**	-0.00414
	(3.95)	(-4.80)	(-5.44)	(-1.77)	(1.31)	(-3.19)	(-2.21)	(-1.31)
Export status	-11.69***	10.27**	98.51***	36.98***	-14.28***	13.06***	109.8***	39.40***
	(-4.32)	(2.20)	(9.13)	(3.12)	(-5.31)	(2.78)	(9.89)	(3.31)
Similar products	-11.23*	16.11**	18.54	57.52***	-11.70**	15.65**	21.94	56.93***
	(-1.96)	(2.13)	(0.99)	(4.99)	(-2.03)	(2.03)	(1.13)	(4.91)
Capital intensity	-0.0464	0.0459	-0.344**	-0.0929	-0.0631	0.0808	-0.0803	-0.0258
	(-0.97)	(0.76)	(-1.99)	(-0.85)	(-1.30)	(1.26)	(-0.46)	(-0.23)
R&D intensity	-0.182	2.288**	3.818	-0.157	-0.224	3.599***	4.597	1.423
	(-0.20)	(2.03)	(1.28)	(-0.10)	(-0.23)	(2.83)	(1.45)	(0.69)
Capacity utilization	0.219***	-0.298**	-0.285	0.197	0.209**	-0.281*	-0.216	0.208
	(2.59)	(-2.10)	(-0.92)	(0.65)	(2.44)	(-1.95)	(-0.68)	(0.68)
Pseudo R^2	0.00982	0.0746	0.0483	0.120	0.00843	0.0664	0.0403	0.115
N	1676	1676	1676	1676	1668	1668	1668	1668

t statistics in parentheses
$^* p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01$

Table 4: Alternative productivity measures

The above testing of Predictions 1-4 indicates that the model may very well describe an important mechanism for firms' use of different sourcing options. In the motivation for this paper, however, I argued for the relevance of sourcing strategies, i.e. the combination of all the sourcing options the firm decides to use. The theoretical model generates clear predictions on which such strategies are compatible with the model, and also in which order they can be ranked by firm productivity. In the model, there can, depending on the form of $t(\omega)$, be 10 different sourcing strategies, ranked by productivity of the firms that will use them, with the most productive firms choosing SI and the least productive choosing NO:[19]

Sourcing strategy	Rank	Percentage of firms
SI	1	0.2%
$SO + SI$	2	0.2%
$NO + SO + SI$	3	4.9%
SO	4	1.6%
$NI + SO$	5	0.1%
$NO + SO$	5	26.9%
$NO + NI + SO$	7	5.8%
NI	8	0.8%
$NO + NI$	9	4.7%
NO	10	45.3%

A first examination of these possible strategies reveal that they cover only 10 of the 15 strategies that could possibly exist in the real world. However, these 10 strategies account for 90.7% of the strategies that the firms in the sample actually choose. Firms with more than 200 employees are overrepresented in the sample, and splitting the sample in large and small firms along this demarcation criteria, I find that the 10 sourcing strategies permitted by the model account for 84.0% of the strategies used by large firms, and 93.3% of the strategies chosen by the small firms. The most important strategy that is excluded by the theoretical model,

---

[19]Not all these strategies can potentially occur for a specific function form of $t(\omega)$, but all strategies are possible under one form or another of $t(\omega)$. The ranking along firm productivity is unambiguous, with the exception of the ranking of NISO and NOSO, as they will never both occur for a specific function form of $t(\omega)$. These two strategies are thus ranked as equal on the productivity scale.

is NONISOSI, sourcing from all the possible sourcing options. This strategy is used by 3.1% of the firms in the data. In the subsamples, this value is 9.1% for the large firms, and 0.8% among the small firms. I thus argue that the sourcing options that are not rationally used, according to the theoretical model, are not commonly used among firms in my data sample either, with the exception of some of the largest firms that use all four sourcing options. Further, the model also predicts a ranking of these most-used sourcing strategies along the productivity axis. Based on this theoretical ranking, an ordered logit regression on sourcing strategy coded inversely to the one shown above should show a positive coefficient for the productivity measures. The result of such a regression is shown in Table 5.

	(14) Ordered logit Sourcing strategies	(15) Tobit Offshoring	(16) Tobit Vertical integration
Productivity	1.077***	73.64**	49.85***
	(2.64)	(2.47)	(2.64)
Productivity squared	-0.115**	-7.709**	-4.905**
	(-2.53)	(-2.44)	(-2.32)
Export status	1.407***	115.28***	18.53***
	(11.82)	(9.88)	(3.73)
Similar product	0.476*	43.06**	34.53***
	(1.93)	(2.10)	(4.46)
Capital intensity	0.0010	0.111	0.151**
	(0.55)	(0.67)	(2.49)
R&D intensity	0.089*	8.892**	2.583**
	(1.83)	(2.45)	(2.15)
Capacity utilization	-0.0015	-0.108	-0.248
	(-0.41)	(-0.32)	(-1.63)
Pseudo R^2	0.0828	0.0416	0.0640
N	1555	1663	1663

t statistics in parentheses
$* p < 0.10$, $** p < 0.05$, $*** p < 0.01$

Table 5: Sourcing strategies

Indeed, the coefficients for productivity in regression (14) show the expected sign, as predicted by the theoretical model. This indicates that firms sort into the

different sourcing strategies based on their productivity.

One question in the debates around the "disintegration of production" has been the degree of offshoring, the moving of jobs out of the country. There are many reasons this topic raises debates, as it is easy to understand that such movement of tasks to foreign countries will affect local labor markets, although as mentioned above, the literature is ambiguous about how it will affect wage levels. In regression (15) I group together the share of intermediate inputs that the firms buy from both related and unrelated firms outside of Spain. Not surprisingly, and in accordance with previous studies, these regressions show that larger, more productive firms offshore more tasks and import more intermediate inputs from abroad than less productive, smaller firms. Again, the effect seems to be concave. In the next regression, (16), I run a similar regression, but on the sum of the shares of intermediate inputs bought from vertically integrated plants both in Spain and abroad. Also this relationship seems to be positive, but concave; larger, more productive firms use more inputs from vertically integrated plants. Also the signs of the control variables seem to be consistent with the above interpretations from the regressions on individual sourcing options.

4   Conclusion

In this paper I have developed a theoretical model for firms' sourcing decisions that incorporates dimensions from the productivity-driven sorting mechanism from Antràs and Helpman (2004), as well as the task-trading from Grossman and Rossi-Hansberg (2008). Combining this with some fixed costs associated with each task I get a framework where firms sort into sourcing strategies, as in the simultaneous use of several sourcing options, both domestically and abroad. To the best of my knowledge, this is the first sourcing model to generate such rich testable predictions on the "disintegration of industrial production." The model also predicts the intensity in the use of each individual sourcing option as functions of output levels and productivity, as well as the degree of both offshoring and the existence of multi-plant firms.

Testing these predictions on firm-level data from Spanish firms, a first observation is that around half of the firms in the sample use more than one sourcing

option, indicating the empirical relevance of models that allow for sourcing strategies that use a combination of sourcing options. In the more detailed testing of the use of sourcing options, and sourcing strategies, I find quite strong support for the model's predictions. A couple of caveats concerning the data should however be pointed out. One is that large firms are overrepresented in the sample, and since firm size is an important factor in predicting sourcing behavior, this might affect the magnitude of the results. Although I have controlled for this as much as possible, and find that the main results hold also for the sub-sample of only firms with less than 200 employees, it would be preferable to estimate the effects on a more representative sample. Secondly, data on sourcing behavior have only been recorded since 2006, and since these are long-term decisions for firms, there is very little variance over time to permit the use of econometric panel data techniques, and causality is thus difficult to establish. This implies that the relation between higher estimated productivity and the use of foreign sourcing and vertical integration, may run both ways, and even be a self-reinforcing effect. Literature on productivity premiums among exporters discusses selection into and learning from exporting, effects that would be highly relevant to estimate for sourcing strategies as well, although at this point in time the data does not permit this line of investigation.

Finally, the empirical testing shows that there is a strong relation between exporting and sourcing strategies. This is not surprising, as in my model the mechanisms that drive the use of sourcing options with higher fixed costs, but lower variable costs, are exactly the same mechanisms that cause firms to export in models like that of Melitz (2003). This means that the empirical relation could possibly just be a spurious one, although I find that hard to believe. There are likely to be synergies between international production and international distribution, and also between lowering marginal costs of production and becoming more competitive in the international market. Integrating my model into a Helpman et al. (2004) type of sorting mechanism for exporting and FDI would thus be an interesting next step in this research.

131

References

Acemoglu, D., Antràs, P., & Helpman, E. 2007. Contracts and Technology Adoption. *American Economic Review*, **97**(3), 916–943.

Antràs, P. 2003. Firms, Contracts, and Trade Structure. *Quarterly Journal of Economics*, **118**(4), 1375–1418.

Antràs, P. 2005. Incomplete Contracts and the Product Cycle. *American Economic Review*, **95**(4), 1054–1073.

Antràs, P., & Helpman, E. 2004. Global Sourcing. *Journal of Political Economy*, **112**(3), 552–580.

Antràs, P., & Helpman, E. 2008. Contractual Frictions and Global Sourcing. *In:* Helpman, E., Marin, D., & Verdier, T. (eds), *The Organization of Firms in a Global Economy.* Harvard University Press: Cambridge, MA.

Autor, D., Katz, L., & Krueger, A. 1998. Computing Inequality: Have Computers Changed the Labor Market. *Quarterly Journal of Economics*, **113**(4), 1169–1213.

Bernard, A., & Jensen, J. 1999. Exceptional Exporter Performance: Cause, Effect, or both? *Journal of International Economics*, **47**(1), 1–25.

Bernard, A., Eaton, J., Jensen, J., & Kortum, S. 2003. Plants and Productivity in International Trade. *American Economic Review*, **93**(4), 1268–1290.

Clerides, S., Lach, S., & Tybout, J. 1998. Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco. *Quarterly Journal of Economics*, **113**(3), 903–947.

Del Gatto, M., Mion, G., & Ottaviano, G. 2007. *Trade Integration, Firm Selection and the Costs of Non-Europe.* CRENoS Discussion Papers 217. Universita' degli Studi di Cagliari.

Dixit, A., & Stiglitz, J. 1977. Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, **67**(3), 297–308.

Federico, S. 2010. Outsourcing versus Integration at Home or Abroad and Firm Heterogeneity. *Empirica*, **37**(1), 47–63.

Feenstra, R. 1998. Integration of Trade and Disintegration of Production in the Global Economy. *The Journal of Economic Perspectives*, **12**(4), 31–50.

Feenstra, R., & Hanson, G. 1996. Globalization, Outsourcing, and Wage Inequality. *American Economic Review*, **86**(2), 240–245.

Feenstra, R., & Hanson, G. 1999. The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979-1990. *Quarterly Journal of Economics*, **114**(3), 907–940.

Girma, S., & Görg, H. 2004. Outsourcing, Foreign Ownership, and Productivity: Evidence from UK Establishment-level Data. *Review of International Economics*, **12**(5), 817–832.

Grossman, G., & Helpman, E. 2002. Integration versus Outsourcing in Industry Equilibrium. *Quarterly Journal of Economics*, **117**(1), 85–120.

Grossman, G., & Rossi-Hansberg, E. 2008. Trading Tasks: A Simple Theory of Offshoring. *American Economic Review*, **98**(5), 1978–1997.

Helpman, E. 2006. Trade, FDI, and the Organization of Firms. *Journal of Economic Literature*, **44**(3), 589–630.

Helpman, E., Melitz, M., & Yeaple, S. 2004. Export Versus FDI with Heterogenous Firms. *American Economic Review*, **94**(1), 300–316.

Hummels, D., Ishii, J., & Yi, K. 2001. The Nature and Growth of Vertical Specialization in World Trade. *Journal of International Economics*, **54**, 75–96.

Kohler, W., & Smolka, M. 2009. *Global Sourcing Decisions and Firm Productivity: Evidence from Spain*. CESifo Working Paper Series 2903. CESifo.

Kohler, W., & Wrona, J. 2010. *Offshoring Tasks, Yet Creating Jobs?* CESifo Working Paper Series 3019. CESifo.

Leamer, E. 1996. Wage Inequality from International Competition and Technological Change: Theory and Country Experience. *American Economic Review*, **86**(2), 309–314.

Lommerud, K., Meland, F., & Straume, O. 2009. Can Deunionization Lead to International Outsourcing? *Journal of International Economics*, **77**(1), 109–119.

Melitz, M. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, **71**(6), 1695–1725.

Nunn, N., & Trefler, D. 2008. The Boundaries of the Multinational Firm: An Empirical Analysis. *In:* Helpman, E., Marin, D., & Verdier, T. (eds), *The Organization of Firms in a Global Economy.* Harvard University Press: Cambridge, MA.

Olley, S., & Pakes, A. 1988. The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, **64**(6), 1263–1298.

Skeels, C., & Vella, F. 1999. A Monte Carlo Investigation of the Sampling Behavior of Conditional Moment Tests in Tobit and Probit Models. *Journal of Econometrics*, **92**(2), 275–294.

Tomiura, E. 2007. Foreign Outsourcing, Exporting, and FDI: A Productivity Comparison at the Firm Level. *Journal of International Economics*, **72**(1), 113–127.

Appendix A1: Proof that variable costs are decreasing in production intensity

The optimal constant marginal costs of production are

$$
\begin{aligned}
c^s(s) &= \int_0^{\omega_{NONI}} t(\omega) w^N d\omega + \int_{\omega_{NONI}}^1 w^N d\omega \\
c^m(s) &= \int_0^{\omega_{NOSO}} t(\omega) w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SONI}} t(\omega) w^S d\omega + \int_{\omega_{SONI}}^1 w^N d\omega \\
c^l(s) &= \int_0^{\omega_{NOSO}} t(\omega) w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SOSI}} t(\omega) w^S d\omega + \int_{\omega_{SOSI}}^1 w^S d\omega
\end{aligned}
$$

It is possible to show how these change wrt x by looking at the indirect effect of this through the cutoff ω's. The above expressions are for situations where $0 < \omega_{NONI}, \omega_{NOSO}, \omega_{SONI}, \omega_{SOSI} < 1$. If this is not the case, the expressions will reach corner solutions. Consider for example the possible situation for medium-sized firms if $\omega_{NOSO} < 0 < 1 < \omega_{SONI}$. In this case these firms will source all intermediate inputs through foreign outsourcing, and their variable costs of production will thus be $\int_0^1 t(\omega) w^S d\omega$, which is independent of x. Such corner solutions can, however, never lead to a positive relationship between x and $c(s)$, so by showing that $\frac{\partial c(s)}{\partial x}$ for all the interior solutions described above I will have proved that $c(s)$ is non-increasing in x for all possible values of x, including corner solutions.

For small firms the change in costs from a change in x will be

$$
\begin{aligned}
\frac{\partial c^s(s)}{\partial x} &= \frac{\partial c^s(s)}{\partial \omega_{NONI}} \frac{\partial \omega_{NONI}}{\partial x} = \frac{\partial \omega_{NONI}}{\partial x} \frac{\partial}{\partial \omega_{NONI}} \left[\int_0^{\omega_{NONI}} t(\omega) w^N d\omega + \int_{\omega_{NONI}}^1 w^N d\omega \right] \\
&= \frac{\partial \omega_{NONI}}{\partial x} \frac{\partial}{\partial \omega_{NONI}} \left[T(\omega_{NONI}) - T(0) + 1 - \omega_{NONI} \right] w^N \\
&= \frac{\partial \omega_{NONI}}{\partial x} \left[t(\omega_{NONI}) - 1 \right] w^N,
\end{aligned}
$$

where $T(\omega) = \int t(\omega)$. Inserting for $t(\omega_{NONI})$ we get

$$
\frac{\partial c^s(s)}{\partial x} = \underbrace{\frac{\partial \omega_{NONI}}{\partial x}}_{0>} \underbrace{\frac{f_I^N - f_O^N}{x}}_{>0} < 0.
$$

Since $t(\omega)$ is monotonically increasing in ω, the same must be true for the inverse

function. This implies that since

$$\frac{\partial t\left(\omega_{NONI}\right)}{\partial x} = -\frac{f_I^N - f_O^N}{w^N x^2} < 0 \implies \frac{\partial \omega_{NONI}}{\partial x} < 0.$$

With $\frac{f_I^N - f_O^N}{x} > 0$ and $\frac{\partial \omega_{NONI}}{\partial x} < 0$ it must be the case that $\frac{\partial c^s(s)}{\partial x} < 0$, and I have shown that costs are monotonically decreasing in x for small firms. Following the above approach, it is also easy to show that the same must be the case for medium and large firms. For medium-sized firms

$$
\begin{aligned}
\frac{\partial c^m\left(s\right)}{\partial x} &= \frac{\partial c^m\left(s\right)}{\partial \omega_{NOSO}}\frac{\partial \omega_{NOSO}}{\partial x} + \frac{\partial c^m\left(s\right)}{\partial \omega_{SONI}}\frac{\partial \omega_{SONI}}{\partial x} \\
&= \frac{\partial \omega_{NOSO}}{\partial x}\frac{\partial}{\partial \omega_{NOSO}}\left[\int_0^{\omega_{NOSO}} t\left(\omega\right)w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SONI}} t\left(\omega\right)w^S d\omega + \int_{\omega_{SONI}}^1 w^N d\omega\right] \\
&\quad + \frac{\partial \omega_{SONI}}{\partial x}\frac{\partial}{\partial \omega_{SONI}}\left[\int_0^{\omega_{NOSO}} t\left(\omega\right)w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SONI}} t\left(\omega\right)w^S d\omega + \int_{\omega_{SONI}}^1 w^N d\omega\right] \\
&= \underbrace{\frac{\partial \omega_{NOSO}}{\partial x}}_{0>}\underbrace{\frac{f_O^S - f_O^N}{x}}_{>0} - \underbrace{\frac{\partial \omega_{SONI}}{\partial x}}_{>0}\underbrace{\frac{f_O^S - f_I^N}{x}}_{>0} < 0,
\end{aligned}
$$

since

$$
\begin{aligned}
\frac{\partial t\left(\omega_{NOSO}\right)}{\partial x} &= -\frac{f_O^S - f_O^N}{\left(w^N - w^S\right)x^2} < 0 \\
\frac{\partial t\left(\omega_{SONI}\right)}{\partial x} &= \frac{f_O^S - f_I^N}{w^S x^2} > 0.
\end{aligned}
$$

For large firms the change in costs from a change in x will be:

$$
\begin{aligned}
\frac{\partial c^l(s)}{\partial x} &= \frac{\partial c^l(s)}{\partial \omega_{NOSO}}\frac{\partial \omega_{NOSO}}{\partial x} + \frac{\partial c^l(s)}{\partial \omega_{SOSI}}\frac{\partial \omega_{SOSI}}{\partial x} \\
&= \frac{\partial \omega_{NOSO}}{\partial x}\frac{\partial}{\partial \omega_{NOSO}}\left[\int_0^{\omega_{NOSO}} t(\omega)w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SOSI}} t(\omega)w^S d\omega + \int_{\omega_{SOSI}}^1 w^S d\omega\right] \\
&\quad + \frac{\partial \omega_{SOSI}}{\partial x}\frac{\partial}{\partial \omega_{NOSO}}\left[\int_0^{\omega_{NOSO}} t(\omega)w^N d\omega + \int_{\omega_{NOSO}}^{\omega_{SOSI}} t(\omega)w^S d\omega + \int_{\omega_{SOSI}}^1 w^S d\omega\right] \\
&= \underbrace{\frac{\partial \omega_{NOSO}}{\partial x}}_{0>}\underbrace{\frac{f_O^S - f_O^N}{x}}_{>0} + \underbrace{\frac{\partial \omega_{SOSI}}{\partial x}}_{0>}\underbrace{\frac{f_I^S - f_O^S}{x}}_{>0} < 0,
\end{aligned}
$$

since

$$
\begin{aligned}
\frac{\partial t(\omega_{NOSO})}{\partial x} &= -\frac{f_O^S - f_O^N}{(w^N - w^S)x^2} < 0 \\
\frac{\partial t(\omega_{SOSI})}{\partial x} &= -\frac{f_I^S - f_O^S}{w^S x^2} < 0.
\end{aligned}
$$

I have thus shown that $\frac{\partial c(s)}{\partial x} \le 0 \forall x$.

Appendix A2: Production intensity and productivity

The firms' profit function is given by

$$
\pi = A^{1-\rho}(\theta x)^\rho - c(x)x - f(x),
$$

with corresponding first order conditions:

$$
\rho A^{1-\rho}\theta^\rho x^{\rho-1} - \frac{\partial c(x)}{\partial x}x - c(x) - \frac{\partial f(x)}{\partial x} = 0.
$$

Total differentiation of this FOC yields:

$$\frac{dx}{d\theta} = \frac{\rho^2 A^{1-\rho}\theta^{\rho-1}x^{\rho-1}}{\rho(1-\rho)A^{1-\rho}\theta^{\rho}x^{\rho-2} + \frac{\partial^2 c(x)}{\partial x^2}x + 2\frac{\partial c(x)}{\partial x} + \frac{\partial^2 f(x)}{\partial x^2}}.$$

The numerator in this expression is always positive, whereas the denominator contains both positive and some potentially negative elements. For the FOC to denote a maximum, however, it must be that the second-order condition holds; $\frac{\partial^2 \pi}{\partial x^2} < 0$:

$$\rho(1-\rho)A^{1-\rho}\theta^{\rho}x^{\rho-2} + \frac{\partial^2 c(x)}{\partial x^2}x + 2\frac{\partial c(x)}{\partial x} + \frac{\partial^2 f(x)}{\partial x^2} > 0,$$

which is simply the denominator in the expression above. This implies that in optimum, an increase in productivity will increase the optimal production intensity, and $\frac{dx}{d\theta} > 0$. There will be a unique optimum if the SOC holds for all values of x. Intuitively, this will be the case if the reduction in marginal costs from an increase in x is never larger than the reduction in marginal revenue from the same change in x. In such a situation, a marginal increase in x would lead to a drop in variable costs of production sufficiently large to cause an even larger increase in x, and thus lead to a self-reinforcing process of falling costs and increasing production. If there are never such self-reinforcing effects, $\frac{\partial^2 \pi}{\partial x^2} < 0 \forall x$, and the equilibrium is unique.

Appendix A3: Mapping of the sourcing strategies

With four different sourcing options (NO, NI, SO, SI) there should exist six different $t(\omega)$-loci. However, the variable costs of both NI and SI are indifferent to $t(\omega)$, which means that for a given production intensity, one of these will be preferred to the other for all ω. More specifically, for $x < x_{NISI} = \frac{F_I^S - F_I^N}{w^N - w^S} \Longrightarrow NI \succ SI$ and vice versa. This leaves me with five $t(\omega)$-loci, which should imply a total of ten cutoff-values for x. Again there is a special case, as it turns out that $\omega_{SOSI} > \omega_{NONI} \ \forall \ x$. There are thus nine x-cutoffs that determine the rankings of $t(\omega)$-loci, plus x_{NISI} that determine when $NI \succ SI$. The complete list of these

x-values in descending order is:

$$x_{NOSINONI} = x_{SOSISONI} = x_{NISI} = \frac{f_I^S - f_I^N}{w^N - w^S}$$

$$x_{SONINOSI} = \frac{w^N \left(f_O^S - f_I^N\right) + w^S \left(f_I^S - f_O^N\right)}{\left(w^N - w^S\right)\left(w^N + w^S\right)}$$

$$x_{SONINONI} = x_{NOSOSONI} = x_{NOSONONI} = \frac{w^N \left(f_O^S - f_I^N\right) + w^S \left(f_I^N - f_O^N\right)}{w^N \left(w^N - w^S\right)}$$

$$x_{NOSONOSI} = x_{NOSISOSI} = x_{NOSOSOSI} = \frac{w^S \left(f_I^S - f_O^N\right) - w^N \left(f_I^S - f_O^S\right)}{w^S \left(w^N - w^S\right)}.$$

Category	x
1	$x_{NOSONOSI} > x$
2	$x_{NOSOSONI} > x > x_{NOSONOSI}$
3	$x_{SONINOSI} > x > x_{NOSOSONI}$
4	$x_{NOSINONI} > x > x_{SONINOSI}$
5	$x > x_{NOSINONI}$

This implies that for the smallest firms with $x < x_{NOSONOSI}$ it will be the case that

$\omega_{NOSO} > \omega_{NOSI}$	$\omega_{NOSI} > \omega_{SONI}$	$\omega_{SOSI} > \omega_{SONI}$	$\omega_{NONI} > \omega_{SONI}$
$\omega_{NOSO} > \omega_{SOSI}$	$\omega_{NOSI} > \omega_{NONI}$	$\omega_{SOSI} > \omega_{NONI}$	
$\omega_{NOSO} > \omega_{SONI}$	$\omega_{NOSI} > \omega_{SOSI}$		
$\omega_{NOSO} > \omega_{NONI}$			

This unambiguously determines the complete order to be

$$\omega_{NOSO} > \omega_{NOSI} > \omega_{SOSI} > \omega_{NONI} > \omega_{SONI}.$$

For a given production intensity x, this ranking defines six regions along the ω-axis. Below I have shown these regions for ascending ω values. Each column shows all pairwise rankings of sourcing options for that given range of ω values.[20] In each column there is one sourcing option that dominates all other sourcing options, and will be the one the firm will use for that range of tasks. The chosen sourcing option

---

[20] $NI \succ SI \; \forall \; \omega$ since $x < x_{NISI}$

is shown in the last row of the table.

$\omega < \omega_{SONI}$	$\omega_{SONI} < \omega < \omega_{NONI}$	$\omega_{NONI} < \omega < \omega_{SOSI}$	$\omega_{SOSI} < \omega < \omega_{NOSI}$	$\omega_{NOSI} < \omega < \omega_{NOSO}$	$\omega_{NOSO} < \omega$
$SO \succ NI$	$NI \succ SO$	$NI \succ SO$	$NI \succ SO$	$NI \succ SO$	$NI \succ SO$
$NO \succ NI$	$NO \succ NI$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$
$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SI \succ SO$	$SI \succ SO$	$SI \succ SO$
$NO \succ SI$	$NO \succ SI$	$NO \succ SI$	$NO \succ SI$	$SI \succ NO$	$SI \succ NO$
$NO \succ SO$	$NO \succ SO$	$NO \succ SO$	$NO \succ SO$	$NO \succ SO$	$SO \succ NO$
$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$
NO	NO	NI	NI	NI	NI

For these smallest firms, we can thus see that domestic outsourcing and domestic vertical integration will be the only sourcing options ever used. Further, the table shows that the cutoff between the two sourcing options is, quite naturally, ω_{NONI}. Similar tables for the other four production intensity categories are shown below.

Sourcing when $x_{NOSONOSI} < x < x_{SONINONI}$:

$\omega < \omega_{SONI}$	$\omega_{SONI} < \omega < \omega_{NONI}$	$\omega_{NONI} < \omega < \omega_{NOSO}$	$\omega_{NOSO} < \omega < \omega_{NOSI}$	$\omega_{NOSI} < \omega < \omega_{SOSI}$	$\omega_{SOSI} < \omega$
$SO \succ NI$	$NI \succ SO$	$NI \succ SO$	$NI \succ SO$	$NI \succ SO$	$NI \succ SO$
$NO \succ NI$	$NO \succ NI$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$
$NO \succ SO$	$NO \succ SO$	$NO \succ SO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$
$NO \succ SI$	$NO \succ SI$	$NO \succ SI$	$NO \succ SI$	$SI \succ NO$	$SI \succ NO$
$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SI \succ SO$
$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$
NO	NO	NI	NI	NI	NI

Sourcing when $x_{SONINONI} < x < x_{NOSISONI}$:

$\omega < \omega_{NOSO}$	$\omega_{NOSO} < \omega < \omega_{NONI}$	$\omega_{NONI} < \omega < \omega_{SONI}$	$\omega_{SONI} < \omega < \omega_{NOSI}$	$\omega_{NOSI} < \omega < \omega_{SOSI}$	$\omega_{SOSI} < \omega$
$NO \succ SO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$
$NO \succ NI$	$NO \succ NI$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$
$SO \succ NI$	$SO \succ NI$	$SO \succ NI$	$NI \succ SO$	$NI \succ SO$	$NI \succ SO$
$NO \succ SI$	$NO \succ SI$	$NO \succ SI$	$NO \succ SI$	$SI \succ NO$	$SI \succ NO$
$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SI \succ SO$
$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$
NO	SO	SO	NI	NI	NI

140

Sourcing when $x_{NOSISONI} < x < x_{NOSINONI}$:

$\omega < \omega_{NOSO}$	$\omega_{NOSO} < \omega < \omega_{NONI}$	$\omega_{NONI} < \omega < \omega_{NOSI}$	$\omega_{NOSI} < \omega < \omega_{SONI}$	$\omega_{SONI} < \omega < \omega_{SOSI}$	$\omega_{SOSI} < \omega$
$NO \succ SO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$
$NO \succ NI$	$NO \succ NI$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$
$NO \succ SI$	$NO \succ SI$	$NO \succ SI$	$SI \succ NO$	$SI \succ NO$	$SI \succ NO$
$SO \succ NI$	$SO \succ NI$	$SO \succ NI$	$SO \succ NI$	$NI \succ SO$	$NI \succ SO$
$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SI \succ SO$
$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$	$NI \succ SI$
NO	SO	SO	SO	NI	NI

Sourcing when $x_{NOSINONI} < x$:

$\omega < \omega_{NOSO}$	$\omega_{NOSO} < \omega < \omega_{NOSI}$	$\omega_{NOSI} < \omega < \omega_{NONI}$	$\omega_{NONI} < \omega < \omega_{SOSI}$	$\omega_{SOSI} < \omega < \omega_{SONI}$	$\omega_{SONI} < \omega$
$NO \succ SO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$	$SO \succ NO$
$NO \succ SI$	$NO \succ SI$	$SI \succ NO$	$SI \succ NO$	$SI \succ NO$	$SI \succ NO$
$NO \succ NI$	$NO \succ NI$	$NO \succ NI$	$NI \succ NO$	$NI \succ NO$	$NI \succ NO$
$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SO \succ SI$	$SI \succ SO$	$SI \succ SO$
$SO \succ NI$	$SO \succ NI$	$SO \succ NI$	$SO \succ NI$	$SO \succ NI$	$NI \succ SO$
$SI \succ NI$	$SI \succ NI$	$SI \succ NI$	$SI \succ NI$	$SI \succ NI$	$SI \succ NI$
NO	SO	SO	SO	SI	SI

Together, these five tables and the cutoff production intensities show all the information needed to draw Figure 1 in the paper.