Essays on Knowledge Flows, International Economics, and Entrepreneurship

Christian Fons-Rosen

London School of Economics and Political Science

London, September 2009

A thesis submitted to the Department of Economics of the London School of Economics and Political Science for the degree of Doctor of Philosophy

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly with me is clearly identified in it). The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I certify that chapter 2 of my thesis, "Relative Factor Endowments and International Portfolio Choice", was co-authored with Alejandro Cunat. I, Christian Fons-Rosen, contributed in excess of 50 percent to the genesis of the project, the work on the model, the data analysis, and the writing of the text.

Christian Fons-Rosen

Abstract

This thesis consists of three essays on either knowledge flows, international economics, or entrepreneurship. The first chapter focuses on knowledge flows and foreign direct investment. The second chapter aims to understand the pattern of cross-country equity portfolio allocations. The third chapter focuses on how entrepreneurship practices across countries is affected by bureaucratic circumstances.

Chapter I investigates whether FDI is a channel through which knowledge spills over from the foreign multinational to the host country. I analyse whether patents developed by local inventors in Central and Eastern Europe (CEE) cite the stock of patents of FDI multinationals more often after these companies have established themselves in CEE. Using a newly hand-collected data sample on privatisation cases resolved during the 1990s, I find that winning bidders experience a 20% greater increase in citations received by the host country compared to the losing bidder.

Chapter II presents a model of international portfolio choice based on cross-country differences in relative factor abundance. The change in factor prices after a positive shock in a particular country provides insurance to countries that have dissimilar factor endowment ratios. The main prediction is that countries with similar relative factor endowments have a stronger incentive to invest in one another for insurance purposes. Empirical evidence supports our theory.

Chapter III presents a model of corporate finance that incorporates bureaucratic start-up costs and where sectors differ in their need of external capital. The main theoretical prediction is that a reduction in start-up costs leads to an increase in the share of value added and number of firms in sectors with greater external finance. Intuitively, the sector with high external finance experiences a greater improvement in economies of scale, thereby making it more attractive to consumers. Using sector-country level data on manufacturing production, I find support for the predictions of the model.

Acknowledgements

I would like to thank my supervisor, Alwyn Young, and my advisor, Gianluca Benigno, for their great support, guidance, and constructive comments. Particularly in the final years of my PhD, I have also benefited from the Labour and Productivity groups at the LSE, and in particular from Guy Michaels, Steve Pischke, Barbara Petrongolo, Luis Garicano, John Van Reenen, and Radha Iyengar. I am also very grateful for comments and discussions with Nobuhiro Kiyotaki and Kosuke Aoki early on my PhD.

The Centre for Economic Performance has provided me with an intellectually rich academic environment, great colleagues to work with, and a very efficient and helpful IT group who were always ready to help me out.

I am very thankful for the financial support by the Bank of Spain, Fundacion Ramon Areces, and Fundacion Rafael del Pino, who have sponsored me at most stages of my PhD.

I am also indebted to a large number of people. To my family, and very especially to my parents, who have always shown unconditional love, support, and trust, especially in periods when my work was not progressing as we all had hoped for. Their encouraging and calming words have made this journey much more enjoyable and worthwhile. To my friends at the LSE, especially Ander Perez, Mariano Bosch, Ashley Taylor, Eva Vourvachaki, and Mirko Draca. Their friendship and affection have been very present at all stages of my work. Finally I also want to remember my friends who, even though not directly involved in my PhD or the LSE, were always available and very present in my life; a very special thank you to Vicente Mataix-Pastor, Damia Pujol, and Quique Andreu.

Christian Fons-Rosen

Contents

List of Figures

List of Tables

Chapter 1

Knowledge Flows Through FDI Activity: The Case of Privatisations in Central and Eastern Europe

1.1 Introduction

With the development of endogenous growth theory (Romer 1986, 1990; Lucas 1988; Grossman and Helpman 1991; Aghion and Howitt 1995), the economic profession came to accept the view that innovation, knowledge spillovers, and R&D were key factors for self-sustained, long-term economic growth and industrial development.

A key driver for a country to benefit from innovation generated abroad is its absorptive capacity, that is, the ability to tap into the world technology pool. Trade flows, foreign direct investment (FDI), and labour mobility are among the best conduits for absorbing and diffusing knowledge and technology from abroad. FDI not only exposes local firms to global best practice technology and management techniques, but also exerts competitive pressure on corporate governance.

Policy makers in many transition and developing economies place FDI inflows high on their agenda, expecting it to bring new technologies and management skills, apart from generating productivity spillovers to domestic entities. Many governments even offer foreign companies more favourable conditions than those granted to domestic firms through subsidies and tax holidays.

Despite being very important from a policy perspective, there is little evidence suggesting that domestic firms benefit from FDI. Rodrik (1999) concluded that "today's policy literature is filled with extravagant claims about positive spillovers from FDI but the evidence is sobering", and ten years later the picture is not much more optimistic: "what is rarer is evidence of productivity spillovers to domestic firms" (World Bank, 2008a). As a conclusion, empirical research based on firm-level panel data tends to produce mixed results and the question is far from solved.

While these statements hold in general across all regions, they are particularly true in Central and Eastern Europe (CEE). Before 1990 these countries were almost completely closed to foreign activity. They had a well educated population and a very important legacy of scientific knowledge compared to most of non-OECD countries. Among the enormous structural changes that these countries have undergone during the 1990s, there has been a growing presence of FDI activity. Figure (1.1) shows the evolution of FDI inflows as a percentage of GDP[1] in the Czech Republic, Hungary, and Poland. Before 1990, FDI was non-existent, but by the year 2000 the three countries had FDI inflows above 5% of GDP on a fairly stable basis.

In this paper, I use a new methodological approach and identification technique to analyse whether the knowledge pool of foreign multinationals becomes more accesible to local firms in CEE *after* FDI takes place. While the previous literature on FDI in CEE follows the convention of estimating a production function, this paper rather uses patent citations as an indicator of knowledge flows.[2] I analyse whether patents generated by

[1] Source: World Bank, World Development Indicators (Internet download June 2009)

[2] For evidence that patent citations incorporate a substantial signal component reflecting patterns of knowledge flows, see Jaffe et al. (1998) and Jaffe et al. (2000).

foreign FDI firms experience a relative increase in citations received by local CEE inventors, compared to foreign non-FDI firms. In other words, I address the question of whether patents generated by local CEE inventors cite the stock of patents of foreign FDI multinationals more often *after* these enterprises have established themselves in CEE.

In the first part of the paper, I make extensive use of two datasets: a patents dataset and an ownership database. I use the patents data from the European Patent Office (EPO) and the United States Patents and Trademark Office (USPTO) between 1978 and 2007 that can be found in PATSTAT (April 2008). In particular, I select the universe of patents developed by inventors located in either the Czech Republic, Hungary, or Poland, and identify all the citations made to patents belonging to foreign companies. Then, I construct measures of yearly citations received by each foreign company.[3]

In order to identify which foreign firms are doing FDI, we use the ownership information provided by Bureau Van Dijk (BvD)'s Amadeus/Orbis datasets.[4] They provide information on firm ownership, including the identification of the ultimate owner, which identifies the one single firm/person/entity that ultimately owns the firm. Variables available for the ultimate owner include, among others, the country of origin and the BvD identification number. To identify which foreign companies do FDI in CEE, we take the population of BvD firms in the Czech Republic, Hungary, and Poland, and identify their ultimate owner. The subsample of ultimate owners coming from foreign countries are, by definition, doing FDI in CEE.[5]

After a matching process by which each firm is allocated the same BvD identification number in both PATSTAT and Amadeus/Orbis datasets, our group of foreign FDI

[3]We limit our sample to foreign companies that have been cited at least once by inventors in our three CEE countries.

[4]For more detailed information on the patents and ownership databases, see Appendix.

[5]Since our panel of ownership information starts in 1995 and some firms might have started their FDI activities in earlier years, we have completed the information with two datasets on changes in ownership: Thomson Financial SDC and BvD's Zephyr. Additionally, we went to the national accounts of each of these three host countries to obtain the starting year of FDI for the remaining FDI companies.

firms is characterised by doing FDI in at least one of the CEE countries, apart from receiving at least one citation by inventors from this same CEE country.[6]

In the second part of the paper, this study goes beyond the existing literature in terms of the identification technique by selecting the comparison group of non-FDI firms in a more precise way. One of the main challenges of the literature on FDI firms and knowledge flows is the choice of an appropriate comparison group. FDI firms are not a random sample of foreign multinationals, and the reasons to undertake FDI are often unobserved to the econometrician. Since it is very hard to identify a non-FDI firm that is as similar as possible to an FDI firm, the coefficient estimates obtained from a standard difference-in-difference methodology can be misleading.

In order to circumvent this concern, I collected data on the privatisation processes that took place in the Czech Republic, Hungary, and Poland during the 1990s, which included a variety of privatisation methods. An important number of state-owned companies were privatised through an open tender or competitive bid in which foreign companies could place bids to acquire the control of the company about to be privatised. Among all the bids received, the national authority usually proceeded with a selection process until two or three final bids were still considered in the final round. Finally, the authorities chose the winning bidder among the final remaining bids. The identifying assumption is that losing bidders are the closest possible firms to the winning bidders, and thereby form a valid counterfactual. We rely among the revealed ranking of firms to identify a valid counterfactual.

The main finding is that patents generated by foreign FDI firms experience a relative increase in citations received, compared to foreign non-FDI firms, *after* FDI takes place. That is, patents generated by local CEE inventors cite the stock of patents of foreign FDI multinationals more often *after* these enterprises have established themselves in CEE, compared to the stock of patents developed by foreign non-FDI firms. Interestingly,

[6]Similarly, our group of foreign non-FDI firms receives at least one citation by CEE countries, but is not the ultimate owner of any company in the Czech Republic, Hungary, and Poland.

starting FDI in itself does not lead to a substantial increase in citations received in all of our econometric specifications. Rather, what is needed is that the FDI firm has spent enough time in the host country, thereby allowing for its ideas and knowledge to be diffused and absorbed by local inventors. In other words, the increase in citations received by the FDI firms is positively related to how long the firm has been present in the host country.

This paper is related to four streams of the literature. First, the empirical literature on FDI and knowledge spillovers runs a production function equation[7] in which the regressor of interest is the share of FDI in a given sector or region. A number of survey papers have concluded that the evidence supporting these positive knowledge spillovers is either non-existent or very weak.[8] This is also true for studies that focus particularly on CEE. Djankov and Hoekman (2000) find rather a negative effect of FDI on domestic firms in the Czech Republic; Konings (2001) casts doubt on horizontal FDI spillovers in Bulgaria, Romania, and Poland; Damijan et al. (2003) look at a number of transition economies and conclude that FDI does not generate positive intra-industry spillovers for domestic firms. Bosco (2001) similarly claims that the evidence for technological spillovers in Hungary does not allow for clear-cut conclusions. More optimistic is the message of Javorcik (2004) who finds backward spillovers in Lithuania, and of Kinoshita (2001) who reports that some Czech manufacturing sectors benefit from horizontal spillovers of FDI.

Compared to this literature, we focus on patent citations instead of estimating a production function equation. Apart from shifting the analysis from knowledge spillovers to knowledge flows, our approach will rather focus on the larger companies of these economies, given that our focus is on firms active in R&D. An interesting fact to support the patent citation approach is that the contribution of multinationals' R&D to total

[7]As a general benchmark, firm-level output levels or growth rates are regressed on inputs (e.g. capital, labour, materials), leading to a residual interpreted as total factor productivity. FDI spillovers are found if multinational presence are positively correlated with the productivity residual.

[8]See Gorg and Strobl (2001), Crespo and Fontoura (2007), Gorg and Greenaway (2004).

R&D is substantial in our countries of interest, reaching more than 60 percent in Hungary (World Bank, 2008b).

Even though the previous literature has almost unanimously rejected any positive effect of FDI on local firms in CEE, some anecdotal evidence on absorptive capacity still suggests the possibility of some positive effect that the previous literature was unable to capture. First, Kinoshita (2001) finds that the learning effect (i.e. absorptive capacity) of R&D in Czech manufacturing firms is far more important than the innovative effect in explaining the productivity growth of a firm. Second, Javorcik and Sparateanu (2005) find that in a survey of enterprises, almost a quarter of respondents in the Czech Republic and 15% in Latvia learned from multinationals about new technologies.[9]

Secondly, there is a very limited literature on FDI and patent citations. We highlight Branstetter (2006), who does a before-after analysis of 200 Japanese multinationals. He tests the hypothesis that FDI is a channel of knowledge spillovers for Japanese multinationals undertaking FDI in the U.S. and finds that FDI increases the flow of knowledge spillovers both from and to the investing Japanese firms.[10]

Thirdly, the literature on patent citations and geography started with the seminal paper by Jaffe et al. (1993). Inventors cite other inventors living in geographical proximity more than proportionally. Griffith et al. (2007) examine the "home bias" of international knowledge spillovers measured by the speed of patent citations and find that the geographical localisation of knowledge spillovers has fallen over time. Griffith et al. (2006) analyse the relationship between U.S. and U.K. and provide evidence of knowledge spillovers associated with technology sourcing.[11]

Finally, the identification strategy used in the second part of this paper is closest

[9]On theoretical literature related to FDI and absorptive capacity, see Javorcik and Saggi (2003), and Leahy and Neary (2007).

[10]Singh (2005) uses patent citations to understand the role of multinational subsidiaries in the diffusion of knowledge. Greater subsidiary activity increases cross-border knowledge flows between host and source country.

[11]Audretsch and Feldman (1996) and Keller (2002) give evidence in favour of the fact that mobility of engineers across firms matters for localised spillovers.

to Greenstone and Moretti (2004) and Greenstone et al. (2007). In these papers, U.S. counties compete for a large plant to locate within their boundaries, and the authors compare the winning county to one or two runner-up counties to analyse the effect on property values, productivity, and welfare.

The structure of the paper is as follows. Section 1.2 introduces the data and provides summary statistics. Section 1.3 explains the econometric methodology, while Section 1.4 shows results. Section 1.5 concludes.

1.2 Data and Summary Statistics

1.2.1 Data

The database on patents comes from the European Patent Office (EPO) and the U.S. Patent and Trademark Office (USPTO), and is included in the Worldwide Patent Statistical Database (PATSTAT). The PATSTAT database contains information on all patent applications to the EPO and USPTO, including information about applicant (name and location), inventors (name and location), granted status, technology class, year of application, and citations made and received. The data dates back to 1978 for EPO (i.e. the year when EPO was launched) and much earlier for the USPTO.[12] We are able to track the number of patents awarded and citations received for each firm over time. Of particular interest for this project is the fact that for each citation we can identify both the *citing* applicant and the *cited* applicant, apart from the year in which the citation takes place.

The database on ownership comes from BvD's Amadeus and Orbis. For European countries, we use firm-level data from the Amadeus (Analyse MAjor Databases from EUropean Sources) database. This standardised commercial data is collected by about

[12]These DVDs are provided twice a year, and the version used for this paper is April 2008, which incorporates all the population of patents since the beginning of EPO in 1978. For the USPTO, we will also use data from 1978 onwards.

50 vendors (generally the office of register of companies) across Europe. The database contains financial information on about 8 million firms from 34 countries, including all the European Union countries and Eastern Europe. Additionally, for other major foreign investors in CEE, we obtain the equivalent information from BvD's Orbis dataset, which is the extension of Amadeus to the rest of the world.[13] Amadeus/Orbis also provides information on firm ownership, including the ultimate owner. In particular, it identifies the one single firm/person/entity that ultimately owns the firm.[14] To define an ultimate owner, BvD analyses the shareholding structures of companies that, according to the independence indicator, are not independent from their shareholders. BvD looks for the shareholder with the highest percentage of ownership. If this shareholder is independent, it is recognised as the ultimate owner. If the highest shareholder is not independent, the same process is repeated until the ultimate owner if found.

For this particular project, it is necessary to identify the foreign companies that do FDI in CEE. We take the population of firms in the Czech Republic, Hungary, and Poland, and identify their ultimate owner, specifically by the ID number provided by BvD. All the ultimate owners coming from foreign countries are, by definition, doing FDI in CEE.[15]

For the dataset on the bidders of privatisation cases in the Czech Republic, Hungary, and Poland during the 1990s, we needed to collect the data ourselves. The information is not archived or available from any central governmental agency, so that public or official channels were only marginally helpful. Given the diverse nature of privatisation methods chosen by each country, we had to resort to a variety of procedures to gather the data. A number of research assistants were allocated to each host country to perform

[13]While we obtain Amadeus information from a number of DVDs over time, for Orbis we used the web interface download available to LSE since 2008/2009.

[14]Variables available for the ultimate owner include: country of incorporation/origin; ID number (if the ultimate owner is present in the Amadeus/Orbis database); type (e.g., family, industrial firm, employees/managers, financial institution, state).

[15]For a description of how the matching was done between Patstat and Amadeus/Orbis, see Appendix A.

preliminary research and investigation. After the preparatory work, we begun interviewing numerous academic scholars and national officials who played a role during the privatisations undertaken during the 1990s. The interviewees either participated in the decision making process carried out by the "Evaluation Committee" or contributed to academic literature on privatisation. We also conducted further interviews with specific company managers, and gathered valuable information from national archives, privatisation agencies, state audit departments, magazines and journals providing general privatisation information or industry-specific case-studies, or daily financial papers focusing on Central and Eastern Europe, among others.

1.2.2 Summary Statistics

Table (1.1) provides broad descriptive features of the data. Based on Table (1.1)A, the total number of citations in the sample is approximately 13,000.[16] While only 8% of the cited foreign firms do FDI according to our ownership information, these firms obtain 21% of the citations. This percentage rises to 33% in the case of Poland, where 11% of the firms in our sample do FDI. In Table (1.1)B, the correlation between a firm's stock of patents and its citations received from outside the Czech Republic, Hungary, and Poland is 0.75. Interestingly, this correlation goes down to 0.31 between the stock of patents and the citations received from the three host countries analysed in this paper. The correlation between citations received from the three CEE countries and citations received from the rest of the world is only 0.39.

In Table (1.2) we provide separate information for FDI and non-FDI firms. A higher fraction of European firms does FDI compared to U.S. firms, which makes intuitive sense due to the fact that the barriers to FDI (e.g. geographical and cultural distance) are smaller for European companies. In terms of technology classes, six IPC classes cover the whole group of firms, each of these classes covering less than 25% of firms in

[16]Self-citations have been excluded.

our sample. Consequently, it is reassuring that our results will not be driven by a single technological category.

In Table (1.3), we compare FDI firms to non-FDI firms in terms of their main variables used in our empirical analysis. In all host countries, we observe that FDI firms have a higher stock of patents and receive more citations on a yearly basis compared to non-FDI firms. These differences are especially important for Poland.

Figure (1.2) shows the evolution over time of the average citations received per firm, decomposed into FDI firms and non-FDI firms. Until the beginning of the 1990s, we do not observe substantial differences. It is only from the mid-1990s onwards that the gap widens substantially: FDI firms start to receive many more citations compared to non-FDI firms.

What remains to be seen is for whom these *citing* inventors located in CEE work for. Figure (1.3) provides an answer to this question. Until the beginning of the 1990s, almost all citations were made by inventors working for local CEE companies. This is due to the fact that foreign ownership was almost non-existent. But during the mid-1990s we observe a huge increase in citations made by inventors employed in foreign companies, which shows that FDI did not only aim at CEE for lower wages, but also targetted a new location to start R&D activity. Before 2000, we already observe that most of the citations are made by inventors working for foreign companies, and this trend intensified during the following years.

1.3 Empirical Methodology

1.3.1 Broad Sample

In our first specification, $\ln(c_{cit})$ is (the log of) the number of citations made by inventors of host country c to the patent stock of firm i in year t:

$$\ln(c_{cit}) = \beta_0 + \beta_1 FDI_{ci} + \beta_2 Post_{cit} + \varepsilon_{cit} \qquad (1.1)$$

where FDI_{ci} is a dummy variable taking the value of 1 when firm i does FDI in country c at any point in time. $Post_{cit}$ is a dummy variable taking the value of 1 for FDI firms i in the subsample of years in which FDI is actually taking place in country c.[17] Our interest will be on the coefficient β_2. Do inventors of country c increase the number of citations made to firm i *after* this firm started doing FDI in country c? A positive and significant value of β_2 would support this prediction.

In our second specification, we expand the formulation in the following way:

$$\ln(c_{cit}) = \beta_0 + \beta_1 FDI_{ci} + \beta_2 Post_{cit} + TotYears_{cit} + \varepsilon_{cit} \qquad (1.2)$$

where $TotYears_{cit}$ is the cumulated number of years that firm i has been doing FDI in country c until time t.[18] This specification allows for additional flexibility to assess whether the increase in citations received happens right after FDI starts to take place, i.e. $\beta_2 > 0$, or whether it rather only increases gradually with time, i.e. $\beta_3 > 0$.[19]

From a statistical viewpoint, for both (1.1) and (1.2) one observes a large number of observations for which the dependent variable is zero, so that the log of the dependent variable remains undefined. I deal with this concern in two ways. First, when using an OLS specification, I rewrite the dependent variable as $\ln(1 + c_{cit})$.[20] Second, in our preferred econometric specification we use the Poisson estimator, which is especially suitable for count data.[21] Figure (1.4) illustrates the probability density function of both

[17]Therefore, $FDI_{ci} = 1$ is a necessary condition for observing $Post_{cit} = 1$.

[18]Therefore, $TotYears_{cit} > 0$ whenever $Post_{cit} = 1$. $TotYears_{cit}$ takes the value of 5 when the firm is in its fifth year of FDI, and so on.

[19]One can think of β_2 as the intercept and β_3 as the slope.

[20]This solution has been widely adopted in the international economics literature when regressing a gravity-type equation.

[21]See Santos-Silva and Tenreyro (2005) in support of the Poisson estimator when the dependent variable has a large amount of zeroes. When using the Poisson estimator, our dependent variable will not appear in logarithms.

the observed values in the dependent variable and the predicted values of the Poisson estimated values. We see that the Poisson fits very well the distribution of the true data.

In terms of control variables, four types of fixed effects are incorporated: (i) year; (ii) source country; (iii) host country; (iv) technology. Among other things, year fixed effects should control for the fact that in later years we observe many more citations. The (source and host) country fixed effects control for the average number of citations made and received by a firm of a certain country. Furthermore, the PATSTAT dataset provides an IPC code identifying the technology class of each patent. By aggregating this measure to the firm level, we can also control for the average number of citations received by a firm in each technology class.[22] In addition to these four types of fixed effects, we will also have fixed effects at the firm level for each host country separately, i.e. firm-host dummies.[23] Since source, host, and technology dummies are constant at the firm-host level, we will rather interact them with year dummies whenever firm-host dummies are present.

An additional control variable will be (the log of) the stock of patents that a firm has accumulated at each point in time. This should control for the fact that the number of citations received by a firm might depend on how much R&D investment has been done by this firm in the past.

Even though in our most complete specification we already incorporate firm-host fixed effects, apart from source*year, host*year, and technology*year dummies, it could still be the case that a firm decides to do FDI *as a consequence* of having been very successful in their R&D activity in previous years.[24] If high success in a firm's R&D activities leads to both a higher firm-level productivity level together with greater ex-

[22] We allocate a technology class to each firm in the following way: first, we allocate an IPC code to each patent among eight possible codes. Second, we identify the most common IPC code for the patents belonging to a firm, and allocate this code to this firm.

[23] In other words, if a company does FDI in both Hungary and Poland, this company will be allocated two different firm-level fixed effects.

[24] In the framework by Helpman et al. (2004), the sunk cost of FDI is larger than the one of exporting, so that only the most productive firms will choose to do FDI.

pected profits from expanding their activities to other countries, it might be that only the most innovative and profitable firms self-select to doing FDI. In that case, an increase in citations received would partly be the consequence of great innovative success, rather than of doing FDI. In order to rule out this alternative explanation, for each firm-year observation, we control for the citations received by this firm from the rest of the world.

1.3.2 Privatisation Cases

The number of observations in the second part of the paper is substantially reduced. We will only use information about privatisation cases that underwent a competitive bid procedure and for which we know both the winning and the losing bidders. Therefore, any case of greenfield FDI or foreign companies investing in CEE through stock market investments will not be used. The benefit of this new strategy is to use a difference-in-difference econometric specification in which a comparison group (losing bidders) is carefully selected and allocated to each treatment group (winning bidders).

The econometric specifications will be very similar to the ones used for the broader sample and for simplicity we will follow previous notation. In the simplest specification, a dummy variable $Post_{cit}$ equals one to indicate the subset of years after the privatisation decision has taken place:

$$\ln(c_{cbit}) = \beta_0 + \beta_1 FDI_{cbi} + \beta_2 Post_{cbit} + \beta_3 (FDI_{cbi} * Post_{cbit}) + \varepsilon_{cbit}. \quad (1.3)$$

A positive value of the new coefficient of interest β_3 can be interpreted in the following way: after the competitive bid of case b was resolved, the winning bidder experiences a greater increase in citations received from host country c, compared to the increase in citations received by the losing bidder. Apart from this new coefficient, all the remaining coefficients keep the same interpretation as in the regressions used for the broader sample.

In our second specification, we drop the dummy variable $Post_{cbit}$ and instead use the variable $TotYears_{cbit}$, which takes the same value as the cumulated number of years passed since the privatisation tender was resolved:

$$\ln\left(c_{cbit}\right) = \beta_0 + \beta_1 FDI_{cbi} + \beta_2 TotYears_{cbit} + \beta_3 \left(FDI_{cbi} * TotYears_{cbit}\right) + \varepsilon_{cbit}. \tag{1.4}$$

The new interpretation of coefficient β_3 is slightly different. A positive and statistically significant value would suggest that the greater increase in citations received by the winning bidder is especially observed after the foreign company has been present in the host country long enough for its knowledge stock to be diffused.

Our final specification simply combines the previous two expressions,

$$\begin{aligned} \ln\left(c_{cbit}\right) &= \beta_0 + \beta_1 FDI_{cbi} + \beta_2 Post_{cbit} + \beta_3 TotYears_{cbit} \\ &\quad + \beta_4 \left(FDI_{cbi} * Post_{cbit}\right) + \beta_5 \left(FDI_{cbi} * TotYears_{cbit}\right) + \varepsilon_{cbit}. \end{aligned} \tag{1.5}$$

The goal is to assess which of the previous two effects has a stronger impact from a statistical viewpoint. A positive β_4 would suggest that the increase in citations received by the winning bidder already expresses itself shortly after the privatisation tender was resolved. Rather, a positive β_5 would provide support for the fact that it takes time for the knowledge stock of the winning bidder to diffuse across the investors located in the host country.

1.4 Results

1.4.1 Broad Sample

Table (1.4) provides results of our benchmark specification using the Ordinary Least Squares (OLS) estimator. Equations (1)-(4) include year, host and source country, and technology fixed effects, while equations (5)-(8) rather incorporate interactions of year dummies with the last three dummy categories. Throughout all eight specifications, our coefficient of interest on "Dummy for post-FDI years" is approximately 0.1, and statistically significant at the 1% level. In terms of economic significance, it means that FDI firms experience a 10% increase in citations received from inventors located in host country after they start doing FDI in that country. The coefficient on "Dummy for firms doing FDI" changes sign once we add the control variables on the stock of patents and citations received in the rest of the world. It seems to suggest that even though FDI firms are cited more often than non-FDI firms at all times, it is completely driven by their greater size in terms of patents and citations. In addition, the coefficient on citations received by the rest of the world is positive and very strongly statistically significant in all specifications. Nevetheless, the coefficient on "Dummy for post-FDI years" is greater in absolute value.

As it was expected from the correlation between these two variables, companies that receive more citations in the rest of the world also receive more citations in our three host countries. But the increase in citations received by FDI firms in CEE is robust to controlling for how "attractive" these firms are from an innovation viewpoint to the rest of the world. It rules out the possibility that the increase in citations is a consequence of very productive R&D efforts in recent years, that in turn has led these firms to start FDI activities.

Table (1.5) presents OLS results allowing for the change in citations received to depend on the cumulated years that the FDI firm has been present in the host country. Under this new specification, the coefficient on "Dummy for post-FDI years" is still positive and statistically significant, but is about half in size. This reduction in the coefficient is now captured by the coefficient on "Stock of years doing FDI". Even in the most demanding specification with firm-level fixed effects, and year dummies interacted with host country, source country, and technology class, we observe a positive and statistically significant coefficient implying that each additional year of FDI leads to an increase in citations received of half a percentage point. In terms of the remaining control variables, we do not observe any substantial change, neither in size nor in significance.

Table (1.6) presents the equivalent regressions to Table (1.4), but rather uses the Poisson estimator. In equations (1)-(4), our coefficient of interest is still positive and strongly statistically significant. But as opposed to the OLS estimations, equations (5)-(8) do not consistently provide evidence on FDI leading to an increase in citations received. Similarly to Table (1.4), the coefficient on "Dummy for firms doing FDI" changes sign once we add our two main control variables, and the coefficient on citations received by the rest of the world is positive and very statistically significant.

Table (1.7) presents results for the Poisson estimator including the variable "Stock of years doing FDI". The picture is very clear throughout all specifications: starting FDI in itself does not lead to an increase in citations received. Rather, what is needed is that the FDI firm has spent enough time in the host country, thereby allowing for its ideas and knowledge to be diffused and absorbed by local inventors. This conclusion also explains why in equations (5)-(8) of Table (1.6) we rather struggled to find support for our prediction. It is remarkable that the coefficient on "Dummy for post-FDI years" is always close to zero and never statistically significant. On the other hand, the coefficient on "Stock of years doing FDI" is always positive and statistically significant.

1.4.2 Privatisation Cases

Before we describe the results of the main regressions with bidders, let us first get a visual idea of the data. Figure (1.5) shows the average yearly citations per firm decomposed into winning and losing bidders. Until the starting year of FDI, i.e. the year in which the privatisation case was resolved, both types of firms seem to receive more or less the same number of citations from inventors located in our CEE countries. Interestingly, once the winning bidder enters the host country, the number of citations it receives rises substantially, while the citations received by the losing bidder follows a much flatter trend. This result still holds in Figure (1.6) once we include dummies for years, host countries, and source countries.

Table (1.8) provides descriptive information about the sample of firms. We have 47 winners and 89 losers, and the main represented countries are the U.S., France, and Germany. Table (1.9) compares both the level and the growth rate of citations received by winners and losers. The first two rows provide information about the level effect. On average, winning bidders receive more citations overall in the EPO/USPTO datasets, and also receive more citations when we limit the sample to citations made by inventors located in CEE countries. Notice that the econometric specification will already control for differences in levels, so that these differences do not undermine our identification technique. Rather, losers form a valid counterfactual if we do not observe any differences in growth rates (or trends), compared to the winning bidders.

The third and fourth rows of Table (1.9) use overall citations available in EPO/USPTO. Before the privatisation process has been resolved, winning bidders seems to have a slightly higher growth rate in citations received, while the opposite is true after FDI has starts to take place. But these differences are not statistically significant.

In the last two rows of Table (1.9) we restrict the sample to citations made by inventors located in either the Czech Republic, Hungary, or Poland. Before the year in which the privatisation case is resolved, we do not observe any statistically significant difference between the growth in citations received by winning bidders and losing bidders.[25] On the other hand, after the privatisation resolution the winning bidder experiences a growth rate in the number of citations received that is substantially greater than the one of the losing bidder. The difference is statistically significant at the 1% level.

Table (1.10) provides OLS results for the three specifications described previously. Equations (1)-(3) provide benchmark results with the following economic interpretation: winning bidders experience a 20% greater increase in citations received by the host country after the privatisation case was resolved, compared to losing bidders. Interestingly, most of this effect does not seem to take place right after the foreign company starts its FDI activity in the host country. Rather, every additional year of FDI seems to increase the gap between citations received by the two bidders at a rate of 2% with strong statistical significance. Equations (4)-(6) show the same regressions after additionally including fixed effects for years, host countries, and source countries. Both the statistical and the economic significance does not seem to be affected, while the R-squared has risen from 0.16 to 0.22. In terms of control variables, the (log of the) stock of patents enters with a positive and statistically significant sign. The dummy variable for FDI firms enters with a negative sign, suggesting that at all times these firms are cited less than the losing bidders. While our descriptive statistics suggest the opposite, once we control for the stock of patents we also observed a similar result in our broader sample. A possible interpretation is that foreign firms with very little exposure to CEE before FDI takes place are especially interested in having access to this new region.

Table (1.11) presents the same results as Table (1.10), except that we also include case-specific dummies. The idea is to control for the possible impact that particular characteristics or circumstances of privatisation cases could have on the citations received by these companies. The coefficient values of our interaction terms of interest

[25]The growth rate is measured in the following way: we take the citations received by each firm in the first five years of our dataset, and compute the growth rate by taking the citations received by this same firm in the five years before the privatisation case is resolved.

remain almost unaffected. The main difference compared to Table (1.10) is that the dummy variable for FDI firms is not statistically significant anymore, implying that winning and losing bidders seem to receive approximately the same number of citations at all times. In our more demanding specification, Table (1.12) substitutes case dummies for firm dummies. The R-squared reaches 0.32 and our interaction term between the FDI firms and the variable capturing the cumulated number of years of FDI presence in the host country is still statistically significant at the 1% level.

Tables (1.13)-(1.15) present results for the identical specifications as in Tables (1.10)-(1.12), except that now the Poisson estimator is used instead of OLS. Equations (1)-(3) provide a familiar result. When included separately, the interaction terms for both of our variables of interest are positive and strongly statistically significant. Once the two interaction terms are included in the same regression, the main driver of our result is again the interaction term including the cumulated stock of years that the winning bidder has been present in the host country. Equations (4)-(6) include dummies for years, host countries, and source countries, and provide the same intuition, except for a lower statistical significance in the last equation. Finally, Tables (1.14) and (1.15) present results with case dummies and firm dummies, respectively. Both the economic and statistical significance of our results is unaltered, while the R-squared reaches 0.41.

1.5 Conclusion

I have used patent citations data to evaluate whether foreign direct investment is an important channel by which knowledge flows are transmitted from the foreign multinational to the host country. Apart from presenting results for a broad sample of FDI multinationals in CEE countries, I also use a novel identification technique with hand-collected data from the privatisation processes that took place in these countries during the 1990s. In particular, I gather data on winning and losing bidders of the privatisation cases that we resolved by a public tender or a competitive bid, whereby the losing bidder is assumed to be a valid counterfactual for the winning bidder. Both sets of results support the conclusion that patents developed by local inventors in CEE cite the stock of patents of FDI multinationals more often after these companies have established themselves in CEE.

The literature on knowledge flows and FDI is almost completely silent about the precise transmission mechanism by which host inventors benefit from the presence of foreign multinationals.[26] Furthermore, FDI can take many different forms with potentially very different implications for the host economy, even though this aspect is not fully understood yet.

For this reason, in future work I will move from a panel of firms to a panel of inventors. By tracking inventors over time and location, apart from identifying their co-authors and citations pattern, I aim to address the following questions: (i) does it matter whether FDI firms just focus on production activities or rather also open an R&D lab in the host country?; (ii) how does the interaction between inventors located in the home and host countries affect the intensity of knowledge flows?; (iii) Figure (1.3) shows that most inventors in CEE are now working for foreign multinationals. How much local innovation is crowded out when FDI multinationals hire top local inventors?

[26]An exception is Veugelers and Cassiman (2004) who use Belgian survey data and find that foreign subsidiaries are not more likely to transfer technology to the local economy as compared to local firms.

Bibliography

[1] Aghion, P. and P. Howitt (1995). "Technical Progress in the Theory of Economic Growth." In Jean-Paul Fitoussi, ed., *Economics in a Changing World: Proceedings of the 10th World Congress of the International Economic Association*, Vol. 5. London: Macmillan.

[2] Audretsch, D. and M. P. Feldman (1996). "R&D Spillovers and the Geography of Innovation and Production". The American Economic Review, vol. 86, no. 3; pp. 630-640.

[3] Bosco, M. G. (2001). "Does FDI Contribute to Technological Spillovers and Growth? A Panel Data Analysis of Hungarian Firms". Transnational Corporations, vol. 10, no. 1; pp. 43-68.

[4] Branstetter, L. (2006). "Is Foreign Direct Investment a Channel of Knowledge Spillovers? Evidence from Japan's FDI in the United States". Journal of International Economics, vol. 68, pp. 325-344.

[5] Crespo, N. and M. P. Fontoura (2007). "Determinant Factors of FDI Spillovers - What Do We Really Know?". World Development, vol. 35, no. 3; pp.410-425.

[6] Damijan, J.P., M. Knell, B. Majcen, M. Rojec (2003). "The Role of FDI, R&D Accumulation and Trade in Transferring Technology to Transition Countries: Ev-

idence from Firm Panel Data for Eight Transition Countries". Economic Systems, vol. 27; pp. 189-204.

[7] Djankov, S. and B. Hoekman (2000). "Foreign Investment and Productivity Growth in Czech Enterprises". World Bank Economic Review, vol. 14, no. 1; pp. 49-64.

[8] Gomez-Salvador, R., Messina, J. and Vallanti, G. (2004). "Gross job flows in Europe", Labour Economics, vol. 11,pp. 469–85.

[9] Gorg, H. and D. Greenaway (2004). "Much Ado about Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment?". The World Bank Research Observer, vol. 19, no. 2; pp.171-197.

[10] Gorg, H. and E. Strobl (2001). "Multinational Companies and Productivity Spillovers: A Meta-Analysis". Economic Journal, vol. 111, pp. 723-739

[11] Greenstone, M. and E. Moretti (2004). "Bidding for Industrial Plants: Does Winning a "Million Dollar Plant" Increase Welfare?". MIT Working Paper 04-39.

[12] Greenstone, M., R. Hornbeck, and E. Moretti (2007). "Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants". MIT Working Paper 07-31.

[13] Griffith, R., R. Harrison, J. Van Reenen (2006). "How Special is the Special Relationship? Using the Impact of U.S. R&D Spillovers on U.K. Firms as a Test of Technology Sourcing". The American Economic Review, vol. 96, no. 5; pp. 1859-1875.

[14] Griffith, R., S. Lee, J. Van Reenen (2007). "Is Distance Dying at Last? Home Bias in Fixed Effects Models of Patent Citations". CEP Discussion Paper No. 818.

[15] Grossman, G. and E. Helpman (1991). "Trade, Knowledge Spillovers, and Growth". Working Paper No. 3485. National Bureau of Economic Research, Cambridge, Massachusetts.

[16] Helpman, E., M. Melitz, S. Yeaple (2004). "Exports versus FDI with heterogeneous firms". American Economic Review, vol. 94; pp. 300-316.

[17] Jaffe, A., M. Trajtenberg, and M. Fogarty (2000). "The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees", NBER Working Paper No. 7631.

[18] Jaffe, A., M. Trajtenberg, and R. Henderson (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations". Quarterly Journal of Economics, vol. 108, no. 3; pp. 577-598.

[19] Jaffe, A., M. Fogarty, and B. Banks (1998). "Evidence from patents and patent citations on the impact of NASA and other federal labs on commerical innovation". Journal of Industrial Economics, vol. 46, no. 2; pp. 183-205.

[20] Javorcik, B. (2004). "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages". American Economic Review, vol. 94, no. 3; pp. 605-627.

[21] Javorcik, B. and K. Saggi (2003). "Technological Asymmetry among Foreign Investors and Mode of Entry". World Bank Working Paper Series 3196.

[22] Javorcik, B. and M. Sparateanu (2005). "Disentangling FDI Spillovers Effects: What Do Firm Perceptions Tell Us?" In *Does Foreign Direct Investment Promote Development?* by T. Moran, E. Graham, and M. Blomstrom, eds., Institute for International Economics, Washington D.C., 2005

[23] Keller, W. (2002). "Geographic Localization of International Technology Diffusion". The American Economic Review, vol. 92, no. 1; pp.120-142.

[24] Kinoshita, Y. (2001). "R&D and Technology Spillovers via FDI: Innovation and Absorptive Capacity". CEPR Discussion Paper 2775.

[25] Konings, J. (2001). "The Effects of Foreign Direct Investment on Domestic Firms". Economics of Transition, vol. 9, no. 3; pp. 619-633.

[26] Leahy, D. and J. P. Neary (2007). "Absorptive Capacity, R&D Spillovers, and Public Policy". International Journal of Industrial Organization, vol. 25, no. 5; pp. 1089-1108.

[27] Lucas, R. (1988). "On the Mechanics of Economic Development". Journal of Monetary Economics, vol. 22, no. 1; pp. 3-42.

[28] Rodrik, D. (1999). "The New Global Economy and Developing Countries: Making Openness Work". Overseas Development Council (Baltimore, MD) Policy Essay No. 24.

[29] Romer, P. (1986). "Increasing Returns and Long-run Growth". Journal of Political Economy, vol. 94, no. 5; pp. 1002-1037.

[30] Romer, P. (1990). "Endogenous Technological Change". Journal of Political Economy, vol. 98, no. 5; pp. 71-102.

[31] Santos-Silva, J.M.C. and S. Tenreyro (2005). "The Log of Gravity". The Review of Economics and Statistics, vol. 88, no. 4, pp. 641-658.

[32] Singh, J. (2005). "Asymmetry of Knowledge Spillovers between MNCs and Host Country Firms". Unpublished.

[33] Veugelers, R. and B. Cassiman (2004). "Foreign Subsidiaries as a Channel of International Technology Diffusion: Some Direct Firm Level Evidence from Belgium". European Economic Review, 48, pp. 455-476.

[34] World Bank (2008a). "Globalization and Technology Absorption in Europe and Central Asia" by I. Goldberg, L. Branstetter, J. G. Goddard, S. Kuriakose. World Bank Working Paper No. 150.

[35] World Bank (2008b). "Global Economic Prospects: Technology Diffusion in the Developing World".

1.A Appendix: Matching between Amadeus/Orbis and PATSTAT

The appendix has three sections. The first section gives general information on the Amadeus and Orbis databases of financial accounts provided by Bureau Van Dijk (BvD). While Amadeus is limited to Europe, from Orbis we additionally gather information on other parts of the world. Particularly for this paper, we have added information on the United States, Japan, Canada, South Korea, and China, given their important role as foreign investors in Central and Eastern Europe (CEE).

The second section describes the Amadeus/Orbis ownership database. In our context, the ownership structure is crucially to identify foreign firms doing FDI in CEE.

The third section describes the matching process between Amadeus/Orbis and PATSTAT. That is, we explain how we aim to allocate a BvD identification number to each company that applies for a patent.

1.A.1 General information on Amadeus/Orbis

For European countries, we use firm-level data from the Amadeus (Analyse MAjor Databases from EUropean Sources) database, created by Bureau Van Dijk (BvD). This standardised commercial data is collected by about 50 vendors (generally the office of register of companies) across Europe. The database contains financial information on about 8 million firms from 34 countries, including all the European Union countries and Eastern Europe. Additionally, for other major foreign investors in CEE, we obtain the equivalent information from BvD's Orbis dataset, which is the straightforward extension of Amadeus to other parts of the world.[27]

Among the key advantages of Amadeus/Orbis over other data sources are its large

[27] While we obtain Amadeus information from a number of DVDs over time, for Orbis we used the web interface download available to LSE since 2008/2009.

coverage of small and medium sized firms and its unique accounting information on private firms. It covers both listed and unlisted firms of a wide variety of size and age categories, all industries, and ownership types. Coverage varies by country and generally improves over time. The firm and industry coverage of Amadeus is an order of magnitude better compared to other existing firm samples as argued by Gomez-Salvador et al. (2004).

The accounting database includes items from the balance sheet (22 items) and income statement (22 items). No information is available from the cash flow report (i.e., investment or capital expenditure data is not available). The accounting data is harmonized by BvD to enhance comparison across countries. This comparison becomes easier over time due to the improvement in the European Union harmonization of accounting standards. The main descriptive items are country of incorporation, legal form (public vs. private), listing and activity status, date of incorporation, types of accounts (consolidated vs. unconsolidated), product market activity codes (primary and secondary). In addition, for a relatively large number of firms, we observe the number of employees.

The Amadeus information comes from a number of DVDs that we have been collecting. An important feature of Amadeus is the criteria for dropping firms from the database over time. First, the firm's accounts data are followed for up to ten years; each Amadeus DVD contains only the latest 10 years (if available) of financial data. Second, as long as a firm continues to file its financial statements, it continues to appear in Amadeus. In case a firm stops filling its financial statements, it is kept in the database for four extra years. For example, a firm that stops filling its reports in 2003 (i.e., 2002 is the last year for which a financial statement was reported) remains in the database until 2006 included. In 2007 the firm is dropped from the sample (all observations of the specific firm are taken out from the Amadeus database in the 2007 update).

1.A.2 Amadeus/Orbis ownership database

Amadeus/Orbis also provides information on firm ownership. BvD processes the raw data to give information along three dimensions:

(i) Independence indicator: Qualifies the degree of independence of a company with regard to its shareholders.

(ii) Ultimate owner: Identifies the one single firm/person/entity that ultimately owns the firm. Variables available for the ultimate owner include: country of incorporation/origin; ID number (if the ultimate owner is present in the Amadeus/Orbis database); type (e.g., family, industrial firm, employees/managers, financial institution, state). To define an ultimate owner, BvD analyses the shareholding structures of companies that, according to the independence indicator, are not independent from their shareholders.[28] BvD looks for the shareholder with the highest percentage of ownership. If this shareholder is independent, it is recognised as the ultimate owner. If the highest shareholder is not independent, the same process is repeated until the ultimate owner if found.

(iii) Shareholders: Lists shareholders of a given company. In addition to variables available for the ultimate owner, we observe shareholder's direct and total percentage stakes in the firm. Control relationships are followed rather than patrimonial relationships. When there are two categories of shares voting/non-voting, the percentages recorded are those attached to the category voting shares.

A monthly DVD issue of Amadeus contains, for each company, only the last ownership data available with the date (month and year) at which BvD verified this information as valid. In order to construct the panel of ownership, we use 7 different DVD updates from 7 consecutive years to extract ownership data: May 2001 (update 80); May 2002 (92); July 2003 (106); May 2004 (116); October 2005 (133); September 2006 (144); and May 2007 (152). The resulting ownership records dataset spans the

[28]To be independent, the shareholder must be independent by itself (i.e., having one of the following type: Individuals and families, Public authorities, Employees/Managers) or must be an entity with an independence indicator A+, A, or A- (i.e., an entity with no shareholder in control of more than 25%).

period 1995-2007 and gives unique breadth of cross-sectional coverage over time.

1.A.3 The matching process between Amadeus/Orbis and PATSTAT

In this section we describe the matching between each EPO/USPTO patent applicant and an Amadeus/Orbis firm. The match is done by a coincidence both in company name and country of location. We do not consider patent applicants that are individuals or other legal entities like foundations or hospitals, because these will not receive any ID number by the BvD datasets Amadeus and Orbis.

A number of difficulties arise during the process, among which we emphasise:

(i) misspelling of company names (e.g. BAYER versus BAYAR)

(ii) same company name can be written in different ways (e.g. BAYER, BAYER AG, BAYER A.G., BAYER AKTIENGES., BAYER AKTIENGESELLSCHAFT)

(iii) a large number of corporate extensions have to be standardised across countries (e.g. LIMITED, CORPORATION, GMBH, AG, SA, SL)

In order to account for these matters, the company name of each first patent applicant is standardised by using an algorithm in order to come up with two different versions of the cleaned name:[29]

(i) standard name: includes the standardized corporate extension

(ii) stem name: excludes the standardized corporate extension

After the company names in Amadeus/Orbis are standardised in the same way, the matching by cleaned company name and country takes place. For each match, we provide information on whether it was matched by standard name or stem name.[30]

[29] The initial algorithm was available to me through the CEP Productivity and Innovation Group, including the Derwent (2000) industrial standard for converting corporate extensions to standard formats for many different countries, and my task has been to improve this algorithm, especially to increase the coverage of matching of Central and Eastern European countries. This meant standardising company extensions for the Czech Republic, Hungary, or Poland, among other tasks.

[30] Given that the standardisation process cannot correct spelling errors, some minor manual name matching has been done by CEP research assistants. The research assistants were allocated countries based on their language skills.

In a number of cases, multiple ID matches were found for a given patent applicant, identified by its name and country. These cases are resolved by supplementary information (e.g. ownership or address information), where available.

The match between patent applicants and company names in Amadeus/Orbis includes the following countries: all European countries, U.S., Japan, Canada, South Korea, and China. For the paper, unmatched patent applicants will be excluded from the analysis. Consequently, our new population of firms includes only patent applicants that have been allocated an ID.

Therefore, the subsample of firms that we identify as doing FDI activity in CEE are company names with an ID, and that additionally are the Ultimate Owner of a company in the Czech Republic, Hungary, or Poland.

Figure 1.1. Evolution of FDI inflows as a percentage of GDP

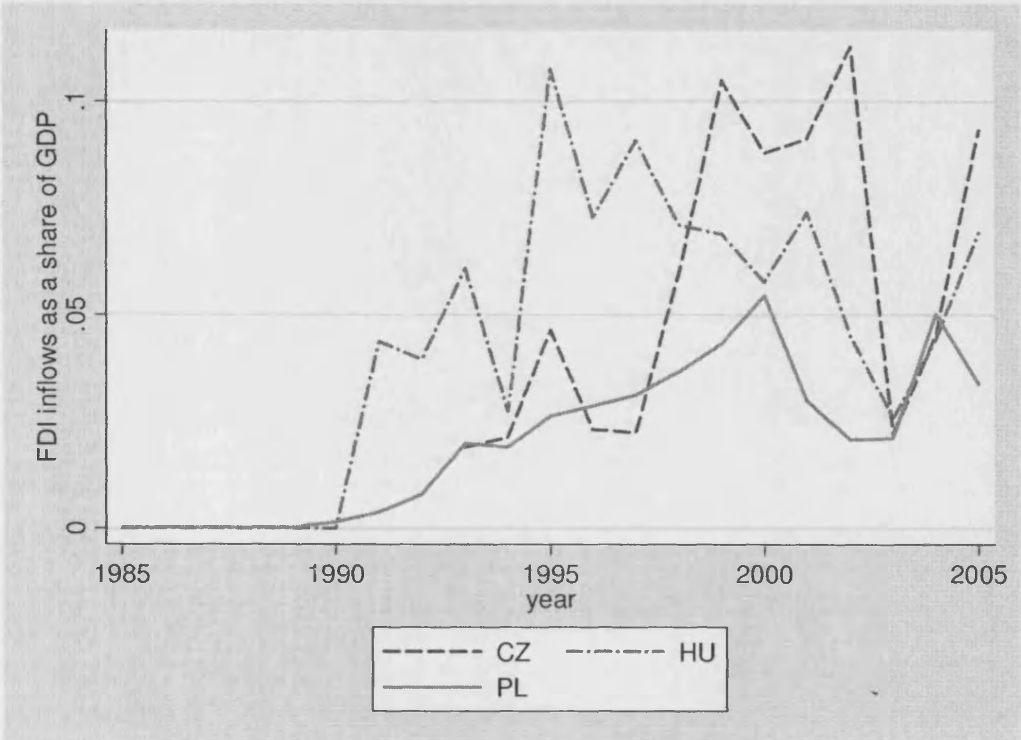

Figure 1.2. Average citations per firm-year

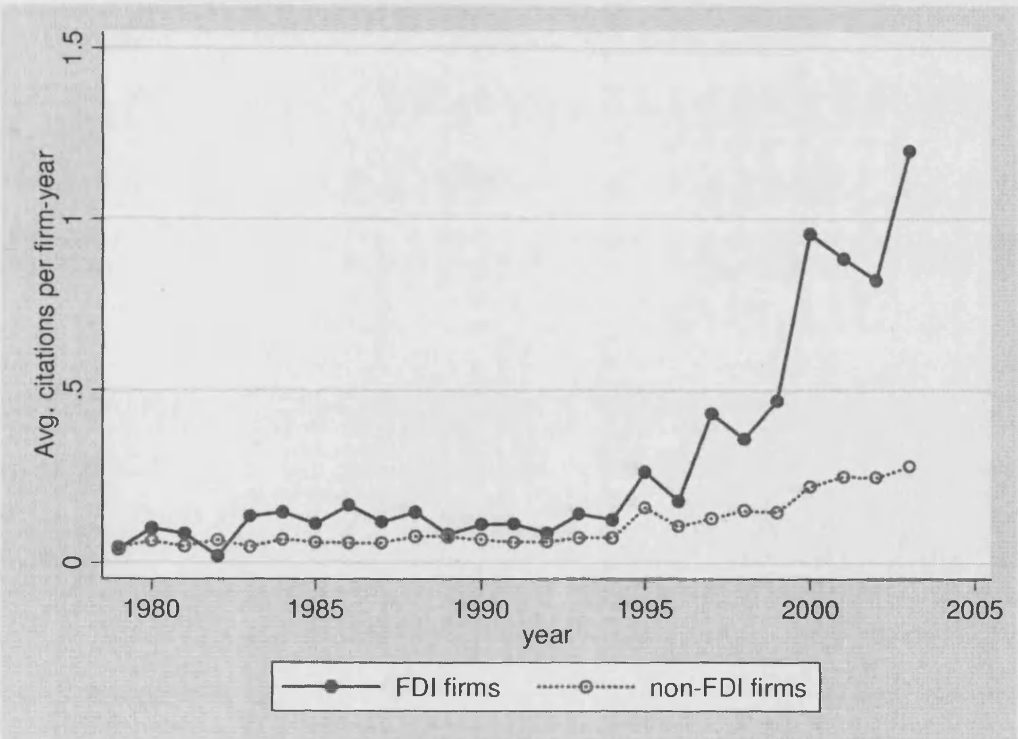

Figure 1.3. Decomposition of CEE citations by nationality of firm

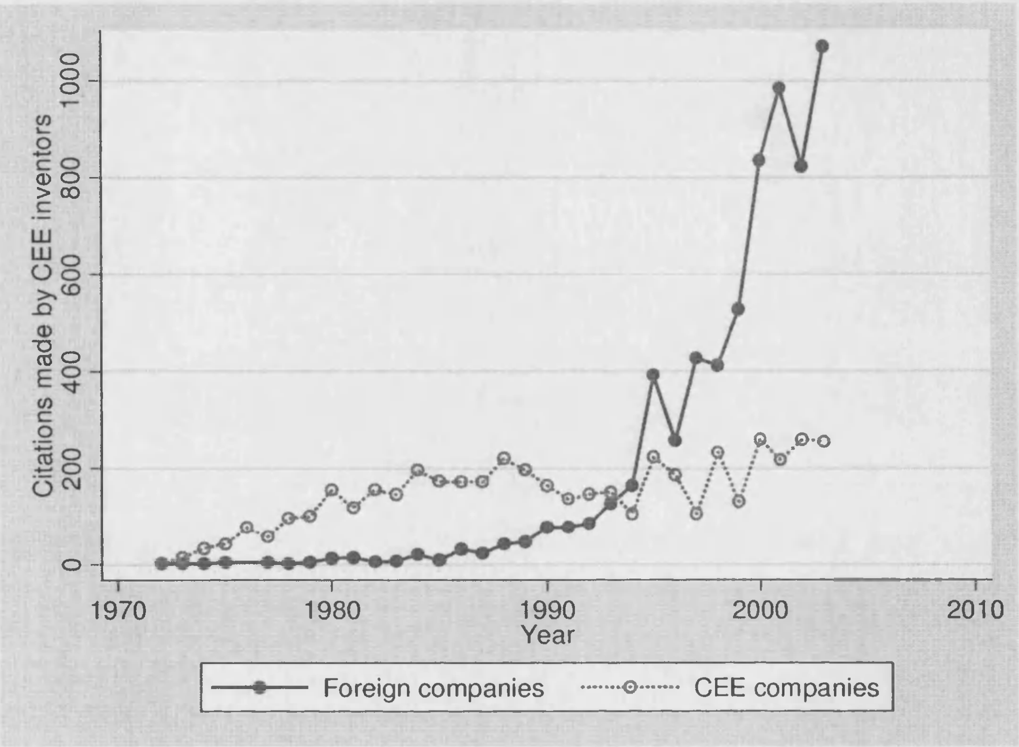

Figure 1.4. Comparison of probability distribution functions (observed vs estimated data)

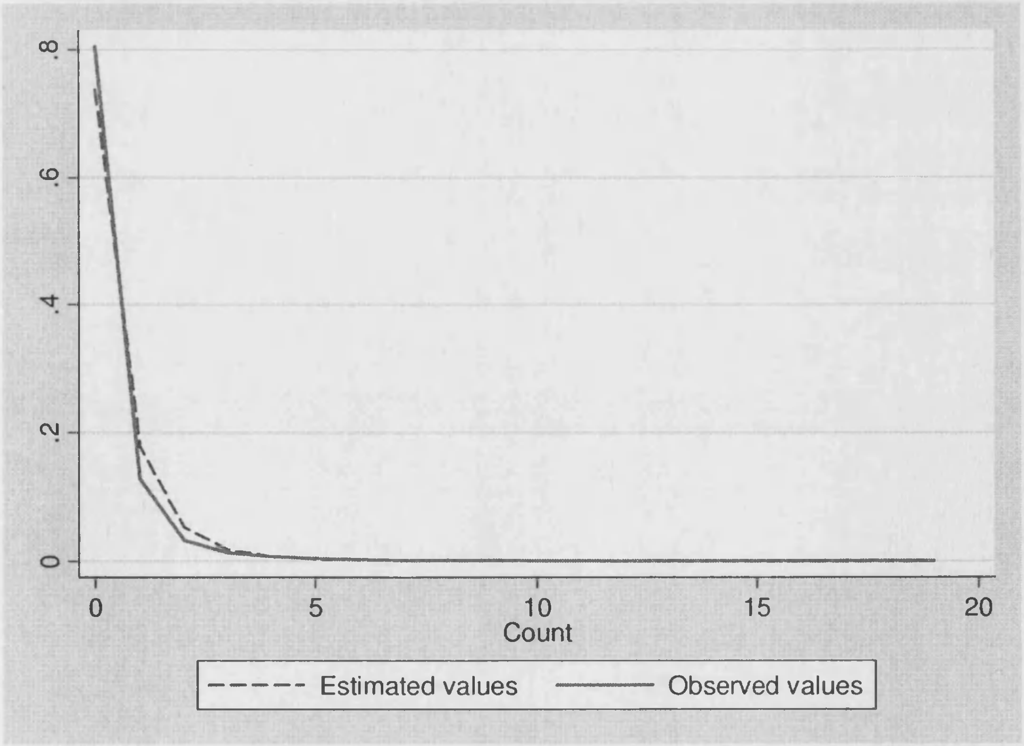

Figure 1.5. Bidders. Average citations per firm-year (raw data)

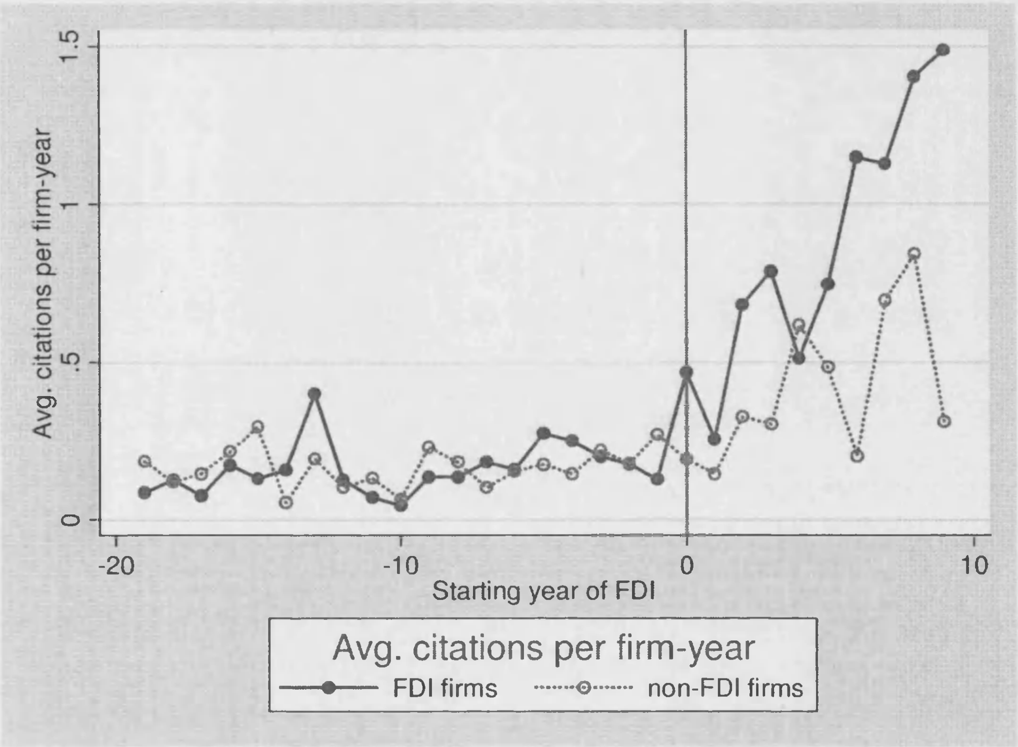

Figure 1.6. Bidders. Average citations per firm-year (with year, host, and source dummies)

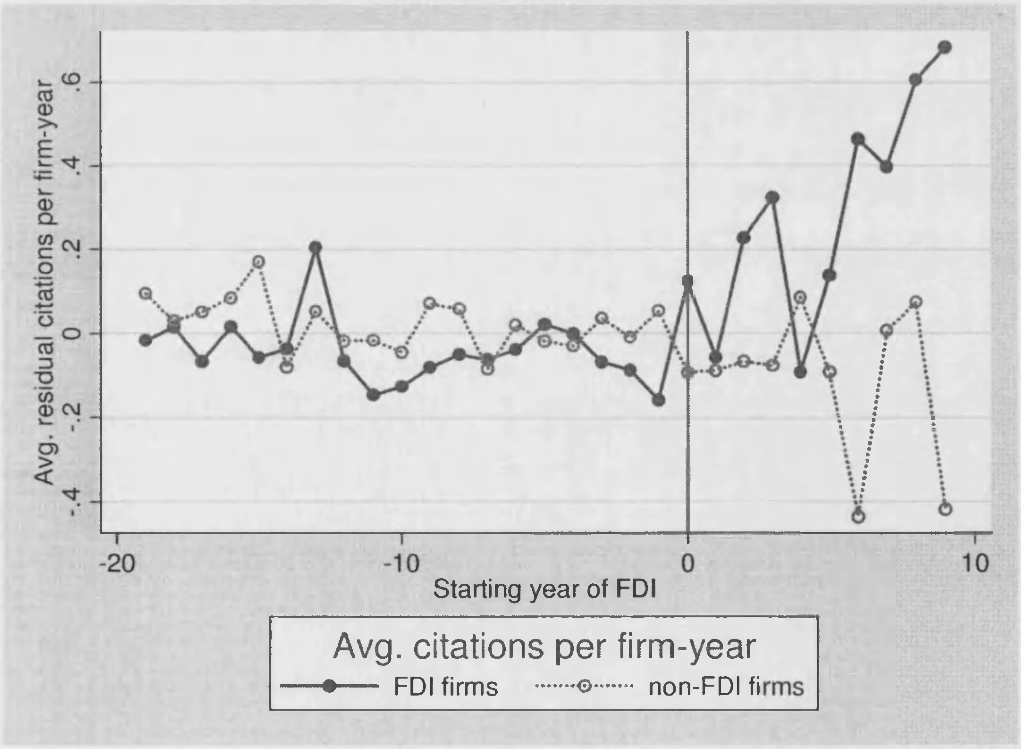

Table 1.1. Descriptive Statistics

A. Summary statistics of aggregate variables

	CZ	HU	PL	Total
Citing firms	658	902	596	2156
Cited firms	1321	1499	1152	3972
...out of which doing FDI:	*97 (7.3%)*	*81 (5.5%)*	*126 (10.9%)*	*304 (7.7%)*
Citing patents	1309	2178	1055	4542
Cited patents	3337	4678	2524	10539
Citations	4134	5418	3612	13164
...out of which received by FDI firms:	*716 (17%)*	*814 (15%)*	*1218 (33%)*	*2748 (21%)*

B. Correlations:

	Stock of patents	Citations per firm-year	Citations per firm-year outside CEE
Stock of patents	1		
Citations per firm-year	0.31	1	
Citations per firm-year outside CEE	0.75	0.39	1

Table 1.2. Decomposition of firms

A. Decomposition of firms by source country

	FDI firms		non-FDI firms	
	Total	%	*Total*	%
Europe	127	42%	979	27%
U.S.	132	43%	2100	57%
Japan	38	13%	503	14%
Rest	7	2%	86	2%
Total	304	100%	3668	100%

B. Decomposition of firms by technology sector (IPC Class):

	FDI firms		non-FDI firms	
	Total	%	*Total*	%
A	43	14%	638	17%
B	60	20%	619	17%
C	65	21%	775	21%
D	3	1%	62	2%
E	0	0%	85	2%
F	44	14%	331	9%
G	37	12%	594	16%
H	52	17%	537	15%
n.a.	0	0%	27	1%
Total	304	100%	3668	100%

Description of IPC codes:

A Human Necessities
B Performing Operations: Transporting
C Chemistry; Metallurgy
D Textiles; Paper
E Fixed Constructions
F Mechanical Engineering; Lighting; Heating
G Physics
H Electricity

Table 1.3. Summary statistics of key variables, divided by FDI and non-FDI firms:

Variable		Mean	S.D.	Min	Median	Max	Observations
Full Sample							
Stock of patents per firm	All firms	746.42	2748.69	0	56	70622	94789
	FDI firms	2948.13	6298.88	0	757	70622	8382
	non-FDI firms	532.84	1980.89	0	42	70622	86407
Citations per firm-year	All firms	0.14	0.83	0	0	82	94789
	FDI firms	0.33	1.78	0	0	82	8382
	non-FDI firms	0.12	0.67	0	0	56	86407
CZ							
Stock of patents per firm	All firms	762.88	2776.02	0	64	70622	31050
	FDI firms	2381.52	4938.85	0	635	51038	2667
	non-FDI firms	610.78	2422.78	0	49	70622	28383
Citations per firm-year	All firms	0.13	0.76	0	0	50	31050
	FDI firms	0.27	0.99	0	0	16	2667
	non-FDI firms	0.12	0.73	0	0	50	28383
HU							
Stock of patents per firm	All firms	662.77	2552.89	0	47	70622	36832
	FDI firms	2974.02	6870.32	0	669	70622	2254
	non-FDI firms	512.11	1869.65	0	40	51957	34578
Citations per firm-year	All firms	0.15	0.76	0	0	56	36832
	FDI firms	0.36	1.29	0	0	25	2254
	non-FDI firms	0.13	0.71	0	0	56	34578
PL							
Stock of patents per firm	All firms	841.93	2963.69	0	60	70622	26907
	FDI firms	3367.89	6793.45	0	914	70622	3461
	non-FDI firms	469.06	1479.26	0	39	31935	23446
Citations per firm-year	All firms	0.13	0.99	0	0	82	26907
	FDI firms	0.35	2.42	0	0	82	3461
	non-FDI firms	0.1	0.52	0	0	17	23446

Table 1.4. OLS. Full sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy for post-FDI years	**0.108**	**0.087**	**0.098**	**0.105**	**0.099**	**0.079**	**0.090**	**0.099**
	[0.016]* **	**[0.016]*** **	**[0.015]*** **	**[0.016]*** **	**[0.016]*** **	**[0.016]*** **	**[0.015]*** **	**[0.015]*** **
Dummy for firms doing FDI	0.033	-0.017	-0.028		0.040	-0.010	-0.020	
	[0.010]***	[0.009]*	[0.008]***		[0.010]***	[0.009]	[0.008]**	
ln(Stock of Patents)		-0.008	-0.007	0.009		-0.005	-0.003	0.010
		[0.001]***	[0.001]***	[0.002]***		[0.001]***	[0.001]**	[0.002]***
ln(Citations received by non-CEE)		0.035	0.034	0.031		0.031	0.029	0.025
		[0.002]***	[0.002]***	[0.002]***		[0.002]***	[0.001]***	[0.002]***
Observations	84719	84718	84718	84718	84719	84718	84718	84718
R-squared	0.04	0.08		0.05	0.06	0.09		0.07
Firm-level RE/FE			RE	FE			RE	FE
Year FE	Y	Y	Y	Y				
Host Country FE	Y	Y	Y					
Source Country FE	Y	Y	Y					
Technology Class FE	Y	Y	Y					
Year*Host Country FE					Y	Y	Y	Y
Year*Source Country FE					Y	Y	Y	Y
Year*Technology Class FE					Y	Y	Y	Y

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level; for regressions with firm RE and FE, std. errors are robust to heteroskedasticity and serial correlation.

Table 1.5. OLS. Full sample. Includes (Stock of years doing FDI)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy for post-FDI years	0.067	0.048	0.055	0.057	0.066	0.046	0.054	0.058
	[0.019]***	[0.019]**	[0.018]***	[0.018]***	[0.019]***	[0.018]**	[0.017]***	[0.017]***
Dummy for firms doing FDI	0.032	-0.017	-0.029		0.039	-0.010	-0.020	
	[0.010]***	[0.009]*	[0.008]***		[0.010]***	[0.009]	[0.008]**	
Stock of years doing FDI	**0.006**	**0.005**	**0.006**	**0.007**	**0.004**	**0.004**	**0.005**	**0.006**
	[0.002]**	**[0.002]****	**[0.002]*****	**[0.002]*****	**[0.002]***	**[0.002]***	**[0.002]****	**[0.002]*****
ln(Stock of Patents)		-0.009	-0.008	0.008		-0.005	-0.003	0.009
		[0.001]***	[0.001]***	[0.002]***		[0.002]***	[0.001]**	[0.002]***
ln(Citations received by non-CEE)		0.035	0.034	0.031		0.031	0.029	0.026
		[0.002]***	[0.002]***	[0.002]***		[0.002]***	[0.001]***	[0.002]***
Observations	84719	84718	84718	84718	84719	84718	84718	84718
R-squared	0.04	0.08		0.05	0.06	0.10		0.07
Firm-level RE/FE			RE	FE			RE	FE
Year FE	Y	Y	Y	Y				
Host Country FE	Y	Y	Y					
Source Country FE	Y	Y	Y					
Technology Class FE	Y	Y	Y					
Year*Host Country FE					Y	Y	Y	Y
Year*Source Country FE					Y	Y	Y	Y
Year*Technology Class FE					Y	Y	Y	Y

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level; for regressions with firm RE and FE, std. errors are robust to heteroskedasticity and serial correlation.

Table 1.6. Poisson. Full Sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy for post-FDI years	**0.715**	**0.494**	**0.562**	**0.440**	**0.518**	**0.255**	**0.259**	**0.182**
	[0.193]*	**[0.170]***	**[0.156]***	**[0.147]***	**[0.191]***	**[0.157]**	**[0.111]**	**[0.108]***
Dummy for firms doing FDI	0.559	-0.224	-0.440		0.728	-0.012	-0.129	
	[0.146]***	[0.131]*	[0.130]***		[0.143]***	[0.121]	[0.099]	
ln(Stock of Patents)		-0.145	-0.236	0.664		-0.082	-0.163	0.753
		[0.035]***	[0.030]***	[0.075]***		[0.035]**	[0.029]***	[0.075]***
ln(Citations received by non-CEE)		0.569	0.614	0.592		0.498	0.520	0.490
		[0.039]***	[0.034]***	[0.048]***		[0.036]***	[0.031]***	[0.046]***
Observations	84719	84718	84718	85282	84719	84718	84718	84715
R-squared								
Firm RE/FE			RE	FE			RE	FE
Year FE	Y	Y	Y	Y				
Host Country FE	Y	Y	Y					
Source Country FE	Y	Y	Y					
Technology Class FE	Y	Y	Y					
Year*Host Country FE					Y	Y	Y	Y
Year*Source Country FE					Y	Y	Y	Y
Year*Technology Class FE					Y	Y	Y	Y

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level; for regressions with firm RE and FE, std. errors are robust to heteroskedasticity and serial correlation.

Table 1.7. Poisson. Full sample. Includes (Stock of years doing FDI)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy for post-FDI years	0.385	0.259	0.226	0.100	0.247	0.074	0.009	-0.117
	[0.218]*	[0.189]	[0.152]	[0.151]	[0.212]	[0.177]	[0.121]	[0.124]
Dummy for firms doing FDI	0.551	-0.229	-0.462		0.722	-0.016	-0.148	
	[0.147]***	[0.131]*	[0.132]***		[0.143]***	[0.122]	[0.099]	
Stock of years doing FDI	**0.043**	**0.030**	**0.047**	**0.052**	**0.035**	**0.023**	**0.035**	**0.046**
	[0.018]**	**[0.014]****	**[0.015]*****	**[0.018]*****	**[0.019]***	**[0.015]**	**[0.014]****	**[0.015]*****
ln(Stock of Patents)		-0.147	-0.241	0.644		-0.084	-0.166	0.740
		[0.035]***	[0.031]***	[0.075]***		[0.035]**	[0.029]***	[0.074]***
ln(Citations received by non-CEE)		0.570	0.617	0.600		0.499	0.523	0.497
		[0.039]***	[0.034]***	[0.049]***		[0.037]***	[0.031]***	[0.046]***
Observations	84719	84718	84718	85282	84719	84718	84718	84715
R-squared								
Firm-level RE/FE			RE	FE			RE	FE
Year FE	Y	Y	Y	Y				
Host Country FE	Y	Y	Y					
Source Country FE	Y	Y	Y					
Technology Class FE	Y	Y	Y					
Year*Host Country FE					Y	Y	Y	Y
Year*Source Country FE					Y	Y	Y	Y
Year*Technology Class FE					Y	Y	Y	Y

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level; for regressions with firm RE and FE, std. errors are robust to heteroskedasticity and serial correlation.

Table 1.8. The Bidders' Sample

Number of Companies:

Winning bidders:	47
Losing bidders:	89

Decomposition by FDI starting year:

1990	2
1991	12
1992	6
1993	9
1994	1
1995	7
1996	4
1997	1
1998	3
1999	2

Decomposition of winners and losers by home country:

Winners:

AT	1
CH	2
DE	9
FR	6
GB	3
HU	3
KR	2
NL	2
SE	2
US	17
Total:	47

Losers:

ARG	1
AT	2
CA	3
CH	4
DE	16
DK	1
FI	2
FR	19
GB	8
IT	4
JP	3
KR	1
NL	1
SE	2
US	22
Total:	89

Table 1.9. The Bidders' Characteristics by Winner Status

	Winners (1)	Losers (2)	(1)-(2) t-stat [p-value]
(1) Levels of Variables:			
Total citations received	11569.6	2748.3	4.3
	(18058.9)	(4901.4)	[0.00]***
Total citations received from CEE	3.1	2.5	0.4
	(6.1)	(7.5)	[0.71]
(2) Growth rates of Variables:			
Growth rate in total citations before FDI bidding year	15.1	10.2	0.81
	(29.3)	(32.2)	[0.42]
Growth rate in total citations after FDI bidding year	3.9	8.2	1.43
	(11.4)	(24.5)	[0.16]
Growth rate in CEE citations before FDI bidding year	0.5	0.3	1.35
	(1.1)	(1.1)	[0.18]
Growth rate in CEE citations after FDI bidding year	3.4	0.6	3.35
	(3.4)	(1.4)	[0.00]***

Table 1.10. OLS without case dummies.

	(1)	(2)	(3)	(4)	(5)	(6)
D(FDI firms)	-0.060	-0.055	-0.060	-0.067	-0.059	-0.068
	[0.030]**	[0.032]*	[0.030]**	[0.033]**	[0.033]*	[0.033]**
D(post-FDI years)	0.016		0.069	-0.050		0.009
	[0.026]		[0.036]*	[0.042]		[0.039]
D(FDI firms)*D(post-FDI years)	**0.195**		**0.043**	**0.191**		**0.045**
	[0.058]* **		**[0.067]**	**[0.062]*** **		**[0.070]**
Stock of years doing FDI		-0.001	-0.007		-0.009	-0.009
		[0.002]	[0.003]**		[0.009]	[0.008]
D(FDI firms)*Stock of years doing FDI		**0.022**	**0.019**		**0.022**	**0.018**
		[0.006]* **	**[0.007]*** **		**[0.006]*** **	**[0.007]** **
ln(Stock of Patents)	0.054	0.055	0.054	0.061	0.062	0.061
	[0.008]***	[0.008]***	[0.008]***	[0.010]***	[0.010]***	[0.010]***
Observations	4098	4098	4098	4098	4098	4098
R-squared	0.16	0.16	0.16	0.21	0.22	0.22
Year FE				Y	Y	Y
Host Country FE				Y	Y	Y
Source Country FE				Y	Y	Y
Case dummies						
Firm dummies						

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level.

Table 1.11. OLS with case dummies.

	(1)	(2)	(3)	(4)	(5)	(6)
D(FDI firms)	-0.034	-0.030	-0.033	-0.048	-0.041	-0.049
	[0.029]	[0.027]	[0.029]	[0.030]	[0.027]	[0.031]
D(post-FDI years)	0.034		0.073	-0.030		0.019
	[0.026]		[0.038]*	[0.037]		[0.036]
D(FDI firms)*D(post-FDI years)	**0.180**		**0.039**	**0.185**		**0.045**
	[0.062]* **		**[0.070]**	**[0.065]*** **		**[0.071]**
Stock of years doing FDI		0.001	-0.005		-0.002	-0.002
		[0.002]	[0.003]*		[0.009]	[0.008]
D(FDI firms)*Stock of years doing FDI		**0.021**	**0.018**		**0.021**	**0.018**
		[0.006]* **	**[0.007]** **		**[0.006]*** **	**[0.007]** **
ln(Stock of Patents)	0.048	0.050	0.048	0.050	0.051	0.050
	[0.007]***	[0.007]***	[0.007]***	[0.008]***	[0.008]***	[0.008]***
Observations	4098	4098	4098	4098	4098	4098
R-squared	0.24	0.24	0.24	0.28	0.28	0.28
Year FE				Y	Y	Y
Host Country FE				Y	Y	Y
Source Country FE				Y	Y	Y
Case dummies	Y	Y	Y	Y	Y	Y
Firm dummies						

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level.

Table 1.12. OLS with firm dummies.

	(1)	(2)	(3)	(4)	(5)	(6)
D(FDI firms)	-0.023	-0.018	-0.024	-0.019	-0.012	-0.020
	[0.045]	[0.045]	[0.045]	[0.038]	[0.034]	[0.038]
D(post-FDI years)	0.017		0.066	-0.044		0.010
	[0.032]		[0.042]	[0.038]		[0.038]
D(FDI firms)*D(post-FDI years)	**0.196**		**0.040**	**0.194**		**0.043**
	[0.065]*		**[0.070]**	**[0.066]***		**[0.072]**
Stock of years doing FDI		-0.003	-0.007		-0.008	-0.007
		[0.002]	[0.003]**		[0.009]	[0.009]
D(FDI firms)*Stock of years doing FDI		**0.023**	**0.020**		**0.023**	**0.019**
		[0.006]*	**[0.007]***		**[0.006]***	**[0.007]***
ln(Stock of Patents)	0.053	0.063	0.054	0.047	0.049	0.049
	[0.010]***	[0.011]***	[0.010]***	[0.014]***	[0.014]***	[0.013]***
Observations	4098	4098	4098	4098	4098	4098
R-squared	0.28	0.28	0.28	0.31	0.32	0.32
Year FE				Y	Y	Y
Host Country FE				Y	Y	Y
Source Country FE				Y	Y	Y
Case dummies						
Firm dummies	Y	Y	Y	Y	Y	Y

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level.

Table 1.13. Poisson without case dummies.

	(1)	(2)	(3)	(4)	(5)	(6)
D(FDI firms)	-0.777	-0.590	-0.779	-0.472	-0.403	-0.479
	[0.371]**	[0.318]*	[0.373]**	[0.328]	[0.300]	[0.322]
D(post-FDI years)	0.082		0.509	-0.334		0.029
	[0.255]		[0.263]*	[0.337]		[0.313]
D(FDI firms)*D(post-FDI years)	**0.954**		**0.379**	**0.747**		**0.199**
	[0.334]* **		**[0.421]**	**[0.280]*** **		**[0.459]**
Stock of years doing FDI		-0.021	-0.053		-0.062	-0.058
		[0.018]	[0.017]***		[0.059]	[0.050]
D(FDI firms)*Stock of years doing FDI		**0.079**	**0.070**		**0.080**	**0.068**
		[0.020]* **	**[0.027]*** **		**[0.023]*** **	**[0.037]***
ln(Stock of Patents)	0.580	0.605	0.581	0.713	0.706	0.705
	[0.060]***	[0.065]***	[0.064]***	[0.062]***	[0.061]***	[0.061]***
Observations	4098	4098	4098	4098	4098	4098
R-squared	0.18	0.18	0.19	0.29	0.29	0.30
Year FE				Y	Y	Y
Host Country FE				Y	Y	Y
Source Country FE				Y	Y	Y
Case dummies						
Firm dummies						

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level.

Table 1.14. Poisson with case dummies.

	(1)	(2)	(3)	(4)	(5)	(6)
D(FDI firms)	0.080	0.233	0.072	0.183	0.171	0.189
	[0.238]	[0.209]	[0.239]	[0.281]	[0.258]	[0.282]
D(post-FDI years)	0.235		0.542	-0.173		0.268
	[0.217]		[0.251]**	[0.309]		[0.322]
D(FDI firms)*D(post-FDI years)	**0.654**		**0.210**	**0.574**		**-0.069**
	[0.246]* **		**[0.356]**	**[0.265]** **		**[0.372]**
Stock of years doing FDI		-0.006	-0.040		0.100	0.101
		[0.018]	[0.021]*		[0.065]	[0.064]
D(FDI firms)*Stock of years doing FDI		**0.055**	**0.055**		**0.066**	**0.071**
		[0.018]* **	**[0.029]* **		**[0.027]** **	**[0.038]* **
ln(Stock of Patents)	0.589	0.632	0.589	0.738	0.735	0.733
	[0.068]***	[0.076]***	[0.072]***	[0.091]***	[0.089]***	[0.089]***
Observations	4098	4098	4098	4098	4098	4098
R-squared	0.29	0.29	0.30	0.41	0.41	0.41
Year FE				Y	Y	Y
Host Country FE				Y	Y	Y
Source Country FE				Y	Y	Y
Case dummies	Y	Y	Y	Y	Y	Y
Firm dummies						

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level.

Table 1.15. Poisson with firm dummies.

	(1)	(2)	(3)	(4)	(5)	(6)
D(FDI firms)	0.172	0.190	0.082	0.303	0.296	0.343
	[0.467]	[0.493]	[0.454]	[0.354]	[0.379]	[0.372]
D(post-FDI years)	-0.045		0.351	-0.088		0.187
	[0.204]		[0.206]*	[0.302]		[0.295]
D(FDI firms)*D(post-FDI years)	**0.503**		**0.080**	**0.392**		**-0.126**
	[0.254]**		**[0.327]**	**[0.284]**		**[0.373]**
Stock of years doing FDI		-0.053	-0.067		0.036	0.030
		[0.027]**	[0.027]**		[0.057]	[0.056]
D(FDI firms)*Stock of years doing FDI		**0.047**	**0.052**		**0.059**	**0.068**
		[0.022]**	**[0.028]***		**[0.030]****	**[0.040]***
ln(Stock of Patents)	0.910	1.145	1.037	1.278	1.211	1.210
	[0.099]***	[0.188]***	[0.178]***	[0.234]***	[0.215]***	[0.215]***
Observations	4098	4098	4098	4098	4098	4098
R-squared	0.29	0.29	0.30	0.41	0.41	0.41
Year FE				Y	Y	Y
Host Country FE				Y	Y	Y
Source Country FE				Y	Y	Y
Case dummies						
Firm dummies	Y	Y	Y	Y	Y	Y

Note: The dependent variable is yearly citations received by each firm for the sample period 1978-2007. The statistical significance is as follows: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the firm level.

Chapter 2

Relative Factor Endowments and International Portfolio Choice

2.1 Introduction

This paper presents a model of international portfolio choice based on cross-country differences in relative factor abundance. Countries have varying degrees of similarity in their factor endowment ratios, and are subject to aggregate productivity (country-specific) shocks. Risk averse consumers can insure against these shocks by investing their wealth at home and abroad. In a many-good setup, the change in factor prices after a positive shock in a particular country provides insurance to countries that have dissimilar factor endowment ratios, but is bad news for countries with similar factor endowment ratios, since their incomes will worsen. A positive productivity shock in a capital-abundant country, for example, will raise wage rates and reduce the return to capital, thus raising the incomes of labour-abundant countries and harming the incomes of other capital-abundant countries. Therefore countries with similar relative factor endowments have got a stronger incentive to invest in one another for insurance purposes than countries with dissimilar endowments.

Since our theoretical mechanism works through the effects of shocks on prices, the size of the country suffering the shock (and selling assets) is obviously a relevant consideration. In a generalization of our model, we study how endowment similarity interacts with country size. We show that under standard assumptions a country invests relatively *more* in a large-similar country than in a small-similar country, and relatively *less* in a large-dissimilar country than in a small-dissimilar country.

We first frame this intuition within a complete asset markets setup, in which countries trade Arrow-Debreu securities prior to the realization of uncertainty. However, our results do not hinge on many of the strong assumptions (Arrow-Debreu securities, complete asset markets, absence of home bias in portfolios) we make for tractability purposes. When we replace the Arrow-Debreu setup with a more 'realistic' financial side, the model yields predictions similar to those of our stylised model: we assume that countries can exchange claims on their GDPs before uncertainty is realised, and that investing abroad is subject to frictions that reduce the expected return of foreign assets. This obviously generates a home bias in the portfolios of countries. By the same line of reasoning as above, investing in countries with similar factor endowment ratios provides better insurance to a country with a home-biased portfolio.

Empirical evidence linking bilateral international portfolio investment positions to a proxy for relative factor endowment similarity supports our theory: after controlling for commodity and asset market frictions, the similarity of host and source countries in their relative capital-labour ratios is estimated to have a positive effect on the source country's investment position in the host country. The magnitude of this effect depends on the host country's GDP size, as larger countries have a stronger effect on prices.

Figures (2.1) and (2.2) show the relationship between bilateral equity investment (vertical axis) and similarity in factor endowments (horizontal axis), after controlling for country fixed effects and a set of standard controls.[1] Figure (2.1) illustrates the case

[1] Small countries are those with GDP below the median of our sample. The vertical axis is the residual of an OLS regression of bilateral equity investment on source- and host- country fixed effects, log of

of small host countries: the similarity in factor endowments between the source and the host country does not seem to be a determinant of equity investment. The coefficient describing how changes in similarity affect equity investment is statistically insignificant and close to zero. In other words, for countries that cannot significantly change world relative prices, it does not seem to be relevant whether investor and recipient countries are similar or not. Figure (2.2) illustrates the case for large host countries. We obtain a positive and statistically significant coefficient: conditional on the host country being relatively large, increases in factor endowment similarity lead to greater equity investment positions of the source country into the host country.

The idea that relative price changes may act as an insurance mechanism can be traced back to Cole and Obstfeld (1991), who argued this might explain the lack of international diversification of country portfolios. In their model, two completely specialised countries trade with each other in assets and outputs. But asset trade is almost redundant, as changes in the terms of trade after a shock act as insurance. By allowing for many countries with varying degrees of factor endowment similarity, we turn this intuition into a theory of international portfolio choice. Unlike Cole and Obstfeld (1991), however, the emphasis of our model is not on the terms of trade, but on factor prices. Think of the standard indeterminacy problem of the production structures of countries in a Heckscher-Ohlin model with more goods than production factors. In that environment, it is impossible to talk about the terms of trade of countries, as the latter depend on the countries' production structures. But the model does have instead unambiguous predictions about the behavior of factor prices, as these do not depend on production structures.

On the empirical side, an additional problem associated to terms of trade movements

bilateral trade, log of distance, common legal origin, common border, common language, colony dummy, regional trade agreement, correlation in GDP per capita, and a number of controls on equity markets and informational frictions. (See the controls used in Table 4, column (12).) The horizontal axis is the residual of an OLS regression of similarity in factor endowments (as we define in section 3 below) on the same set of control variables.

is the difficulty to isolate the extent to which the comovement of the terms of trade of two countries is due to their similarity in factor endowment ratios from, among others, other common sources of comparative advantage; the occurrence of sectoral shocks; the correlation of aggregate productivity shocks; or exchange-rate movements. Working with proxies for factor endowment similarity, as we do, has the advantage to bypass these thorny issues, as it relates our endogenous variable, the portfolio positions of countries, to its 'primitive' determinant (according to our theory).

Our model consists of endowment economies as in Lucas (1982) and Svensson (1988). We allow countries to differ in their patterns of specialization according to their relative factor endowments, in a manner similar to Helpman and Razin (1978) and Helpman (1988). In comparison with these references, however, we only allow for country-specific aggregate productivity shocks in our analysis.

Our work adds to a growing body of research that attempts to explain the international portfolio choices of countries. Obstfeld and Rogoff (2001) and Lane and Milesi-Ferretti (2008) have put emphasis on commodity trade costs; Martin and Rey (2000) and (2004) have focused on the role of size; and Portes and Rey (2005) have highlighted the importance of informational costs for investment flows.[2] In comparison with these references, our paper highlights that bilateral portfolio positions not only depend on frictions between countries, but also on other country-pair specific characteristics: in our theory, even when bilateral frictions (and productivity shocks) are equally correlated across all country-pairs is it possible that a country finds it optimal not to invest the same amount across countries.

Finally, the causal direction from asset trade to production specialization has been addressed by Kalemli-Ozcan *et al.* (2003) and Koren (2005). Both argue that international asset market integration favours specialisation in production, as it enables countries to insure against sector-specific productivity shocks. Our paper complements this

[2]For other approaches to international portfolio choice, see also Kraay and Ventura (2000) and (2002), Kraay *et al.* (2005), and Lane and Milesi-Ferretti (2001) and (2002).

literature by pointing that causality might also run in the opposite direction: the real side of the economy also determines the international portfolio positions of countries.

The rest of the paper is structured as follows: Section 2.2 discusses a stylised model linking production specialisation and international portfolio choice, and lays the ground for the empirical analysis. Section 2.3 discusses our empirical strategy, while Section 2.4 discusses empirical evidence supportive of the model. Some concluding remarks follow. Finally, the appendix provides proofs and extensions of the model discussed in Section 2.2.

2.2 The Model

Let us denote countries with $j \in J$. Abusing notation, we will also use J to denote the number of countries. Each country has got a representative consumer, who maximises expected utility $E[U(C_j)]$. $E(\cdot)$ is the expectations operator, and $U(\cdot)$ is the utility function, which we assume concave: $U'(\cdot) > 0$, $U''(\cdot) < 0$. C denotes consumption of a freely traded final composite good $C_j = C_{1j}^{\frac{1}{2}} C_{2j}^{\frac{1}{2}}$, where C_{ij} denotes country j's consumption of freely traded intermediate good i, $i = 1, 2$.[3] Preferences are identical across countries.

Technologies in the intermediate good industries are also identical across countries. We simplify by assuming linear production functions:[4] $y_{1j} = A_j K_j$ and $y_{2j} = A_j L_j$, where y_{ij} denotes production of good i in country j, and $A_j > 0$ denotes country j's aggregate productivity level. We can think of $A_j K_j$ and $A_j L_j$ as production factors (capital and labour) measured in efficiency units. We assume perfect competition.

Each country has got exogenously given endowments of the two production factors,

[3] We constrain the elasticity of substitution to be equal or larger than one, so as to avoid 'immiserizing growth' issues. The Cobb-Douglas assumption is made here for tractability. As we discuss below, most of our results do not depend on it.

[4] In Appendix B we show that this assumption is harmless: a model with neoclassical production functions yields similar insights.

which are internationally immobile and supplied inelastically. We distinguish two subsets of countries, which we denote with k and l: $J_k \cup J_l = J$, $J_k \cap J_l = \varnothing$.[5] For all $k \in J_k, l \in J_l$,

$$\begin{aligned} K_k &= \phi_k \left(1/2 + \mu\right), && (2.1) \\ L_k &= \phi_k \left(1/2 - \mu\right), && (2.2) \\ K_l &= \phi_l \left(1/2 - \mu\right), && (2.3) \\ L_l &= \phi_l \left(1/2 + \mu\right). && (2.4) \end{aligned}$$

$\mu \in [0, 1/2]$. Notice this implies countries in J_k have their production structures biased towards good 1 relative to countries in J_l. For the sake of simplicity, we assume an equal number of countries of each type: $J_k = J_l = J/2$. The parameter $\phi_j > 0$ is a scaling factor that allows for cross-country differences in size. We assume that the distributions of this scaling factor within J_k and J_l are symmetric.

A_j is *ex-ante* uncertain. We assume there are J states of nature (denoted by s, $s = 1, ..., J$), each with identical probability $\pi(s) = 1/J$. States of nature are characterised by productivity level vectors $A(s) = [A_1(s), A_2(s), ..., A_J(s)]$. In particular,

$$\begin{aligned} A(1) &= (1 + a, 1, ..., 1), \\ A(2) &= (1, 1 + a, ..., 1), \\ &\;\;\;\cdots \\ A(J) &= (1, ..., 1, 1 + a), \end{aligned}$$

where $a > 0$ is a constant.[6]

There is a world market in which agents can buy or sell Arrow-Debreu contingent

[5]To avoid confusion, we will spare the indices j and j' for when we refer to any country in J; we will use k and k' to refer to countries in J_k; and l and l' to refer to countries in J_l.

[6]This is similar to what Acemoglu and Zilibotti (1997) and Martin and Rey (2000, 2004) assume in differents contexts.

claims before uncertainty is realized. These claims have payoffs that depend on the state of nature: the owner (seller) of the security receives (pays) worth one unit of the final good if state s occurs, but nothing in any other state. We assume asset-market completeness.

2.2.1 Goods Market Equilibrium

Given the homotheticity of $C(\cdot)$, relative demands depend only on relative prices. Goods market equilibrium is therefore determined by

$$\frac{y_{1W}}{y_{2W}} = \frac{C_{1W}}{C_{2W}} = \frac{C_{1j}}{C_{2j}} = \frac{p_2}{p_1} = \frac{w}{r}, \tag{2.5}$$

where $C_{iW} \equiv \sum_{j\in J} C_{ij}$ and $y_{iW} \equiv \sum_{j\in J} y_{ij}$. Notice that p_i is also the price of the factor used in industry i when factors are measured in efficiency units. This can be seen from the equilibrium pricing conditions: $p_1 = r$ and $p_2 = w$, where r and w denote, respectively, the price of factor AK and factor AL. Taking the final good as the numeraire,

$$r = \frac{1}{2}\left(\frac{L_W}{K_W}\right)^{\frac{1}{2}}, \tag{2.6}$$

$$w = \frac{1}{2}\left(\frac{K_W}{L_W}\right)^{\frac{1}{2}}, \tag{2.7}$$

where $K_W \equiv \sum_{j\in J} A_j K_j$ and $L_W \equiv \sum_{j\in J} A_j L_j$. Obviously, $w/r = K_W/L_W$. Notice that free trade and the pricing conditions imply factor price equalization across countries, as in Trefler (1993) and Ventura (1997).

2.2.2 Asset Market Equilibrium

Let $B_j(s)$ denote country j's net purchase of state-s Arrow-Debreu securities. Let $p(s)$ denote the price of one such security. Each country's utility maximization problem can be expressed as

$$\max_{\{B_j(s)\}_{s=1}^{J}} \quad \sum_s \pi(s)\, U\left[Y_j(s) + B_j(s)\right], \tag{2.8}$$

subject to budget constraints $\sum_s p(s)\, B_j(s) = 0$ and $C_j(s) = Y_j(s) + B_j(s)$. Manipulating the first order conditions for states s and s',

$$\frac{\pi(s)\, U'\left[C_j(s)\right]}{\pi(s')\, U'\left[C_j(s')\right]} = \frac{p(s)}{p(s')}. \tag{2.9}$$

Market clearing requires $\sum_j B_j(s) = 0$ and $Y_W(s) = \sum_j C_j(s)$ for all s, where Y_W denotes world production of the final good. Finally, we close the model with the no-arbitrage condition $\sum_s p(s) = 1$.

Under log-utility ($U(C) = \ln(C)$), for example, the model yields the following equilibrium asset prices and portfolio choices:

$$p(s) = \frac{\left[Y_W(s)\right]^{-1}}{\sum_{s'} \left[Y_W(s')\right]^{-1}}, \tag{2.10}$$

$$B_j(s) = \frac{1}{J}\left[\sum_{s'} \frac{Y_j(s')}{Y_W(s')}\right] Y_W(s) - Y_j(s). \tag{2.11}$$

The intuition underlying these expressions is rather straightforward. The price of a security $p(s)$ depends negatively on the relative abundance of the final good in the corresponding state of nature. Regarding the first term on the right-hand side of equation (2.11), the term in square brackets reflects the fact that the size of country j's portfolio will be larger the higher its average output relative to the world's output. That is, a country's wealth and its consumption posssibilities are a positive function of its expected output. The term $Y_W(s)$ captures the fact that countries will be able to consume more in states of nature with high world output. As for the second term, country j's purchase of state-s security is inversely related to country j's state-s final-good output due to the representative agent's interest in smoothing consumption across states of nature.

2.2.3 International Portfolio Choice

We now discuss the effects of *ex-ante* uncertainty in the goods markets on the portfolio choices of countries. To build up intuition, we discuss the model's implications on endowment similarity and country size separately. We start by assuming that all countries are of equal size. We then relax this assumption.

The Role of Endowment Similarity

Let us initially simplify the model by assuming away country-size effects: $\phi_j = 1$ for all $j \in J$. Define a country's gross domestic product as $Y_j = rA_jK_j + wA_jL_j$. Without loss of generality, consider country $k \in J_k$. In states of nature in which any country $l \in J_l$ has got a high productivity level, country k's GDP improves due to a price effect, whereas states of nature in which any country $k' \in J_k$, $k' \neq k$, has got a high productivity draw bring about a negative effect on country k's income through the resulting change in factor prices. Country k's GDP is highest when its own productivity level is high: the negative effect of the change in factor prices is smaller than the positive effect on output induced by the productivity increase. Appendix A shows

$$Y_k(k) > \frac{1}{J}Y_W > Y_k(l) > Y_k(k'), \tag{2.12}$$

where $Y_W \equiv \sum_j Y_j(s)$ is constant across states of nature due to the model's symmetry. A country therefore has got a stronger incentive to insure against states of nature in which countries with similar factor endowment ratios have got a high productivity level. And the obvious provider of such insurance is the country that experiences high productivity: the model's symmetry implies $Y_{k'}(k') > \frac{1}{J}Y_W > Y_l(k') > Y_k(k')$.[7]

[7]Notice that the model points to a negative correlation between the incomes of similar countries. However, this is not due to a negative correlation between productivity shocks as a negative correla-

Given the model's symmetry and the absence of aggregate uncertainty, we conjecture the equilibrium exhibits full insurance. It is easy to find asset prices, consumption and portfolio allocations such that all the equilibrium conditions hold and countries manage to fully insure: $p(s) = 1/J$, $C_j(s) = \frac{1}{J} Y_W$, and

$$B_j(s) = \frac{1}{J} Y_W - Y_j(s), \tag{2.13}$$

for all j, s. It is worth noting that this result not only holds for log-utility, but for any concave utility function.

We can now characterise the international portfolios of countries. Consider state of nature k'. From (2.12) and (2.13), $B_k(k') > B_l(k') > 0 > B_{k'}(k')$ for $\mu > 0$: country k' sells insurance against state k' to all other countries. The model's symmetry implies $B_k(k') > B_k(l) > 0 > B_k(k)$: the share in country k's international portfolio is larger for assets issued by a country with a similar factor endowment ratio than for assets issued by the other type of country. In Appendix A we show $B_k(k') - B_l(k') = Y_l(k') - Y_k(k') = \frac{2a}{Y_W} \mu^2 \geq 0$. Thus, for $\mu = 0$, $B_k(k') = B_l(k')$: when relative factor endowment differences are small, countries k and l do not differ in their investment decisions regarding country k'. For low values of μ, all countries are very similar in their relative factor endowments. Thus, a shock to any particular country will hardly have an important effect on factor prices; in this case, any two countries will take identical positions in any third country.[8]

Define the following elasticity:

$$\beta \equiv \frac{\frac{B_k(k') - B_k(l)}{B_k(l)}}{-\frac{\left|\frac{K_{k'}}{L_{k'}} - \frac{K_k}{L_k}\right| - \left|\frac{K_l}{L_l} - \frac{K_k}{L_k}\right|}{\left|\frac{K_l}{L_l} - \frac{K_k}{L_k}\right|}} > 0. \tag{2.14}$$

tion between real outputs (that is, outputs measured at constant prices) would show, but to the effect of country-specific shocks on world prices.

[8]For $\mu = 0$ (or for ε infinite) our model is similar to the one-good standard textbook treatment. See, for example, Obstfeld and Rogoff (1995), chapter 5.

β measures the responsiveness of country k's portfolio position to a rise in endowment similarity from country-pair kl to country-pair kk'. It is easy to show that the elasticity β rises with μ: $d\beta/d\mu > 0$.[9] As μ increases, differences in factor endowment similarity between country pairs become more pronounced; since in this case the effects of shocks on relative prices become stronger, country k invests a larger share of its wealth in country k', and a smaller one in country l.

Finally, one can also show that for $\mu = 1/2$, $B_l(k') = 0 < B_k(k')$; or, by symmetry, $B_k(l) = 0 < B_k(k')$: with complete specialisation and a unitary elasticity of substitution, relative prices offer complete insurance against shocks in countries with different specialisation patterns. However, this result is particular to the Cobb-Douglas assumption on preferences: when we simulate the model with $C_j = \left[(C_{1j})^{\frac{\varepsilon-1}{\varepsilon}} + (C_{2j})^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}$, where $\varepsilon > 1$ is the (constant) elasticity of substitution between goods 1 and 2, the model yields a positive $B_l(k')$, as higher elasticities of substitution imply a lower response of prices to productivity shocks. The rest of our results, however, are robust to values of ε larger than one.[10]

In this two-good two-factor model there is an obvious equivalence between factor endowment similarity, production structure similarity, and terms of trade correlations. As we discuss in Appendix B, however, this is a particular feature of the 2x2 model that breaks down if there are more goods than factors; in this case, the production structures of countries and thus their terms of trade are undetermined, but the implications of our model for factor prices remain unaltered.[11]

The Role of Size

We now allow for differences in country size, as we assumed initially. For tractability purposes, we consider the log-utility case (see equations (2.10) and (2.11)). For a

[9] Expressions for $B_k(k')$ and $B_k(l)$ can be found in Appendix A.

[10] We assumed $J = 4$ and $a = 0.02$, and allowed ε and μ to vary in the ranges $(1, 16]$ and $[1/16, 1/2]$, respectively. The corresponding results are available upon request.

[11] In our empirical work below, we control for country-pair similarity in production structures anyway.

given level of endowment similarity, we study how the host country's size affects the positions of investor countries. For this purpose, we compare the portfolio choices of two investor countries, $k \in J_k$ and $l \in J_l$, with the same size ($\phi_k = \phi_l$) across host countries $k', k'' \in J_k$ with different sizes ($\phi_{k'} < \phi_{k''}$).

From (2.10) and (2.11),

$$p(k')\left[B_k(k') - B_l(k')\right] = \frac{1}{\sum_{s'}\left[Y_W(s')\right]^{-1}}\left[\frac{Y_l(k') - Y_k(k')}{Y_W(k')}\right], \tag{2.15}$$

where $k' \in J_k$. The term $\sum_{s'}\left[Y_W(s')\right]^{-1}$ is constant. Hence, all we need to analyse is the behaviour of the term

$$\frac{Y_l(k') - Y_k(k')}{Y_W(k')} = \frac{1}{2}(L_l - L_k)\left[\frac{1}{L_W(k')} - \frac{1}{K_W(k')}\right] > 0, \tag{2.16}$$

as $L_l > L_k$, and $L_W(k') < K_W(k')$ for all $k' \in J_k$. (See Appendix A.) Hence, $p(k')\left[B_k(k') - B_l(k')\right] > 0$. This result simply restates the role of endowment similarity discussed above.

The inequality $p(k'')\left[B_k(k'') - B_l(k'')\right] > p(k')\left[B_k(k') - B_l(k')\right]$ holds if $\frac{1}{L_W(k'')} - \frac{1}{K_W(k'')} > \frac{1}{L_W(k')} - \frac{1}{K_W(k')}$. A sufficient condition for this is

$$\left(\frac{1}{4} - \mu^2\right)a^2 < \frac{\left(\sum_{k\in J_k}\phi_k\right)^2}{\phi_{k''}\phi_{k'}}. \tag{2.17}$$

Two opposite effects are at stake here. A shock to a larger country has a stronger effect on relative factor prices, leading to a larger difference in the security purchases by countries k and l. Country k will want to take a larger position to insure against the negative effect of the shock on its income, whereas country l will take a smaller position due to the implicit insurance it receives through the change in relative prices. (We call this the *quantity effect,* since it relates to the term $B_k(k'') - B_l(k'')$.) At the same time, a shock to a larger country raises world output by more in the corresponding state of nature, leading to a lower price of the associated security. (We call this the *price effect,* since it relates to the term $p\,(k'')$.)

The sufficient condition (2.17) makes sure that the quantity effect is stronger than the price effect. Notice that, for given values of ϕ_j, a higher μ implies a larger quantity effect, as the productivity shock on the large country will translate into a large effect on relative factor prices. As μ decreases, the highest a compatible with the sufficient condition decreases: the less dissimilar countries k and l, the smaller the quantity effect. This sufficient condition is very weak, as the term on the right-hand side of equation (2.17) is larger than one; the term in parenthesis on the left-hand side is smaller than one; and realistic values for a, the percentage increase in productivity experienced by a country in the event of a shock, are far less than one.

2.2.4 International Portfolio Choice without Arrow-Debreu Securities

The model above delivers the key intuitions that explain our empirical findings: other things equal, countries with more similar (dissimilar) relative factor endowments invest more (less) in one another due to better (worse) insurance possibilities. However, many of the model's assumptions and implications are at odds with reality. First of all, most real-life assets are not Arrow-Debreu. Moreover, international consumption correlations are lower than output correlations, which suggests that actual international risk sharing is far from the complete asset market benchmark. (See Backus *et al.* (1992).) Finally, countries tend to invest most of their wealth in their own domestic assets. (See French and Poterba (1991).)

In Appendix C, we show that a similar model with a more realistic financial side also predicts a positive relationship between factor endowment similarity and international portfolio choice. Assume investors can buy ownership claims on countries' GDPs rather than Arrow-Debreu securities. Assume also that holding foreign assets is subject to frictions. This creates a home bias in each country's portfolio, and leads in turn, within the portfolio share that is invested in foreign assets, to a bias towards assets issued by countries with similar relative factor endowments. This is due to the fact that the latter provide a home-biased portfolio with better insurance for the same reasons we discussed above.

2.2.5 Sectoral Shocks

A detailed analysis of the implications of sectoral shocks is beyond the scope of this paper. However, for comparison purposes, we find it worthwhile sketching the investment patterns arising in our setup in the presence of sectoral shocks rather than country-specific shocks.[12] Modeling sectoral shocks is rather straightforward in the two-good, two-factor version of the model: consider production functions $y_{1j} = A_K K_j$ and $y_{2j} = A_L L_j$, where A_K and A_L are now sector-specific productivity levels. Define $A(s) = [A_K(s), A_L(s)]$ and consider states of nature $A(1) = [1+a, 1]$ and $A(2) = [1, 1+a]$ with probabilities $\pi(1) = \pi(2) = 1/2$. The rest of assumptions of our benchmark model remain the same.

Notice that the incomes of countries with similar production structures are now perfectly correlated. Hence there would be no need for them to invest in one another. Regarding the investment flows between dissimilar countries, a number of cases arise that depend on the values of μ and ε:

1. $\mu = 0$, $\varepsilon \geq 1$: When all countries have got identical factor endowment ratios ($\mu = 0$), sectoral shocks do not lead to any portfolio investment, as the incomes of all countries are perfectly correlated. (Recall that with country-specific shocks there would be some portfolio investment in this case: $J-1$ countries would take

[12] In a completely different setup, Campa and Fernandes (2006) argue that country and industry shocks are of comparable size.

the same position in the country suffering the shock.)

2. $\mu \in (0, 1/2]$, $\varepsilon = 1$: It is easy to show that with a Cobb-Douglas final good aggregator countries with different factor endowment ratios need not invest in one another for insurance purposes. This is because relative prices offer all the necessary insurance. (With country-specific shocks instead, countries with similar relative factor endowments would want to invest in one another, as we saw above.)

3. $\mu \in (0, 1/2]$, $\varepsilon > 1$: This is arguably the most interesting case. For a positive μ, as ε rises relative prices become less responsive to shocks. This implies that countries with different factor endowment ratios would invest more in one another as ε grows. (With country-specific shocks instead, as ε grows, the investment positions between dissimilar countries catch up with those between similar countries: the investment positions between dissimilar countries rises, as relative price changes offer little insurance, whereas the investment positions between similar countries falls, as less insurance is needed.)

In general, the crisp predictions of the two-good, two-factor model fail to hold when there are more goods than factors due to the indeterminacy of production structures. In any case, in our empirical work we use a proxy for similarity in production structures so as to control, among other things, for the effects of sectoral shocks on portfolio choice.

2.3 Empirical Strategy

We estimate an equation that relates the amount invested by source country S in host country H to a proxy for relative factor-endowment similarity between countries S and H, and other controls, such as proxies for frictions in commodity and asset markets, as well as host- and source-country fixed effects. Consider the following expression:

$$B_{SH} = e^{\alpha_D D_{SH}} Z_{SH}^{\alpha_Z} u_{SH}, \tag{2.18}$$

where B_{SH} denotes country S's portfolio investment in country H; α denotes parameters; D_{SH} denotes a proxy for factor-endowment similarity between countries S and H; Z_{SH} stands for a country-pair control; and u_{SH} denotes an error term assumed to be statistically independent of the variables on the right-hand side of the equation.[13]

Notice we are allowing for a non-constant elasticity of country S's portfolio investment in country H, B_{SH}, with respect to the similarity proxy D_{SH}: equation (2.18) yields an elasticity $\alpha_D D_{SH}$, which is increasing in D_{SH} in the same way that the responsiveness β of the source-country's portfolio positions rises with the similarity of its factor endowment ratio to that of the host country. (See equation (2.14).)

2.3.1 Estimation

Apart from using the OLS and Tobit estimators to estimate equation (2.18), we also use the Poisson estimator. While equations like (2.18) are usually log-linearised and estimated by OLS, this practice may be inappropriate for a number of reasons. First, B_{SH} can be zero, in which case log-linearisation is unfeasible. (This problem is often solved by adding one to all observations before taking logs.[14]) Second, as Santos-Silva and Tenreyro (2005) have recently pointed out, under heteroskedasticity, the expected value of the log-linearised error will in general be correlated with the regressors, and OLS will therefore be inconsistent. This is because the non-linear transformation changes the properties of the error term, as the conditional expectation of $\ln u_{SH}$ depends on the shape of the conditional distribution of u_{SH}. Santos-Silva and Tenreyro (2005) propose

[13]This is similar to the regression equation in Lane and Milesi-Ferretti (2008).

[14]The Tobit estimator is also often used when the dependent variable takes zero and positive values. (Again, a one is added to all observations before taking logs.) However, in the presence of fixed effects, the Tobit estimator may be biased due to the incidental parameters problem.

the following example as an illustration of this problem: assume u_{SH} is distributed log-normal, with $E\left(u_{SH} \mid D_{SH}, Z_{SH}\right) = 1$ and variance $\sigma^2_{SH} = f\left(D_{SH}, Z_{SH}\right)$.[15] $\ln u_{SH}$ will thus be distributed normal, with $E\left(\ln u_{SH} \mid D_{SH}, Z_{SH}\right) = -\frac{1}{2}\ln\left(1+\sigma^2_{SH}\right)$, which is a function of the regressors.

In the face of this problem, it is more appropriate to estimate equation (2.18) in its non-linear form. After assessing the properties of a number of alternative estimators, Santos-Silva and Tenreyro (2005) propose the Poisson pseudo-maximum likelihood estimator (often used for count data) for this task. This estimator turns out to be consistent under relatively weak assumptions (mainly that the model is well specified), and also provides a natural way to deal with zero values, as no logarithmic transformation is necessary for its implementation.

2.3.2 Accounting for Country Size

Other things equal, a larger country will have a stronger effect on world prices. Thus, countries should invest *more* in a larger country with similar relative endowments than in a smaller country; and countries should invest *less* in a larger country with dissimilar endowment ratios than in a smaller country. Country similarity should not have a positive effect on a country's portfolio at all if the host country cannot affect world prices.

To capture this intuition, we classify host countries into two categories: 'small' (those with host GDP's below the median of the sample) and 'large' (those with host GDP's above the median of the sample). We then consider a separate coefficient on D_{SH} for each of these two categories, and test the null hypothesis that the two coefficients are equal. We expect the coefficient for large host countries to be larger.

[15]The characteristics of the data suggest u_{SH} will be heteroskedastic. Since B_{SH} is non-negative, when its conditional expectation approaches zero, the probability of B_{SH} being positive and its conditional variance must also tend to zero. When the conditional expectation of B_{SH} is large instead, it is possible to observe a greater dispersion, as B_{SH} can now deviate from its conditional expectation in either direction.

As a robustness check, we interact D_{SH} with the host country's log-GDP, $\ln(Y_H)$. The type of equation we estimate in this case has got the following form:

$$B_{SH} = e^{\alpha_D D_{SH}} e^{\alpha_Y \ln(Y_H)} e^{\alpha_I [D_{SH} \ln(Y_H)]} Z_{SH}^{\alpha_Z} u_{SH}. \tag{2.19}$$

Provided that D_{SH} takes positive values when countries are similar and negative values when countries are dissimilar,[16] we expect this interaction coefficient to be positive. Let us consider the cases of similar and dissimilar country pairs separately to discuss this. Countries invest more in each other when they have similar factor-endowment ratios, *i.e.* $D_{SH} > 0$. The greater the size of the host country, the greater the investment for a given level of similarity. Alternatively, two countries with dissimilar factor-endowment ratios ($D_{SH} < 0$) want to invest less in one another because of the insurance mechanism relative prices provide. The greater the size of the host country, the more it influences world relative prices, and the more insurance it provides thereby to the source country. Therefore, less investment is required in a host country with dissimilar endowment ratios if it is a large country.

2.3.3 Variables and Data

Our dependent variable B_{SH} is taken from the IMF's Coordinated Portfolio Investment Survey (CPIS).[17] For each participating country, the CPIS reports data on foreign asset holdings by residence of the issuer. These include both equity and debt, but the CPIS has made an effort to exclude foreign direct investment (FDI) from these data.[18] Following the spirit of our model, we use foreign equity holdings as our dependent variable. Data have been released for end-1997 (with only 29 source countries), and

[16] The explanation on how D_{SH} is constructed can be found in section 3.3.

[17] See Lane and Milesi-Ferretti (2008) for a detailed description of the dataset, as well as a discussion of its potential shortcomings.

See Appendix D for a detailed description of variables and sources.

[18] The CPIS considers an investment as FDI (as opposed to portfolio investment) if the foreign investor owns 10 percent or more of the ordinary shares or voting power.

then yearly from end-2001 (with 67 source countries) to end-2006. According to Lane and Milesi-Ferretti (2008), for those countries that participated in the 1997, 2001 and 2002 surveys, there is considerable persistence in bilateral equity holdings. We focus exclusively on the 2002 edition. Table (2.1) reports some information for the countries in our sample.

Our measure of factor-endowment similarity between countries S and H is based on the following variable:

$$d_{SH} \equiv \left| \ln\left(\frac{K}{L}\right)_S - \ln\left(\frac{K}{L}\right)_H \right|. \tag{2.20}$$

The source for aggregate capital-labour ratios is Caselli and Feyrer (2007).[19] Notice that d_{SH} decreases with the similarity of countries and is always positive. For the reasons discussed above, we need our proxy for factor endowment similarity (i) to rise with similarity, and (ii) to take positive values when countries S and H are 'similar enough' and negative values when they are 'dissimilar enough'. For this purpose, we first compute $d'_{SH} = \max(d_{SH}) - d_{SH}$. Then, we finally rearrange our variable to $D_{SH} = d'_{SH} - med(d'_{SH})$, where $med(d'_{SH})$ is the sample median of d'_{SH}.[20] We interpret $D_{SH} > 0$ as the country pair being *similar* in terms of factor-endowment ratios. Equivalently, $D_{SH} < 0$ implies the two countries have *dissimilar* ratios.[21]

We proxy for commodity and asset trade frictions with distance, and dummies for country pairs in which countries participate in the same regional trade agreement, share a border, the same currency, a common language, a colonial relationship (past

[19]Capital is constructed with the perpetual inventory method from time series data on real investment with PWT 6.1 data using a depreciation rate of 0.06. Labor is defined as the number of workers also using PWT 6.1. It is obtained as RGDPCH*POP/RGDPWOK, where RGDPCH is real GDP per capita computed with the chain method. See Caselli (2005) and Caselli and Feyrer (2007) for more details.

[20]Normalising D_{SH} by the mean rather than the median leads to very similar results.

[21]Using a measure of country similarity based on factor endowments has the additional advantage that it is less likely to suffer from endogeneity problems than a measure of production specialization. Recall that the results by Kalemli-Ozcan *et al.* (2003) and Koren (2005) point out a causation channel from international asset market integration to production specialization.

or present), and a common legal origin. The source for these data is Glick and Rose (2001), but for the common legal origin dummy, which is taken from La Porta *et al.* (2003).

Finally, we use a proxy for similarity in the production structures of countries. For any pair of countries j, j', this variable is constructed as

$$E_{jj'} = 2 - \sum_{i} (s_{ij} - s_{ij'})^2, \tag{2.21}$$

where s_{ij} denotes country j's export share of good i to the world.[22] $E_{jj'}$ is always positive and grows with the similarity of the production structures of countries. Data on manufacturing exports are obtained from the World Trade Flows Database (see Feenstra *et al.* (2005)).

2.4 Results

Tables (2.2)-(2.3) report our estimation results from three different econometric estimators (OLS, Poisson, and Tobit) and two ways of dealing with host-country size (first, the division of the sample into two groups of host countries based on GDP size; second, the interaction of the country similarity proxy D_{SH} with the host-country's ln(GDP)).[23]

Table (2.2) reports results obtained without including any control variables but source- and host-country fixed effects; Table (2.3) reports results from regressions including a group of standard controls (related to distance and other trade barriers, cultural and institutional characteristics, etc.) and a proxy for similarity in production structures.[24]

[22] We use exports by country-industry rather than production, because the former is available at fine levels of disaggregation for many more countries than the latter. The correlation between "similarity in exports" and "similarity in K/L ratio" is around 0.16.

[23] All our tables present results with standard errors clustered at the source country level, *i.e.* investor country level.

[24] To control for outliers, in Tables 3-8 we eliminate single observations that account for more than 30% of the total equity invested or received by a country. This reduces the sample by around 5%. We

In each of Tables (2.2)-(2.3) we present eight regressions. Columns (1)-(4) present results for the host countries separated by a dummy variable based on size, while columns (5)-(8) rather present regressions using the interaction term. Column (1) corresponds to the Tobit estimation, column (2) to the OLS estimation with zeroes; finally, columns (3)-(4) corresponds to the Poisson estimation without and with zeroes, respectively. For the regressions based on the interaction term, the same sequential pattern applies.

In the remaining tables we perform a number of robustness checks to our main specification presented in Table (2.3). In Tables (2.4)-(2.5) we only present results for the Poisson estimator, since this is our preferred econometric specification.[25] Table (2.4) presents results for the whole sample, but including additional control variables that will be described in the next paragraph. Table (2.5) redoes the same regressions as in Table (2.4) by looking at relatively rich source and host countries. The World Bank classification for the year 2000 divides countries into four categories: (1) High Income; (2) Upper Middle Income; (3) Lower Middle Income; (4) Low Income. We restrict our attention to countries included in categories (1) and (2).[26]

In each of Tables (2.4)-(2.5) we present twelve regressions, the first six using the dummy separation of host countries, and the last six using the interaction of the log of host GDP. Compared to the regressions in Table (2.3), columns (1)-(2) additionally include the log of bilateral trade (as an additional proxy for trade frictions)[27] and GDP growth correlations, which proxy for productivity shock correlations across countries, that may affect the portfolio positions of countries.

Columns (3)-(4) additionally include two financial variables that may influence asset holdings in an incomplete financial market scenario.The two variables are the correla-

---

also tried (i) eliminating single observations accounting for more than 70% of the total equity invested or received by a country, and (ii) including all outliers. Results are comparable to the ones we report.

[25]Tables for OLS and Tobit estimators provide similar results, and are available upon request.

[26]That is, we remove Colombia, Peru, Indonesia, and Thailand from the sample.

[27]The source for these data is once again Glick and Rose (2001).

tion in stock returns and the correlation between host-country stock market returns and source-country GDP growth, which takes into account the role of the host country's stock market in potentially hedging against the source country's output fluctuations. These two variables, based on data between 1980 and 1996, were constructed by Lane and Milesi-Ferretti (2008), who are confident that the endogeneity of financial correlations to the size of bilateral financial holdings is not a concern, since most foreign equity investment took place since the mid-1990s.

Finally, columns (5)-(6) repeat the same procedure after including two additional variables that proxy for informational frictions: the difference in time zone across countries (to proxy for informational similarities) and the similarity in the log of GDP per capita (constructed in a manner similar to D_{SH}). For the regressions based on the interaction term, the same sequential pattern applies.

2.4.1 Benchmark Correlation

As mentioned above, in the results of Table (2.2) we omit any control variables. In columns (1)-(4) we divide the sample in two parts based on the host country's GDP level. We allow each subsample to have its own coefficient, and no interaction terms are included. We always test the null hypothesis of equal coefficients for the two subsamples. In columns (5)-(8), we interact the similarity in capital-labour ratios with $\ln(GDP)$.

In columns (1)-(4), we observe that the coefficient related to large host countries is always greater than the one for small countries, even though the difference is not always statistically significant. At this preliminary stage of the empirical analysis, we can already expect that most of the action will be concentrated in large host countries. In columns (5)-(8) we observe that for all three econometric specifications (Tobit, OLS, Poisson) our interaction term is positive. Except for column (5), the estimated coefficient is also statistically significant. However, since the coefficient on the similarity of capital-labour ratio is rather negative (and usually statistically significant), we cannot conclude that the total effect is positive and statistically significant, as our theory predicts, until we do not calculate the value of the combined coefficient, hereafter CC. This coefficient will depend on the host country's GDP size, and is determined in the following way:

$$CC = coef[D_{SH}] + \ln(GDP) * coef[D_{SH} * \ln(GDP)],$$

where $coef[D_{SH}]$ denotes the coefficient corresponding to endowment similarity, and $coef[D_{SH} * \ln(GDP)]$ denotes the coefficient corresponding to the interaction between factor-endowment similarity and host-country size. This combined coefficient becomes positive well before the mean and median of $\ln(GDP)$ in our sample.[28]

2.4.2 Controls

In Table (2.3) we include controls related to international trade frictions. We also include the similarity in exports as a proxy for the similarity in production structures. In columns (1)-(4), where we divide the sample of host countries by their GDP into two groups, the coefficient for small host countries is never significant, while the coefficient for host large countries is always positive and statistically significant at the 1% level. Additionally, the null hypothesis of same coefficient values is always rejected at the 1% significance level. In columns (5)-(8), where we interact our factor-endowment similarity variable with the log of host-country GDP, the interaction term is always positive and statistically significant at the 1% level as our theory predicts.

For simplicity, let us address the economic significance by examining the OLS columns of Table (2.3). In order to assess the effect of country similarity on equity positions, we first focus on the separation of host countries into two groups. For large

[28]The mean and median of $\ln(GDP)$ in our sample are 26.8 and 26.6, respectively.

host countries, an increase in the index of factor-endowment similarity by 10% leads to an increase in equity positions towards the host country by 2.6% under a similarity index value of 0.25.[29] Similarly, when looking at the equations that interact the similarity index with the log of host country GDP, we find that an increase in the index of factor-endowment similarity by 10% leads to an increase in equity positions towards the host country by 1.7% when we are at the mean of host country GDP, *i.e.* $\ln(GDP) = 26.85$, and the similarity index takes the value of 0.25.[30] When the size of the host country increases to $\ln(GDP) = 28$, a typical value for a large host country, then the increase in equity positions is of 2.7%, very much in accordance with the value obtained under the separation of host countries into two groups. Additionally, the combined coefficient, previously defined as CC, is positive for values of $\ln(GDP)$ above 24.8.[31]

2.4.3 Robustness Checks

Full Sample: Additional Controls

In Table (2.4) we allow for a number of additional controls described previously. For the regressions with the dummy separation, columns (1)-(6), we always find that the coefficient for large countries is positive and statistically significant at the 1% level. On the other hand, the coefficient for the small host countries is never significant. Additionally, we test for equality of coefficients and always reject the null hypothesis of equal coefficients between subsamples.[32] For the regressions with the interaction term with log of host GDP (columns (7)-(12)), the interaction term is always positive and statistically significant at the 1% level. And as in previous regressions, the combined

[29]The equation for the non-constant elasticity is: $\frac{d(Eq_{ij})}{d(D_{SH})}\frac{D_{SH}}{Eq_{ij}} = \alpha_D D_{SH}$. Therefore, under $D_{SH} = 0.25$ and a coefficent $\alpha_D = 1.047$, we obtain an elasticity value of 0.26.

[30]The equation for the non-constant elasticity is: $\frac{d(Eq_{ij})}{d(D_{SH})}\frac{D_{SH}}{Eq_{ij}} = (\alpha_D + \alpha_I \ln(Y_H)) D_{SH}$. Therefore, under $D_{SH} = 0.25$, $\ln(Y_H) = 26.8$, and coefficents $\alpha_D = -8.169$, $\alpha_I = 0.33$, we obtain an elasticity value of 0.17.

[31]See "GDP positive combined coefficient" at the bottom of the table.

[32]See "H0: coef[KL_small]=coef[KL_large]" at the bottom of the table.

coefficient becomes positive and statistically significant at the 5% level well before the mean and median of $\ln(GDP)$ in our sample.

Concerning the role played by the additional control variables, we find that the coefficient of the log of bilateral trade is positive and usually statistically significant at the 10% level. Two variables that are about to be significant at the 10% level are the GDP per capita correlations and the informational variable on time difference across countries.

High-income Countries (World Bank)

In Table (2.5) we restrict the sample to countries that belong to the categories High Income or Upper Middle Income, based on the World Bank classification for the year 2000. The main message remains unchanged: similarity in capital-labour ratios matters for equity holdings in the way proposed by our theoretical framework, and this effect becomes stronger as the size of the host country increases. Additionally, the values of the coefficients are similar to the ones with the full sample. For this reason, we can conclude that our results are not driven by developing countries.

The Role of Financial Development

As can be seen from Table (2.1), the stock market capitalisation, a proxy for financial development, differs substantially across countries in our sample. It could be that our results are mainly driven by highly financially developed countries investing in similar countries, especially under the reasonable claim that capital-abundance and financial development are positively correlated.

To make sure that our results are not driven by a subsample of highly developed investor countries, in Table (2.6) we divide the sample of investor countries by the median of stock market capitalization averaged over the period 1985-2000. Columns (1), (3), (5), and (7) present results for the subsample of investor countries with financial development *above* the median of our sample, while the remaining columns show results for the ones *below* the median. Columns (1)-(2) control for the sample variables as in Table (2.3). In columns (3)-(4) we add the variable proxying for the similarity in the production structure, while in columns (5)-(6) we additional control for the log of bilateral trade. Finally, in the last two columns we also control for: correlation in GDP per capita growth, the two financial correlations previously used, and the two information controls (log of time difference and similarity in real GDP per capita).

Throughout all eight specifications based on the full sample, we observe that our theory holds independently of the investor's level of financial development. The coefficient on the dummy for large host countries is always positive and statistically significant at the 1% level, while the dummy for small host countries is never statistically significant. Interestingly, in terms of the significance of control variables, the role played by stock market correlation seems to be much more important than in previous regressions.

2.5 Concluding Remarks

Recent explanations of the international portfolio positions of countries are based on commodity and asset trade frictions: a country invests more in countries with which goods and assets are traded more freely. This paper complements these theories by pointing out that international portfolio decisions are also influenced by the similarity of the capital-labour ratios of countries.

In particular, we introduce a model of international portfolio choice in which countries have varying degrees of similarity in their factor endowment ratios, and are subject to country-specific productivity shocks. The change in relative factor prices after a positive shock in a particular country provides insurance to countries that have dissimilar factor endowment ratios, but harms countries with similar ones. Therefore, countries with similar relative factor endowments have got a stronger incentive to invest in one another for insurance purposes than countries with dissimilar endowments. Since the effect of a shock on relative prices depends on the size of the country, in a generalisation of our model we study how factor endowment similarity interacts with country size.

Our empirical work lends support to this hypothesis: the similarity in relative capital-labour ratios has got a positive effect on the source country's investment position in the host country. The magnitude of this effect depends on the host country's GDP size, as larger countries have a stronger impact on world prices.

Future work should try to elucidate whether and how other sources of comparative advantage also affect the international portfolio decisions of countries.

Bibliography

[1] Acemoglu, D. and F. Zilibotti (1997): "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth," *Journal of Political Economy*, 105, 709-751.

[2] Backus, D., P. Kehoe and F. Kydland (1992): "International Real Business Cycles," *Journal of Political Economy*, 100, 745-775.

[3] Beck, T., A. Demirgüç-Kunt, and R. Levine, (2000), "A New Database on Financial Development and Structure," *World Bank Economic Review* 14, 597-605.

[4] Campa, J.M. and N. Fernandes (2006): "Sources of Gains from International Portfolio Diversification," *Journal of Empirical Finance*, 13, 417-443.

[5] Caselli, F. (2005): "Accounting for Cross-Country Income Differences," *Handbook of Economic Growth.*

[6] Caselli, F. and J. Feyrer (2007): "The Marginal Product of Capital," *Quarterly Journal of Economics*, 122, 535-568..

[7] Cole, H.L. and M. Obstfeld (1991): "Commodity Trade and International Risk Sharing," *Journal of Monetary Economics*, 28, 3-24.

[8] Feenstra, R.C., R.E. Lipsey, H. Deng, A.C. Ma, and H. Mo (2005): "World Trade Flows," *NBER WP 11040.*

[9] French, K. and J. Poterba (1991): "Investor Diversification and International Equity Markets," *American Economic Review, Papers and Proceedings,* 81, 222-226.

[10] Glick, R. and A.K. Rose (2001): "Does a Currency Union Affect Trade? The Time Series Evidence," manuscript.

[11] Helpman, E. (1988): "Trade Patterns under Uncertainty with Country Specific Shocks," *Econometrica*, 56, 645-659..

[12] Helpman, E. and A. Razin (1978): "Uncertainty and International Trade in the Presence of Stock Markets," *Review of Economic Studies*, 45, 239-250.

[13] Kalemli-Ozcan, S., B.E. Sørensen, and O. Yosha (2003): "Risk Sharing and Industrial Specialization: Regional and International Evidence," *American Economic Review,* 93, 903-918.

[14] Koren, M. (2005): "Financial Globalization, Portfolio Diversification, and the Pattern of International Trade," manuscript, Harvard University.

[15] Kraay, A., N. Loayza, L. Servén and J. Ventura (2005): "Country Portfolios," *Journal of the European Economic Association*, 3 914-945.

[16] Kraay, A. and J. Ventura (2000): "Current Accounts in Debtor and Creditor Countries," *Quarterly Journal of Economics*, 115, 1137-1166.

[17] Kraay, A. and J. Ventura (2002): "Trade Integration and Risky Assets," *European Economic Review*, 46, 1023-1048.

[18] La Porta, R., F. López de Silanes and A. Shleifer (2003): "What Works in Securities Laws?", *Journal of Finance*, 61, 1-32.

[19] Lane, P. and G.M. Milesi-Ferretti (2001): "The External Wealth of Nations: Measures of Foreign Assets and Liabilities for Industrial and Developing Countries," *Journal of International Economics*, 55, 263-294.

[20] Lane, P. and G.M. Milesi-Ferretti (2002): "Long-Term Capital Movements," *NBER Macroeconomics Annual 2001,* 16, 73-116.

[21] Lane, P. and G.M. Milesi-Ferretti (2008): "International Investment Patterns," *Review of Economics and Statistics*, 90(3), 538-549

[22] Lucas, R.E. (1982): "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics*, 10, 335-360.

[23] Martin, P. and H. Rey (2000): "Financial Integration and Asset Returns," *European Economic Review*, 44, 1327-1350.

[24] Martin, P. and H. Rey (2004): "Financial Super-markets: Size Matters for Asset Trade," *Journal of International Economics*, 64, 335-361.

[25] Obstfeld, M. and K. Rogoff (1995): *Foundations of International Macroeconomics*, MIT Press.

[26] Obstfeld, M. and K. Rogoff (2001): "The Six Major Puzzles in International Macroeconomics. Is There a Common Cause?" *NBER Macroeconomics Annual 2000*, 15, 339-389.

[27] Portes, R. and H. Rey (2005): "The Determinants of Cross Border Equity Flows," *Journal of International Economics*, 65, 269-296.

[28] Samuelson, P.A. (1954): "The Transfer Problem and Transport Costs," *Economic Journal*, 64, 264-289.

[29] Santos-Silva, J.M.C. and S. Tenreyro (2005): "The Log of Gravity," *Review of Economics and Statistics*, 88, 641-658.

[30] Svensson, L.E.O. (1988): "Trade in Risky Assets," *American Economic Review*, 78, 375-394.

[31] Trefler, D. (1993): "International Factor Price Differences: Leontieff was Right!" *Journal of Political Economy*, 101, 961-987.

[32] Ventura, J. (1997): "Growth and Interdependence," *Quarterly Journal of Economics*, 112, 57-84.

[33] Wooldridge, J.M. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

2.A Appendix A: Proofs

2.A.1 The Role of Endowment Similarity

Assume $\phi_j = 1$ for all $j \in J$.

Proof 1: $Y_k(l) > Y_k(k')$

Since $r(l) = w(k') > r(k') = w(l)$,

$$\begin{aligned} Y_k(l) &= \left(\frac{1}{2}+\mu\right) r(l) + \left(\frac{1}{2}-\mu\right) w(l) > \\ &> \left(\frac{1}{2}+\mu\right) r(k') + \left(\frac{1}{2}-\mu\right) w(k') = Y_k(k'). \end{aligned}$$

Tedious algebra yields

$$Y_k(l) - Y_k(k') = \frac{2a}{Y_W}\mu^2. \tag{2.22}$$

Proof 2: $\frac{1}{J}Y_W > Y_k(l)$

Since we have factor price equalization (à la Trefler (1993)), we can find Y_W from the integrated equilibrium:

$$\begin{aligned} Y_W &= Y_W(l) = [y_{1W}(l)]^{\frac{1}{2}} [y_{2W}(l)]^{\frac{1}{2}} = \\ &= \left(\frac{J}{2}K_k + \frac{J}{2}K_l + aK_l\right)^{\frac{1}{2}} \left(\frac{J}{2}L_k + \frac{J}{2}L_l + aL_l\right)^{\frac{1}{2}}. \end{aligned}$$

Concerning $Y_k(l)$,

$$Y_k(l) = r(l)K_k + w(l)L_k = \frac{1}{2}\left[\frac{y_{2W}(l)}{y_{1W}(l)}\right]^{\frac{1}{2}} K_k + \frac{1}{2}\left[\frac{y_{1W}(l)}{y_{2W}(l)}\right]^{\frac{1}{2}} L_k =$$
$$= \frac{1}{2}\left[\left(\frac{\frac{J}{2}L_k + \frac{J}{2}L_l + aL_l}{\frac{J}{2}K_k + \frac{J}{2}K_l + aK_l}\right)^{\frac{1}{2}} K_k + \left(\frac{\frac{J}{2}K_k + \frac{J}{2}K_l + aK_l}{\frac{J}{2}L_k + \frac{J}{2}L_l + aL_l}\right)^{\frac{1}{2}} L_k\right].$$

Tedious algebra yields

$$Y_W - JY_k(l) = Y_W^{-1}\left(a^2 + Ja\right) L_k L_l = Y_W^{-1}\left(a^2 + Ja\right)\left(\frac{1}{4} - \mu^2\right) > 0. \qquad (2.23)$$

Proof 3: $Y_k(k) > \frac{1}{J} Y_W$

Recall $\frac{1}{J} Y_W = \frac{1}{J}\sum_j Y_j(s) = \frac{1}{J}\sum_s Y_j(s)$. Since $\frac{1}{J} Y_W > Y_k(l) > Y_k(k')$, it follows that $Y_k(k) > \frac{1}{J} Y_W$.

Proof 4: $B_k(k'), B_l(k'), B_{k'}(k')$

From (2.1)-(2.4), (2.13), and (2.23),

$$B_k(l) = B_l(k') = Y_W^{-1}\left(\frac{a^2}{J} + a\right)\left(\frac{1}{4} - \mu^2\right) > 0. \qquad (2.24)$$

As for $B_k(k')$, from (2.13), (2.22), and (2.24),

$$B_k(k') = Y_W^{-1}\left[\left(\frac{a^2}{J} + a\right)\left(\frac{1}{4} - \mu^2\right) + 2a\mu^2\right] > 0. \qquad (2.25)$$

Asset market clearing and equations (2.24) and (2.25) yield

$$B_{k'}(k') = -Y_W^{-1}\left[\left(\frac{a^2}{J} + a\right)\left(\frac{1}{4} - \mu^2\right)(J-1) + 2a\mu^2\left(\frac{J}{2} - 1\right)\right] < 0.$$

2.A.2 The Role of Size

We allow for cross-country differences in size. We assume that the distributions of the scaling factor ϕ_j within J_k and J_l are symmetric.

Proof 5: $L_W(k') < K_W(k')$

$$
\begin{aligned}
K_W(k') &= \sum_{k\in J_k}\phi_k\left(\frac{1}{2}+\mu\right)+\sum_{l\in J_l}\phi_l\left(\frac{1}{2}-\mu\right)+\phi_{k'}a\left(\frac{1}{2}+\mu\right)= \\
&= \sum_{k\in J_k}\phi_k+\phi_{k'}a\left(\frac{1}{2}+\mu\right), \qquad (2.26)\\
L_W(k') &= \sum_{k\in J_k}\phi_k\left(\frac{1}{2}-\mu\right)+\sum_{l\in J_l}\phi_l\left(\frac{1}{2}+\mu\right)+\phi_{k'}a\left(\frac{1}{2}-\mu\right)\\
&= \sum_{k\in J_k}\phi_k+\phi_{k'}a\left(\frac{1}{2}-\mu\right) < K_W(k'). \qquad (2.27)
\end{aligned}
$$

Proof 6: Sufficient Condition for $p(k'')[B_k(k'')-B_l(k'')] > p(k')[B_k(k')-B_l(k')]$

$$
\begin{aligned}
K_W(k'')-K_W(k') &= (\phi_{k''}-\phi_{k'})a\left(\frac{1}{2}+\mu\right) > 0, \qquad (2.28)\\
L_W(k'')-L_W(k') &= (\phi_{k''}-\phi_{k'})a\left(\frac{1}{2}-\mu\right) > 0. \qquad (2.29)
\end{aligned}
$$

Notice $p(k'')[B_k(k'')-B_l(k'')]-p(k')[B_k(k')-B_l(k')] > 0$ if

$$
\begin{aligned}
&\left[\frac{1}{L_W(k'')}-\frac{1}{K_W(k'')}\right]-\left[\frac{1}{L_W(k')}-\frac{1}{K_W(k')}\right]=\\
&=\left[\frac{K_W(k'')-K_W(k')}{K_W(k'')K_W(k')}\right]-\left[\frac{L_W(k'')-L_W(k')}{L_W(k'')L_W(k')}\right] > 0, \qquad (2.30)
\end{aligned}
$$

which is equivalent to

$$
\frac{\frac{1}{2}+\mu}{\frac{1}{2}-\mu} > \frac{\left[\sum_{k\in J_k}\phi_k+\phi_{k''}a\left(\frac{1}{2}+\mu\right)\right]\left[\sum_{k\in J_k}\phi_k+\phi_{k'}a\left(\frac{1}{2}+\mu\right)\right]}{\left[\sum_{k\in J_k}\phi_k+\phi_{k''}a\left(\frac{1}{2}-\mu\right)\right]\left[\sum_{k\in J_k}\phi_k+\phi_{k'}a\left(\frac{1}{2}-\mu\right)\right]}. \qquad (2.31)
$$

Condition (2.17) is sufficient condition for this inequality to hold.

2.B Appendix B: A Many-Good Model

This appendix discusses a many-good generalization of the model in section 2. Our purpose here is to show that the model's key feature driving international portfolio choice is relative factor endowment similarity. In the Heckscher-Ohlin model, production structures are undefined in the presence of more goods than factors. Therefore factor endowment similarity does not necessarily imply similar production structures. On the other hand, a country will still be interested in investing a larger share of its international portfolio in countries with similar factor endowments for insurance purposes.

We maintain most of the model's assumptions, but for the ones we mention here:

1. The final good C is now defined over a continuum of goods, which are aggregated in a Cobb-Douglas fashion:

$$C_j = \exp\left[\int_0^1 \ln C_j(z)\,dz\right], \tag{2.32}$$

   where $C(z)$ denotes consumption of freely traded intermediate good z, $z \in [0,1]$.

2. Each industry employs the two production factors, K and L, which are freely mobile between industries. Production functions are also of the Cobb-Douglas type:

$$y_j(z) = [A_j K_j(z)]^{\alpha(z)} [A_j L_j(z)]^{1-\alpha(z)},$$

   where $y_j(z)$ denotes production of good z in country j; and $\alpha(z) \in [0,1]$. For simplicity, we assume $\alpha(z) = z$.[33]

3. There is an upper limit $\bar{\mu} < 1/2$ to the differences in relative factor endowments we can allow for, as we focus (for simplicity) on the factor price equalization case.

[33] Any symmetric distribution of $\alpha(z)$ such that $\alpha(z) = 1 - \alpha(1-z)$ would yield similar results.

4. We assume equal size for all countries: $\phi_j = 1$ for all $j \in J$.

2.B.1 Goods Market Equilibrium

We again assume factor price equalization à la Trefler (1993). We will therefore find equilibrium prices by solving for the integrated equilibrium; *i.e.*, we assume both commodities and factors are freely mobile in the world, as if the latter were a single (closed) economy.

The integrated equilibrium conditions are the following:

- Pricing:

$$p(z) = b(z,r,w), \tag{2.33}$$
$$b(z,r,w) = \left[\frac{r}{\alpha(z)}\right]^{\alpha(z)} \left[\frac{w}{1-\alpha(z)}\right]^{1-\alpha(z)}, \tag{2.34}$$
$$P = \exp\left[\int_0^1 \ln p(z)\, dz\right], \tag{2.35}$$

  where $b(z,r,w)$ denotes industry z's cost function; r and w are, respectively, the prices of capital and labour in efficiency units; and P denotes the price of the final good, which we will use as numeraire: $P = 1$.

- Commodity market clearing:

$$C_W(z) = \frac{PC_W}{p(z)} = y_W(z), \tag{2.36}$$
$$C_W = Y_W = rK_W + wL_W,$$

  where $K_W \equiv \sum_{j \in J} A_j K_j$ and $L_W \equiv \sum_{j \in J} A_j L_j$.

- Factor market clearing:

$$\int_0^1 \frac{\partial b(z,r,w)}{\partial r} y_W(z)\, dz = K_W, \tag{2.37}$$

$$\int_0^1 \frac{\partial b(z,r,w)}{\partial w} y_W(z)\,dz = L_W. \tag{2.38}$$

Putting conditions (2.33), (2.34), (2.36), (2.37), and (2.38) together, $w/r = K_W/L_W$, and $P = e^{-\frac{1}{2}} r^{\frac{1}{2}} w^{\frac{1}{2}}$. These last two equations and the choice of numeraire yield $r = e^{-\frac{1}{2}} (K_W/L_W)^{-\frac{1}{2}}$ and $w = e^{-\frac{1}{2}} (K_W/L_W)^{\frac{1}{2}}$. It is easy to show that the results we discussed in section 2.1 also hold here. Defining country j's gross domestic product as $Y_j = r(A_j K_j) + w(A_j L_j)$, we obtain $Y_k(k) > \frac{1}{J} Y_W > Y_k(l) > Y_k(k')$. The model's symmetry implies $Y_{k'}(k') > \frac{1}{J} Y_W > Y_l(k') > Y_k(k')$.

2.B.2 Asset Market Equilibrium

The following results are the counterpart to the results discussed in section 2.3.1:

1. Assume $\mu > 0$. Consider state of nature k', $k' \in J_k$: $B_k(k') > B_l(k') > 0 > B_{k'}(k')$. By symmetry, $B_k(k') > B_k(l) > 0 > B_k(k)$.

2. For $\mu = 0$, $B_k(k') = B_l(k')$.

3. $d\beta/d\mu > 0$.

2.C Appendix C: International Portfolio Choice without Arrow-Debreu Securities

This appendix discusses a model without Arrow-Debreu securities that yields results comparable to those we obtained in section 2. We assume the same setup as in section 2 on the goods side (including our assumptions on productivity shocks and states of nature), but consider a completely different asset side.

1. Let us simplify by assuming $J = 4$, $\phi_j = 1$ for all $j \in J$, and complete specialization ($\mu = 1/2$).

2. We assume quadratic utility

$$U(C_j) = C_j - \frac{b}{2}C_j^2, \tag{2.39}$$

where $b > 0$.[34]

3. Before uncertainty is realized countries can only exchange ownership claims on their GDPs.

4. International asset trade is costly: a fraction $\tau_{jj'} = \tau \in (0,1)$ of the payoff that country j receives from its claims on country-j' GDP, $j' \neq j$, is wasted as a cost of keeping foreign assets in country j's portfolio ($\tau_{jj} = 0$ for all j).[35]

Let V_j be the market value of country j's uncertain GDP $Y_j \equiv p_j y_j$. The problem's budget constraints can be written as follows:

$$V_j = \sum_{j'=1}^{J} x_{jj'} V_{j'}, \tag{2.40}$$

$$C_j = \sum_{j'=1}^{J} x_{jj'} (1 - \tau_{jj'}) Y_{j'}, \tag{2.41}$$

where $x_{jj'}$ denotes country j's share of ownership claims on country-j' income.[36] Asset market clearing requires $\sum_{j=1}^{J} x_{jj'} = 1$ for all $j' \in J$. Country j's utility maximization problem can be expressed as:

$$\max_{\{x_{jj'}\}_{j'=1}^{J}} \quad E\left[U\left[\sum_{j'=1}^{J} x_{jj'} (1 - \tau_{jj'}) Y_{j'}\right]\right], \tag{2.42}$$

[34] b must be small enough so that $U'(C) > 0$.

[35] This is the classical 'iceberg' assumption due to Samuelson (1954), which has been used in international finance by Martin and Rey (2000, 2004).

[36] Country j's 'total consumption' of the final good (that is, inclusive of the resources wasted in keeping its international portfolio) is $\sum_{j'=1}^{J} x_{jj'} Y_{j'}$.

subject to $V_j = \Sigma_{j'} x_{jj'} V_{j'}$. The first-order conditions with respect to $x_{jj'}$, $j' = 1, ..., J$, yield

$$\frac{\lambda_j V_{j'}}{1-\tau_{jj'}} = E\left[U'(C_j) Y_{j'}\right] = cov\left[U'(C_j), Y_{j'}\right] + E\left[U'(C_j)\right] E(Y_{j'}), \tag{2.43}$$

$j' \in J$, and where λ_j is the Lagrange multiplier associated to the constraint. Due to the model's symmetry, $\lambda_j = \lambda$, $E(Y_j) = E(Y)$, and $V_j = V$ for all j. The presence of international asset market frictions thus implies $cov\left[U'(C_j), Y_j\right] < cov\left[U'(C_j), Y_{j'}\right]$ for all $j' \neq j$. With quadratic utility, this is equivalent to $cov\left[C_j, Y_j\right] > cov\left[C_j, Y_{j'}\right]$. Thus, portfolios will be home-biased due to the presence of frictions.

We now show $x_{kk'} > x_{kl}$. Consider first $k, k' \in J_k$: since $cov\left[C_k, Y_k\right] > cov\left[C_k, Y_{k'}\right]$ and $var(Y_k) = var(Y_{k'})$,

$$\left[x_{kk} - x_{kk'}(1-\tau)\right] var(Y_k) > \left[x_{kk} - x_{kk'}(1-\tau)\right] cov(Y_k, Y_{k'}). \tag{2.44}$$

Since $cov(Y_k, Y_{k'}) < 0$,[37] $x_{kk} > x_{kk'}(1-\tau)$. Consider now $k, k' \in J_k$ and $l, l' \in J_l$: since, from the first-order condition (2.43), $cov\left[C_k, Y_{k'}\right] = cov\left[C_k, Y_l\right]$,

$$x_{kk'}(1-\tau) var(Y_{k'}) + x_{kk} cov(Y_k, Y_{k'}) = x_{kl}(1-\tau) var(Y_l) + x_{kl'}(1-\tau) cov(Y_l, Y_{l'}). \tag{2.45}$$

By symmetry, $var(Y_{k'}) = var(Y_l) > 0$, $cov(Y_j, Y_k) = cov(Y_l, Y_{l'}) < 0$, and $x_{kl} = x_{kl'}$. Solving for $x_{kl}(1-\tau)$,

$$\begin{aligned} x_{kl}(1-\tau) &= \frac{x_{kk'}(1-\tau) var(Y_{k'}) + x_{kk} cov(Y_k, Y_{k'})}{var(Y_{k'}) + cov(Y_k, Y_{k'})} < \\ &< \frac{x_{kk'}(1-\tau)\left[var(Y_{k'}) + cov(Y_k, Y_{k'})\right]}{var(Y_{k'}) + cov(Y_k, Y_{k'})} = x_{kk'}(1-\tau), \end{aligned} \tag{2.46}$$

[37] Under complete specialization, it is easy to prove that $cov(Y_k, Y_{k'}) < cov(Y_k, Y_l) = 0$.

since $x_{kk} > x_{kk'}(1-\tau)$ and $cov(Y_k, Y_{k'}) < 0$. Hence, $x_{kk'} > x_{kl}$: country k invests a larger share of its wealth in country k' than in country l.[38]

This setup is correctly spelt out only for the case $x_{kk} > x_{kk'} > x_{kl} \geq 0$. How can we make sure that we have no shortselling in equilibrium? Notice that in the absence of asset trade frictions ($\tau = 0$), countries would be able to insure fully by choosing $x_{kk} = x_{kk'} = x_{kl} = 1/4$. We can show that x_{kl} is a continuous function of τ. Hence, by continuity, for a small positive τ, $x_{kl} > 0$. (In any case, we do not observe shortselling in the data.)

2.D Appendix D: Sources and Definitions of Variables

1. Bilateral portfolio equity holdings (in millions of US dollars): Portfolio equity instruments issued by host country residents and held by source country residents. Source: 2002 Coordinated Portfolio Survey (IMF). Year: 2002.

2. GDP (in millions of year-2000 US dollars). Source: Penn-World Tables. Year: 2001.

3. Total population. Source: World Development Indicators. Year: 2001.

4. Log of bilateral trade. Source: Glick and Rose (2001).

5. Distance: Logarithm of great circle distance in miles between the capital cities of source and host countries. Source: Glick and Rose (2001).

6. Common border: Dummy variable taking the value of 1 if source and host countries share a border. Source: Glick and Rose (2001).

[38] According to our computer simulations, allowing for CRRA utility and a higher elasticity of substitution between goods, as well as for a less restrictive distribution of states of nature, yields similar results. These results are available upon request.

7. Regional trade agreement (RTA): Dummy variable taking the value of 1 if source and host countries share the same regional trade agreement. Source: Glick and Rose (2001).

8. Currency area: Dummy variable taking the value of 1 if source and host countries are in the same (strict) currency union. Source: Glick and Rose (2001). (Updated for the euro area by the authors.)

9. Colony/Colonizer: Dummy variable taking the value of 1 if source and host countries ever had a colonial relationship. Source: Glick and Rose (2001).

10. Common language: Dummy variable taking the value of 1 if source and host countries share a common language. Source: Glick and Rose (2001).

11. Common legal origin: Dummy variable taking the value of 1 if source and host countries have a legal system with a common origin (common law, French, German, or Scandinavian). Source: La Porta *et al.* (2003).

12. Correlation in GDP growth rates: Correlation between the real GDP growth rate in the source and host country. Authors' own computations based on real GDP growth rates. Source: World Development Indicators. Period: 1981-2000.

13. Correlation of stock returns. Source: Lane and Milesi-Ferretti (2008). Period 1980-96.

14. Correlation of growth-stock returns. Source: Lane and Milesi-Ferretti (2008). Period 1980-96.

15. Capital-labour ratios. Source: Caselli (2005) and Caselli and Feyrer (2007). Year: 2000.

16. Time difference. Variable constructed as $\ln(0.001 + time\_difference)$. Authors' calculation.

17. Exports by sector to the world at the 2-digit level. Source: Feenstra (NBER database). We only include countries that have at most 4 missing sectoral values.

18. Stock market capitalization (average 1985-2000). Source: Beck *et al.* (2000).

Figure 2.1. Scatter plot for small host countries

slope=-0.13 t-stat=-0.35

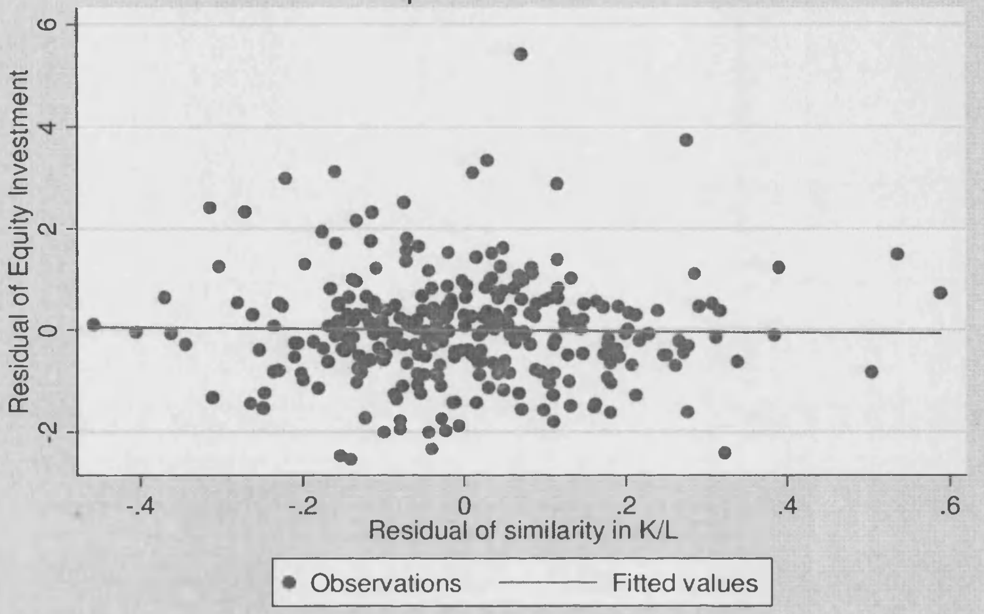

Figure 2.2. Scatter plot for large host countries

slope=1.5 t-stat=4.04

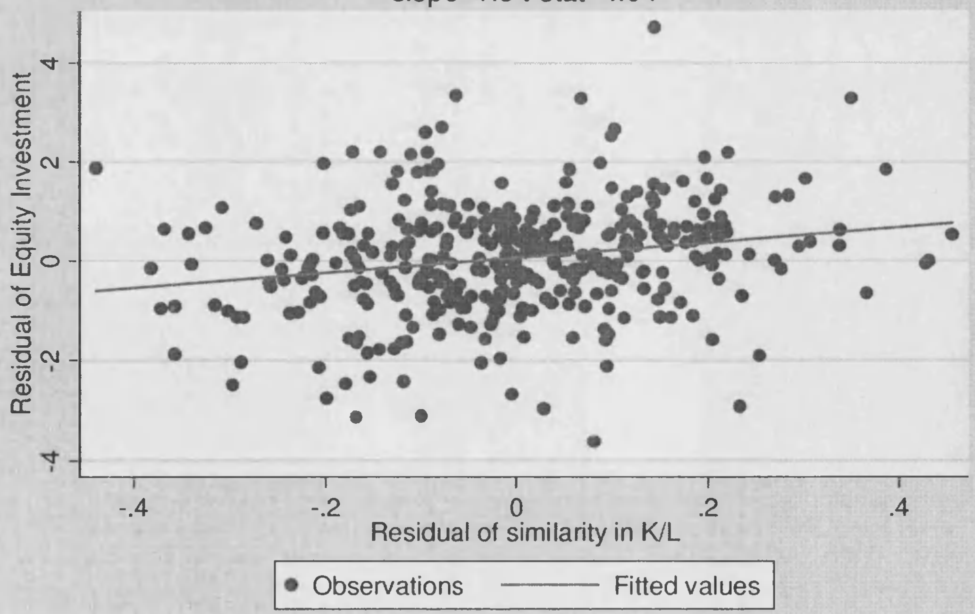

Table 2.1. Information on countries in the sample

Countries	Total source equity	Total host equity	KL ratio	Stock market	GDP
United States	1010832	700712	146391	0.97	9766702
United Kingdom	382012	532698	98998	1.25	1475473
Canada	184005	92856	137800	0.68	834418
Netherlands	179561	181611	134941	0.85	426979
Japan	169489	267749	145955	0.83	3061726
France	135612	217040	142196	0.45	1505280
Ireland	116098	84961	106142	-	101163
Italy	99302	82264	148889	0.25	1321779
Sweden	66628	41224	116631	0.75	225143
Australia	59182	67546	132703	0.63	511186
Belgium	48689	23300	156396	0.48	256399
Norway	43142	12885	177840	0.28	152434
Spain	35424	78611	120603	0.38	826115
South Africa	30911	14895	25692	1.43	360227
Finland	17491	63167	131405	0.67	118716
Austria	15726	3806	150438	0.13	220685
New Zealand	7545	4015	102645	0.43	80377
Argentina	6246	909	50699	0.18	402113
Portugal	4006	9354	87457	0.2	180681
Chile	2166	2715	45687	0.61	178883
Greece	1858	4329	94564	0.28	160250
Israel	1607	11797	115771.	0.37	128604
Korea	1098	50235	108017	0.33	762600
Brazil	939	19778	39218	0.18	1269952
Malaysia	377	6272	58274	1.44	251322
Colombia	245	171	15271	0.98	243894
Indonesia	50	3717	14821	0.15	868471
Thailand	33	5465	38127	0.38	418524
Turkey	8	5605	29238	0.15	363847
Mexico	-	29874	49067	0.22	809757
Peru	-	721	23145	0.12	114061

"GDP" is the real GDP of 2001 in millions of US$ 2000.
"Total source equity" and "Total host equity" are both measured in millions of US$ 2002.
"K/L ratio" is directly taken from Caselli and Feyrer (2007).
"Stock market" is the average stock market capitalization between 1985-2000

Table 2.2. Full Sample. Columns (1)-(4): dummy separation. Columns (5)-(8): ln(GDP) interaction.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Tobit	OLS	Poisson	Poisson	Tobit	OLS	Poisson	Poisson
	Equity>=0	Equity>=0	Equity>0	Equity>=0	Equity>=0	Equity>=0	Equity>0	Equity>=0
Similarity in KL ratio (small host countries)	0.732	0.897	0.624	0.605				
	[0.234]***	[0.156]***	[0.312]**	[0.302]**				
Similarity in KL ratio (large host countries)	1.030	1.320	0.978	0.942				
	[0.187]***	[0.251]***	[0.340]***	[0.318]***				
Similarity in KL ratio					-2.003	-4.157	-5.840	-5.693
					[2.697]	[2.368]*	[3.266]*	[3.249]*
Similarity in KL ratio*ln(GDP)					0.108	0.196	0.239	0.233
					[0.099]	[0.091]**	[0.119]**	[0.120]*
Observations	687	687	560	687	687	687	560	687
H0: coef[KL_small]=coef[KL_large]:	0.26	0.05	0.29	0.32				
GDP positive combined coefficient					18.5	21.2	24.4	24.4
R-squared		0.86				0.86		
Pseudo R-squared	0.38		0.95	0.95	0.38		0.95	0.95

Note: Equity holdings of source/investor country *i* in host/recipient country *j* are measured in millions of US dollars. The dependent variable is ln(1+Equity) in the case of OLS and Tobit, while it is Equity in the Poisson regressions. Regressions include source and host country fixed effects. Standard errors are clustered at the source country level and are reported in parenthesis. *,**,*** indicate statistical significance at the 10%, 5%, 1% level, respectively. The Pseudo R-squared for Poisson is without clustering.

Table 2.3. Full Sample. Columns (1)-(4): dummy separation. Columns (5)-(8): ln(GDP) interaction.

	(1) Tobit Equity>=0	(2) OLS Equity>=0	(3) Poisson Equity>0	(4) Poisson Equity>=0	(5) Tobit Equity>=0	(6) OLS Equity>=0	(7) Poisson Equity>0	(8) Poisson Equity>=0
Log of distance	-0.736	-0.602	-0.161	-0.163	-0.731	-0.589	-0.159	-0.159
	[0.122]***	[0.134]***	[0.074]**	[0.074]**	[0.121]***	[0.136]***	[0.077]**	[0.077]**
Common legal origin	0.459	0.375	-0.092	-0.092	0.455	0.379	-0.056	-0.055
	[0.145]***	[0.151]**	[0.178]	[0.178]	[0.144]***	[0.151]**	[0.181]	[0.182]
Dummy for common border	0.271	0.148	0.430	0.432	0.303	0.198	0.469	0.470
	[0.297]	[0.327]	[0.129]***	[0.131]***	[0.297]	[0.329]	[0.129]***	[0.131]***
Dummy for common language	0.396	0.299	0.492	0.498	0.406	0.290	0.494	0.498
	[0.223]*	[0.242]	[0.175]***	[0.175]***	[0.222]*	[0.248]	[0.174]***	[0.174]***
Dummy for ever colony/colonizer	0.460	0.537	-0.180	-0.184	0.454	0.538	-0.234	-0.238
	[0.293]	[0.342]	[0.206]	[0.205]	[0.293]	[0.355]	[0.199]	[0.198]
Dummy. Strict currency area	0.055	0.046	0.968	0.975	0.072	0.065	0.976	0.982
	[0.270]	[0.315]	[0.221]***	[0.221]***	[0.270]	[0.318]	[0.224]***	[0.224]***
Dummy for regional trade agreement	0.080	0.217	0.196	0.194	0.094	0.233	0.211	0.209
	[0.289]	[0.341]	[0.183]	[0.185]	[0.288]	[0.348]	[0.189]	[0.192]
Similarity in KL ratio (small host countries)	-0.131	0.284	-0.495	-0.495				
	[0.222]	[0.257]	[0.449]	[0.448]				
Similarity in KL ratio (large host countries)	**0.595**	**1.047**	**0.770**	**0.745**				
	[0.179]* **	**[0.311]*** **	**[0.286]*** **	**[0.269]*** **				
Similarity in exports (small host countries)	-0.220	-0.348	-0.866	-0.914				
	[2.185]	[2.008]	[2.271]	[2.250]				
Similarity in exports (large host countries)	-0.602	0.216	0.969	0.883				
	[2.244]	[2.365]	[2.253]	[2.226]				
Similarity in KL ratio					-7.605	-8.169	-7.816	-7.602
					[2.596]***	[2.178]***	[2.198]***	[2.267]***
Similarity in KL ratio*ln(GDP)					**0.293**	**0.330**	**0.299**	**0.291**
					[0.095]* **	**[0.084]*** **	**[0.075]*** **	**[0.078]*** **
Similarity in exports					10.137	4.637	-11.219	-11.176
					[18.767]	[15.829]	[7.867]	[7.640]
Similarity in exports*ln(GDP)					-0.406	-0.195	0.374	0.370
					[0.695]	[0.588]	[0.254]	[0.245]
Observations	650	650	523	650	650	650	523	650
H0: coef[KL_small]=coef[KL_large]:	0.00	0.00	0.00	0.00				
GDP positive combined coefficient					26.0	24.8	26.1	26.1
R-squared		0.88				0.88		
Pseudo R-squared	0.43		0.96	0.96	0.43		0.96	0.96

Note: Equity holdings of source/investor country *i* in host/recipient country *j* are measured in millions of US dollars. The dependent variable is ln(1+Equity) in the case of OLS and Tobit, while it is Equity in the Poisson regressions. Regressions include source and host country fixed effects. Standard errors are clustered at the source country level and are reported in parenthesis. *,**,*** indicate statistical significance at the 10%, 5%, 1% level, respectively. The Pseudo R-squared for Poisson is without clustering.

Table 2.4. Full Sample. Columns (1)-(6): dummy separation. Columns (7)-(12): ln(GDP) interaction.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0
Log of distance	0.006	0.011	0.026	0.027	0.008	0.009	-0.042	-0.034	-0.025	-0.021	-0.039	-0.035
	[0.110]	[0.111]	[0.109]	[0.109]	[0.106]	[0.106]	[0.119]	[0.119]	[0.118]	[0.118]	[0.113]	[0.113]
Common legal origin	-0.121	-0.125	-0.126	-0.129	-0.018	-0.021	-0.071	-0.075	-0.077	-0.079	0.025	0.023
	[0.156]	[0.157]	[0.158]	[0.159]	[0.147]	[0.146]	[0.157]	[0.158]	[0.160]	[0.161]	[0.149]	[0.149]
Dummy for common border	0.397	0.397	0.386	0.383	0.425	0.417	0.435	0.433	0.431	0.428	0.460	0.450
	[0.130]***	[0.132]***	[0.134]***	[0.136]***	[0.134]***	[0.137]***	[0.131]***	[0.133]***	[0.136]***	[0.138]***	[0.138]***	[0.141]***
Dummy for common language	0.502	0.505	0.479	0.483	0.441	0.446	0.492	0.494	0.475	0.478	0.437	0.438
	[0.178]***	[0.177]***	[0.179]***	[0.178]***	[0.165]***	[0.163]***	[0.178]***	[0.177]***	[0.181]***	[0.180]***	[0.165]***	[0.163]***
Dummy for ever colony/colonizer	-0.171	-0.175	-0.153	-0.157	-0.101	-0.111	-0.232	-0.236	-0.221	-0.225	-0.172	-0.181
	[0.207]	[0.206]	[0.210]	[0.209]	[0.206]	[0.204]	[0.201]	[0.200]	[0.205]	[0.204]	[0.203]	[0.201]
Dummy. Strict currency area	0.985	0.989	0.991	0.992	0.994	0.995	0.988	0.991	0.994	0.994	0.996	0.996
	[0.243]***	[0.243]***	[0.240]***	[0.240]***	[0.244]***	[0.244]***	[0.249]***	[0.249]***	[0.246]***	[0.246]***	[0.250]***	[0.249]***
Dummy for regional trade agreement	0.223	0.214	0.237	0.225	0.285	0.285	0.186	0.175	0.199	0.186	0.226	0.217
	[0.107]**	[0.105]**	[0.115]**	[0.113]**	[0.142]**	[0.133]**	[0.131]	[0.127]	[0.137]	[0.132]	[0.148]	[0.142]
Similarity in KL ratio (small host countries)	-0.538	-0.525	-0.561	-0.550	-0.465	-0.437						
	[0.442]	[0.437]	[0.431]	[0.426]	[0.409]	[0.399]						
Similarity in KL ratio (large host countries)	**0.741**	**0.723**	**0.741**	**0.723**	**0.875**	**0.875**						
	[0.283]*	**[0.268]***	**[0.281]***	**[0.266]***	**[0.322]***	**[0.323]***						
Similarity in exports (small host countries)	-0.979	-1.079	-1.195	-1.311	-1.140	-1.272						
	[2.141]	[2.115]	[2.141]	[2.118]	[2.051]	[2.012]						
Similarity in exports (large host countries)	0.818	0.681	0.698	0.560	0.396	0.202						
	[2.122]	[2.081]	[2.105]	[2.068]	[1.948]	[1.907]						
Similarity in KL ratio							-7.737	-7.465	-7.704	-7.443	-7.356	-7.203
							[2.470]***	[2.493]***	[2.491]***	[2.513]***	[2.551]***	[2.572]***
Similarity in KL ratio*ln(GDP)							**0.295**	**0.285**	**0.294**	**0.284**	**0.286**	**0.281**
							[0.084]*	**[0.086]***	**[0.085]***	**[0.087]***	**[0.091]***	**[0.093]***
Similarity in exports							-6.591	-6.434	-7.262	-7.268	-4.892	-3.391
							[9.114]	[8.876]	[8.589]	[8.367]	[9.828]	[9.623]
Similarity in exports*ln(GDP)							0.206	0.196	0.227	0.223	0.137	0.077
							[0.296]	[0.285]	[0.277]	[0.266]	[0.330]	[0.322]
Log of bilateral trade	0.181	0.192	0.188	0.198	0.207	0.212	0.137	0.150	0.143	0.154	0.170	0.180
	[0.121]	[0.121]	[0.118]	[0.119]*	[0.120]*	[0.120]*	[0.130]	[0.130]	[0.130]	[0.129]	[0.126]	[0.126]
GDP growth correlations	-0.256	-0.246	-0.315	-0.304	-0.385	-0.359	-0.217	-0.209	-0.250	-0.241	-0.317	-0.299
	[0.177]	[0.176]	[0.209]	[0.209]	[0.224]*	[0.228]	[0.193]	[0.192]	[0.229]	[0.228]	[0.242]	[0.244]
Correlation stock returns			0.379	0.359	0.389	0.408			0.257	0.239	0.268	0.292
			[0.302]	[0.304]	[0.354]	[0.347]			[0.328]	[0.330]	[0.361]	[0.358]
Correlation GDP growth - stock return			-0.093	-0.115	-0.102	-0.125			-0.025	-0.047	-0.029	-0.050
			[0.203]	[0.204]	[0.200]	[0.201]			[0.221]	[0.222]	[0.219]	[0.219]
Similarity in ln(GDP per capita)					-0.162	-0.286					-0.177	-0.323
					[0.472]	[0.456]					[0.581]	[0.557]
Time difference					0.037	0.036					0.035	0.035
					[0.024]	[0.024]					[0.024]	[0.024]
Observations	523	650	501	624	501	624	523	650	501	624	501	624
H0: coef[KL_small]=coef[KL_large]:	0.00	0.00	0.00	0.00	0.00	0.00						
GDP positive combined coefficient							26.2	26.2	26.2	26.2	25.7	25.6
Pseudo R-squared	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96

Note: Equity holdings of source/investor country *i* in host/recipient country *j* are measured in millions of US dollars. The dependent variable is ln(1+Equity) in the case of OLS and Tobit, while it is Equity in the Poisson regressions. Regressions include source and host country fixed effects. Standard errors are clustered at the source country level and are reported in parenthesis. *,**,*** indicate statistical significance at the 10%, 5%, 1% level, respectively. The Pseudo R-squared for Poisson is without clustering.

Table 2.5. World Bank sample of High Income and Upper Middle Income countries. Columns (1)-(6): dummy separation. Columns (7)-(12): ln(GDP) interaction.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0
Log of distance	0.006	0.008	0.027	0.025	0.008	0.006	-0.042	-0.038	-0.024	-0.023	-0.038	-0.037
	[0.110]	[0.112]	[0.110]	[0.111]	[0.107]	[0.108]	[0.120]	[0.120]	[0.120]	[0.120]	[0.114]	[0.114]
Common legal origin	-0.122	-0.125	-0.127	-0.128	-0.015	-0.017	-0.072	-0.075	-0.078	-0.079	0.026	0.026
	[0.157]	[0.158]	[0.159]	[0.160]	[0.148]	[0.148]	[0.158]	[0.159]	[0.161]	[0.162]	[0.151]	[0.151]
Dummy for common border	0.401	0.396	0.390	0.383	0.428	0.417	0.442	0.436	0.439	0.431	0.466	0.455
	[0.132]***	[0.134]***	[0.137]***	[0.138]***	[0.137]***	[0.140]***	[0.133]***	[0.135]***	[0.138]***	[0.140]***	[0.140]***	[0.143]***
Dummy for common language	0.504	0.507	0.480	0.484	0.443	0.447	0.494	0.496	0.475	0.478	0.437	0.439
	[0.181]***	[0.180]***	[0.182]***	[0.181]***	[0.169]***	[0.167]***	[0.180]***	[0.179]***	[0.184]***	[0.183]***	[0.167]***	[0.166]***
Dummy for ever colony/colonizer	-0.168	-0.174	-0.149	-0.154	-0.098	-0.108	-0.226	-0.232	-0.213	-0.219	-0.166	-0.175
	[0.211]	[0.211]	[0.214]	[0.214]	[0.210]	[0.209]	[0.205]	[0.205]	[0.210]	[0.210]	[0.207]	[0.206]
Dummy. Strict currency area	0.987	0.992	0.992	0.995	0.995	0.997	0.988	0.993	0.995	0.996	0.996	0.998
	[0.244]***	[0.245]***	[0.241]***	[0.241]***	[0.245]***	[0.245]***	[0.250]***	[0.250]***	[0.247]***	[0.247]***	[0.251]***	[0.250]***
Dummy for regional trade agreement	0.234	0.224	0.249	0.237	0.303	0.298	0.205	0.195	0.220	0.208	0.249	0.240
	[0.110]**	[0.108]**	[0.118]**	[0.116]**	[0.145]**	[0.138]**	[0.133]	[0.130]	[0.139]	[0.135]	[0.150]*	[0.145]*
Similarity in KL ratio (small host countries)	-0.556	-0.521	-0.582	-0.547	-0.484	-0.438						
	[0.446]	[0.441]	[0.435]	[0.430]	[0.415]	[0.407]						
Similarity in KL ratio (large host countries)	**0.724**	**0.731**	**0.721**	**0.730**	**0.862**	**0.880**						
	[0.295]**	**[0.298]****	**[0.295]****	**[0.299]****	**[0.339]****	**[0.349]****						
Similarity in exports (small host countries)	-0.982	-1.066	-1.204	-1.304	-1.139	-1.252						
	[2.153]	[2.125]	[2.150]	[2.126]	[2.057]	[2.023]						
Similarity in exports (large host countries)	0.819	0.710	0.696	0.585	0.389	0.241						
	[2.131]	[2.086]	[2.109]	[2.068]	[1.948]	[1.906]						
Similarity in KL ratio							-7.843	-7.560	-7.815	-7.541	-7.484	-7.293
							[2.510]***	[2.512]***	[2.535]***	[2.537]***	[2.585]***	[2.602]***
Similarity in KL ratio*ln(GDP)							**0.298**	**0.289**	**0.297**	**0.288**	**0.290**	**0.284**
							[0.086]***	**[0.086]*****	**[0.087]*****	**[0.088]*****	**[0.093]*****	**[0.094]*****
Similarity in exports							-7.118	-7.082	-7.855	-8.000	-5.108	-4.224
							[9.176]	[9.046]	[8.635]	[8.522]	[9.925]	[9.897]
Similarity in exports*ln(GDP)							0.225	0.220	0.248	0.249	0.144	0.108
							[0.299]	[0.292]	[0.279]	[0.273]	[0.334]	[0.332]
Log of bilateral trade	0.179	0.188	0.185	0.193	0.202	0.207	0.133	0.143	0.138	0.147	0.164	0.173
	[0.122]	[0.123]	[0.119]	[0.120]	[0.121]*	[0.122]*	[0.132]	[0.132]	[0.131]	[0.131]	[0.128]	[0.127]
GDP growth correlations	-0.266	-0.252	-0.328	-0.314	-0.399	-0.376	-0.224	-0.213	-0.259	-0.248	-0.329	-0.312
	[0.179]	[0.180]	[0.214]	[0.215]	[0.229]*	[0.233]	[0.196]	[0.197]	[0.235]	[0.236]	[0.248]	[0.251]
Correlation stock returns			0.386	0.370	0.407	0.414			0.274	0.258	0.293	0.306
			[0.316]	[0.316]	[0.365]	[0.360]			[0.344]	[0.344]	[0.374]	[0.373]
Correlation GDP growth - stock return			-0.099	-0.121	-0.111	-0.135			-0.028	-0.051	-0.035	-0.057
			[0.206]	[0.207]	[0.202]	[0.204]			[0.224]	[0.224]	[0.220]	[0.221]
Similarity in ln(GDP per capita)					-0.198	-0.274					-0.209	-0.308
					[0.480]	[0.462]					[0.591]	[0.574]
Time difference					0.037	0.036					0.036	0.035
					[0.024]	[0.024]					[0.024]	[0.024]
Observations	471	546	449	521	449	521	471	546	449	521	449	521
H0: coef[KL_small]=coef[KL_large]:	0.00	0.00	0.00	0.00	0.00	0.00						
GDP positive combined coefficient							26.3	26.2	26.3	26.2	25.8	25.7
Pseudo R-squared	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96

Note: Equity holdings of source/investor country *i* in host/recipient country *j* are measured in millions of US dollars. The dependent variable is ln(1+Equity) in the case of OLS and Tobit, while it is Equity in the Poisson regressions. Regressions include source and host country fixed effects. Standard errors are clustered at the source country level and are reported in parenthesis. *,**,*** indicate statistical significance at the 10%, 5%, 1% level, respectively. The Pseudo R-squared for Poisson is without clustering.

Table 2.6. Full sample. Separation of source countries based on financial development (FD)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Source countries:	High FD	Low FD	High FD	Low FD	High FD	Low FD	High FD	Low FD
	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0	Equity>0	Equity>=0
Similarity in KL ratio (small host countries)	0.489	0.345	-0.417	-0.266	-0.371	-0.483	-0.137	-0.989
	[0.452]	[0.875]	[0.417]	[1.074]	[0.396]	[1.220]	[0.309]	[0.993]
Similarity in KL ratio (large host countries)	**1.024**	**1.497**	**0.843**	**1.477**	**0.820**	**1.263**	**0.821**	**1.534**
	[0.171]* **	**[0.441]*** **	**[0.313]*** **	**[0.550]*** **	**[0.323]** **	**[0.457]*** **	**[0.274]*** **	**[0.565]*** **
Log of distance			-0.082	-0.527	0.009	-0.125	0.069	-0.463
			[0.083]	[0.110]***	[0.118]	[0.356]	[0.135]	[0.266]*
Common legal origin			0.013	0.157	-0.017	-0.002	-0.040	0.234
			[0.129]	[0.228]	[0.109]	[0.200]	[0.129]	[0.166]
Dummy for common border			0.746	-0.093	0.667	-0.065	0.645	-0.135
			[0.257]***	[0.111]	[0.265]**	[0.237]	[0.264]**	[0.308]
Dummy for common language			0.375	0.732	0.349	0.689	0.339	0.979
			[0.149]**	[0.444]*	[0.140]**	[0.500]	[0.153]**	[0.544]*
Dummy for ever colony/colonizer			-0.265	0.379	-0.272	0.367	-0.225	0.411
			[0.173]	[0.143]***	[0.174]	[0.166]**	[0.180]	[0.196]**
Dummy. Strict currency area			0.909	0.851	0.890	0.992	0.983	0.472
			[0.355]**	[0.303]***	[0.359]**	[0.313]***	[0.358]***	[0.302]
Dummy for regional trade agreement			0.325	0.016	0.118	0.440	0.199	0.526
			[0.201]	[0.226]	[0.238]	[0.311]	[0.182]	[0.309]*
Log of bilateral trade					0.183	0.488	0.212	0.353
					[0.145]	[0.371]	[0.141]	[0.338]
GDP growth correlations					-0.029	-0.667	-0.064	-0.545
					[0.105]	[0.101]***	[0.195]	[0.405]
Correlation stock returns							0.221	-0.856
							[0.521]	[1.126]
Correlation GDP growth - stock return							0.244	-0.174
							[0.227]	[0.487]
Similarity in exports (small host countries)							-2.277	6.918
							[4.037]	[2.119]***
Similarity in exports (large host countries)							6.099	3.483
							[4.409]	[2.244]
Observations	294	202	294	202	294	202	286	189
H0: coef[KL_small]=coef[KL_large]:	0.33	0.36	0.02	0.20	0.03	0.20	0.03	0.04
Pseudo R-squared	0.93	0.94	0.97	0.97	0.97	0.97	0.97	0.97

Note: Equity holdings of source/investor country *i* in host/recipient country *j* are measured in millions of US dollars. The dependent variable is ln(1+Equity) in the case of OLS and Tobit, while it is Equity in the Poisson regressions. Regressions include source and host country fixed effects. Standard errors are clustered at the source country level and are reported in parenthesis. *,**,*** indicate statistical significance at the 10%, 5%, 1% level, respectively. The Pseudo R-squared for Poisson is without clustering.

Chapter 3

Bureaucratic Start-up Costs, External Finance, and a Country's Production Structure

3.1 Introduction

Market economies are characterised by a continuous process of resources being reallocated across agents. One key policy instrument influencing this reallocation of resources is deregulation. Even if at the aggregate level it may be beneficial, deregulation usually comes with strong distributional effects that are not always well understood. In this paper we aim to explore the distributional effects of one specific deregulation instrument: bureaucratic start-up costs to open up new businesses. In particular, the goal is to understand how two main barriers faced by entrepreneurs interact in equilibrium: limited access to external capital and bureaucratic start-up costs.

Understanding how these effects interact is important for a number of reasons. First, it will contribute to assess how entrepreneurs react to the economic environment, which is important from a Schumpeterian perspective of "creative destruction". Second, as many countries are in the process of deregulation, it may contribute to understand the implications of this process. Third, from a political economy perspective, it may help to identify the winners and losers of deregulation.

I adapt the framework by Holmstrom and Tirole (1997) to incorporate start-up costs and where sectors only differ in their need for external capital. In this increasing returns to scale setup, producers' net worth is key to access external capital. The share of value added of each industry is determined by its relative price. In our comparative statics, we obtain the following theoretical prediction: a reduction in the fixed cost faced by entrepreneurs when starting a new firm shifts the number of firms and the share of value added towards the sector with more need for external capital.

Intuitively, a reduction in the bureaucratic start-up cost leads to an equivalent increase in entrepreneurial net worth. Entrepreneurial collateral is especially beneficial in sectors that need to borrow from external lenders, i.e. in sectors with a relatively low level of internally generated cashflow. This is because under asymmetric information between a lender and a borrower, a firm's internal assets will determine the amount of external capital that the lender is willing to provide under the financial contract (for example, see Bernanke and Gertler (1989) or Kiyotaki and Moore (1997)). As a consequence, the relative price of goods produced by sectors with high need of external capital goes down, leading to a proportional increase in their share of value added in the economy.

In the empirical testing of the theoretical prediction, I make use of country-level variation in the bureaucratic cost of starting new firms together with sector-level variation in the need for external finance. In analysing which sectors benefit more from a reduction in the cost of starting a new company, I focus on two dependent variables at the sector-country level: (i) number of establishments; (ii) share of value added. The main empirical finding is consistent with the theoretical prediction: countries with low bureaucratic start-up costs tend to have both a greater share of value added and a greater number of establishments in sectors that are more dependent on external sources of capital.

These findings have two interesting implications. *First*, reductions in bureaucratic costs to start a business and improvements in the financial sector might work as complements. If lower start-up costs shift production towards sectors that have to rely more on capital supplied by the financial sector, then reforms improving the financial system would have a greater positive impact on the economy. Therefore, reductions in bureaucratic start-up costs could bring the right incentives to reform and improve the efficiency of the financial system.

Second, the previous literature on start-up costs found that reductions in the regulatory cost for new companies had a *growth effect* (i.e. higher growth rate in value added). In this paper we find an additional *level effect*, since it also affects the production structure in the long run, both the share of value added and the number of establishments. Therefore, policymakers can use start-up costs to influence the relative importance of sectors across the economy, i.e. the industrial structure of the country.

Regarding the importance of the number of establishments in a given sector, reductions in bureaucratic start-up costs could induce growth by improving the prospects of smaller and younger firms. This could possibly trigger "waves of creative destruction" in the spirit of Schumpeter[1], as data from the U.S. Small Business Administration Report 2003 show that 50 percent of employment in the U.S. is accounted for by firms with less than 500 employees, and they account for 75% of job turnover. Furthermore, we know from Acs and Audretsch (1988, 1990) that small firms are more innovative than large firms in a significant number of industries.

In terms of related literature on the role of start-up costs and growth, Ciccone and Papaioannou (2007) find that countries where it takes less time to register new businesses have more entry in industries that experienced expansionary global demand and

[1] See Rajan and Zingales (1998) for a similar argument on financial development.

technology shifts. Fisman and Sarria-Allende (2004) show that countries with more costly product market regulation see slower entry in industries with growth opportunities as proxied by U.S. industry sales growth. With firm-level data, Klapper et al. (2006) focus on European countries and find that costly entry regulation especially reduces growth in the number of new establishments in industries with high entry in the U.S.

In another strand of literature, the *correlation* between financial development and economic growth is often associated with the work of Goldsmith (1969), McKinnon (1973), and Shaw (1973). In a next step beginning with King and Levine (1993a,b), a number of empirical papers have provided support for financial development *leading to* economic growth. Rajan and Zingales (1998), hereafter RZ, find that sectors more dependent on external finance grow faster in countries with better financial systems. Aghion et al. (2007) use firm-level data of 16 developed and developing countries and find a variety of results confirming that financial development helps potential entrants and small firms.[2]

On the theoretical front, the literature on financial development and economic growth provides mechanisms under which financial development induces faster long run growth, and in many of these papers financial development is modelled as an endogenous outcome. Greenwood and Jovanovic (1990) present a model where financial intermediation and economic growth are both determined endogenously. Saint-Paul (1992) analyses a mechanism which can give rise to multiple equilibria in financial and economic development. In the 'low equilibrium', the underdevelopment of the financial system leads to an inefficient (i.e. less specialised) productive structure which, in turn, justifies the absence of financial development. Similarly, in the book by Hermes and Lensink (1996), Saint-Paul presents a model which explains financial development as being triggered by

[2]Other references include, for example, Rousseau and Wachtel (1998), Demirguc-Kunt and Maksimovic (1996), Levine and Zervos (1998), Braun (2003) or Levine et al. (2000).

an unusual increase in the demand for financial services.[3] I build on this line of reasoning when claiming that shifts in production towards high external finance sectors, that increase the demand for the provision of financial services, can trigger reforms in the financial system.

In a related way, the law and finance literature initiated by LaPorta et al. (1997) focuses on the relationship between financial development and the institutional framework of a country[4]. Financial development is higher in countries with better legal systems and stronger creditor rights. Beck et al. (2003) show that both the legal system and the initial endowment are important determinants of financial development and private property protection.[5] And in a recent speech, Mishkin (2007) describes the steps that must be taken to build an institutional infrastructure that will ensure a well-functioning financial system. He explicitly mentions the bureaucratic cost of starting a new business as one of the main concerns to address, together with other aspects like strong property rights or improvements in corporate governance.[6]

Before moving on to a more rigurous theoretical and empirical analysis, Figures (3.1)-(3.3) give a visual idea of the empirical finding. I calculate the weighted external finance dependence for each country by multiplying each industry's dependence on external finance by the fraction that this industry contributes to total value added in the manufacturing sector in 1990. I then regress this variable against country-level start-up costs. The correlation is negative and strongly statistically significant for the three measures of start-up costs that will be described in more detail in Section 3.

The rest of the paper is structured as follows. Section 3.2 introduces the theoretical

[3] Other references include Blackburn and Hung (1998), Deidda (2006), and Zilibotti (1994). In Deidda (2006), financial development occurs endogenously as the economy reaches a critical threshold of economic development. Also, in Zilibotti (1994), an economy with a capital level over a certain threshold has a 'thick' financial market, which in turn means lower intermediation costs. For a survey, see Levine (1997).

[4] See also LaPorta et al. (1998) and Demirguc-Kunt and Maksimovic (1998).

[5] See also Claessens and Laeven (2003).

[6] Other interesting references on the regulatory environment and firm growth include Kumar et al. (2002), Fonseca and Utrero (2003 and 2004).

framework. In Section 3.3, I present the data and the empirical methodology. In Section 3.4, the empirical results are discussed. Section 3.5 concludes.

3.2 The Model

In this closed economy model based on Holmstrom and Tirole (1997), hereafter HT, there is a continuum of risk-neutral homogeneous agents in the space $[0, K]$, each of them endowed with A units of capital.[7] All agents are protected by limited liability and have access to the same technology.

There are two goods produced, one by each sector. A producer faces a fixed cost $\delta \in (0, A)$ of starting a firm before production can take place.[8] Once the fixed cost is paid, production follows a continuous-investment function in which each unit of investment I_i in sector i yields a return of R_i with probability π, and yields zero return otherwise. The only difference across sectors is in the need of external finance. Sector H has a high need of external finance due to a lower return per unit of investment: $R_H < R_L$. Even though there is idiosyncratic risk at the project level, there will be no aggregate risk considerations in this model.

There are two periods. In the first period, capital is allocated between internal and external capital, i.e. each agent optimally either becomes a producer or a lender. The *producer*'s remaining capital $A - \delta$ is the net worth of the firm. A *lender* lends her assets A to producers at an endogenously determined interest rate r. The allocation of agents in period 1 is symultaneous.[9] In the second period, production takes place, lenders are repaid, and agents consume. Lender's capital will incur an iceberg cost α reflecting the

[7]The value of K can be arbitrarily large. Firms will not have any market power.

[8]A possible interpretation is that δ is the bureaucratic start-up cost that every firm has to pay before running a business. This cost will have a direct mapping in our empirical analysis.

[9]Before firms are established and production takes place, agents can switch between becoming lenders and producers until noone has incentives to move again. Therefore, I eliminate any first-mover advantage of a sequential entry model.

well-established fact that external capital is more costly than internal capital.[10]

3.2.1 Consumer Maximisation

Each risk-neutral agent faces a constant-elasticity-of-substitution (CES) utility function over the two goods, and solves the following maximisation problem:

$$\max_{c_L, c_H} E(U) \;\; s.t. \; Income = P_L c_L + P_H c_H, \tag{3.1}$$

where $U = \left(c_L^{\frac{\varepsilon-1}{\varepsilon}} + c_H^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$. $\varepsilon > 1$ is the elasticity of substitution between goods H and L, and P_i is the price of good i.[11] From standard CES maximisation, optimal relative consumption is given by:

$$\frac{c_H}{c_L} = \left(\frac{P_H}{P_L}\right)^{-\varepsilon} \tag{3.2}$$

where c_i denotes an individual's consumption of good i. Due to the assumed preferences, this condition also holds at the aggregate level:

$$\frac{C_H}{C_L} = \left(\frac{P_H}{P_L}\right)^{-\varepsilon} \tag{3.3}$$

where $C_i = \int_0^K c_i di$. Under these preferences, a reduction in the relative price of a good leads to a more than proportional increase in consumption of this same good.

[10]From the corporate finance literature we know that external funds are generally costlier than internally generated ones. For example, asymmetric information can lead to a monitoring cost (Holmstrom and Tirole, 1997). Alternatively, to ex-ante differenciate good producers from bad ones, a screening cost has to be incurred (Boyd and Prescott, 1986). See also Stiglitz and Weiss (1981) and Myers and Majluf (1984) for references in which outsiders have less knowledge than insiders.

[11]*Income* takes different formulations for producers and lenders, but it is of secondary importance at this point.

3.2.2 Optimal Allocation of Agents

In period 1, the allocation of agents becoming either producers or lenders takes place. In equilibrium, each agent has the same expected rate of return.

Analysis of a given sector

Producer Once the fixed cost δ is paid, each unit of investment I_i gives a return R_i in case the project is successful, and zero otherwise. Producers may deliberately reduce the probability of success by exerting low effort and obtaining a private benefit BI_i in units of the same good.[12] As in the benchmark model of HT, I formalise this moral hazard problem in the following way: the probability of success depends on whether the producer exerts high ($\overline{\pi}$) or low ($\underline{\pi}$) effort, with $\Delta\pi \equiv \overline{\pi} - \underline{\pi} > 0$. Effort is non-observable and non-verifiable for a lender, and the project is only economically viable under high effort:

$$\overline{\pi} P_i R_i I_i > A r_{p,i} + (1+\alpha)(\delta + I_i - A) r_{l,i} > \underline{\pi} P_i R_i I_i + P_i B I_i$$

where $r_{l,i}$ is the lender's rate of return per unit of capital in good i, $r_{p,i}$ is the producer's rate of return per unit of capital in good i, and α is the iceberg cost of external capital.

A necessary condition for the producer to exert high effort is given by the following incentive constraint:

$$\overline{\pi} P_i R_{p,i} I_i \geq \underline{\pi} P_i R_{p,i} I_i + P_i B I_i$$

where $R_{p,i}$ is the return of good i (per unit of investment) that the producer of sector i gets in case of successful outcome. After rearranging, we obtain the same equation

[12] We can think of it as an alternative less productive technology function with a safe return that is not observable by the lender.

as in HT, where the producer must be paid at least the following amount to exert high effort:

$$R_{p,i} \geq \frac{B}{\Delta\pi} \tag{3.4}$$

Consequently, if the producer puts all his capital into the project, the expected producer's rate of return per unit of capital in good i under high effort must be at least

$$E\left(r_{p,i}\right) \equiv \frac{\overline{\pi}P_i R_{p,i} I_i}{A} = \frac{\overline{\pi}P_i B}{\Delta\pi A} I_i \tag{3.5}$$

Lender Lenders allocate external capital into a firm as long as their participation constraint is satisfied:

$$\overline{\pi}P_i I_i\left(R_i - R_{p,i}\right) \geq E(r_{l,i})\left(1+\alpha\right)\left(I_i - A + \delta\right), \tag{3.6}$$

Due to increasing returns to scale (caused by the fixed cost δ), the producer will maximise the firm's size, I_i, by borrowing as much capital as possible. Competitive lenders make no economic profit on the contract that is most advantageous for the borrower. Consequently, in equilibrium the lender's participation constraint will bind, and the upper bound on investment that satisfies the participation constraint is given by:

$$I_i = \frac{A-\delta}{1 - \frac{\overline{\pi}P_i\left(R_i - \frac{B}{\Delta\pi}\right)}{E(r_{l,i})(1+\alpha)}} \tag{3.7}$$

By plugging (3.7) into (3.5) we obtain:

$$E\left(r_{p,i}\right) = \frac{\overline{\pi}P_i B}{\Delta\pi A} \frac{A-\delta}{1 - \frac{\overline{\pi}P_i\left(R_i - \frac{B}{\Delta\pi}\right)}{E(r_{l,i})(1+\alpha)}} \tag{3.8}$$

where $E\left(r_{p,i}\right)$ is expressed as a function of P_i, $E(r_{l,i})$, and parameters.

Equilibrium interest rate $E(r)$

In equilibrium, the expected rates of return to all agents, either producers or lenders, have to be equal within and across sectors. The intuition is that agents allocate themselves simultaneously into the different tasks and nothing stops them from reallocating to another more profitable task before production takes place. As an illustration, take sector i. If $E(r_{p,i}) > E(r_{l,i})$, nothing hinders a lender from becoming a producer. Incentives to deviate stop when $E(r) = E(r_{p,i}) = E(r_{l,i})$. The outside option of a lender is to become a producer, and viceversa, so that any situation in which either $E(r_{p,i}) \neq E(r_{l,i})$ cannot be an equilibrium. Similarly, we also require the same expected return across sectors for a given task: $E(r) = E(r_{p,i}) = E(r_{p,j})$, and $E(r) = E(r_{l,i}) = E(r_{l,j})$ where $i \neq j$.

This framework can be traced back to Kiyotaki (1998). In his model, the real interest rate is determined by the outside option of the lender, which in his scenario is to become a producer at a *predetermined* productivity rate. But in my setting of increasing returns to scale, the productivity of the firm rises with its size. Therefore, the outside option of a lender is *not* to open a new firm without any external capital, but *rather* to additionally attract external capital until the borrowing constraint binds.[13]

Equilibrium investment levels and market clearing conditions

Let us define sector L as the numeraire, $P_L = 1$. Equation (3.8) becomes:

$$E(r_{p,L}) = \frac{\overline{\pi} B}{\Delta \pi A} \frac{A - \delta}{1 - \frac{\overline{\pi}\left(R_L - \frac{B}{\Delta \pi}\right)}{(1+\alpha) E(r_{l,L})}} \tag{3.9}$$

[13]In this equilibrium allocation $E(r_{p,i}) = E(r_{l,i})$, the lender does *not* get her marginal product of capital as a return, but rather less. If each lender obtained her marginal product of capital, the fixed cost would only be paid by the producer, leading to a situation in which $E(r_{p,i}) < E(r_{l,i})$. The method of obtaining the equilibrium interest rate is similar to Krugman (1979), who also makes use of the zero profit condition, together with the fact that all agents in the economy obtain the same wage rate and share the payment of the fixed cost.

By setting $E(r_{p,L}) = E(r_{l,L}) = E(r^*)$, we obtain the equilibrium rate of return per unit of investment:

$$E(r^*) = \overline{\pi}\left[\frac{B}{\Delta\pi}\left(1 - \frac{\delta}{A}\right) + \frac{\left(R_L - \frac{B}{\Delta\pi}\right)}{(1+\alpha)}\right] \tag{3.10}$$

This rate depends negatively on the fixed cost δ and the financial iceberg cost α, and positively on the probability of success $\overline{\pi}$. The equilibrium investment size is found by plugging (3.10) into (3.7),

$$I_L^* = (A - \delta) + \frac{A}{1+\alpha}\left(\frac{\Delta\pi}{B}R_L - 1\right) \tag{3.11}$$

$$I_H^* = (A - \delta) + \frac{A}{1+\alpha}\left(\frac{\Delta\pi}{B}R_H - 1\right) \tag{3.12}$$

These investment sizes depend positively on the producer's net worth $(A - \delta)$, and negatively on the financial iceberg cost α. Since $R_L > R_H$, we observe that the investment size in the sector with high need of external finance is lower, since $I_L^* > I_H^*$.

The relative price of the two goods is obtained by using the equilibrium condition $E(r_{p,L}) = E(r_{p,H}) = E(r^*)$ together with equation (3.5):

$$P_H^* = \frac{I_L^*}{I_H^*} \tag{3.13}$$

Finally, let us solve for the equilibrium ratio of firms across the two sectors, n_H/n_L, where n_i is the number of firms in sector i.[14] This ratio is the first variable that will have a direct mapping in our empirical specification. For a given good i, the market clearing condition for good i is given by

[14] n_i includes firms with both successful and unsuccessful outcomes. In other words, not all n_i firms contribute to output in sector i.

$$C_i = \overline{\pi} n_i^* y_i^* = \overline{\pi} n_i^* \left(R_i I_i^* \right) \tag{3.14}$$

where all successful firms are of equal size y_i^*, and C_i is the aggregate demand for good i. By using (3.13), the consumption ratio, C_L/C_H, is defined as:

$$\frac{C_L}{C_H} = \frac{n_L^*}{n_H^*} \frac{R_L}{R_H} P_H^* \tag{3.15}$$

and using the optimal consumption ratio provided by equation (3.3) leads to the equilibrium ratio of firms across the two sectors:

$$\frac{n_H^*}{n_L^*} = \left(P_H^* \right)^{-(\varepsilon - 1)} \frac{R_L}{R_H} \tag{3.16}$$

The second variable in our empirical specification is proxied by the share of value added, $P_H C_H / C_L$. We can easily obtain this ratio from (3.15) and (3.16):

$$\frac{P_H^* C_H^*}{C_L^*} = \left(P_H^* \right)^{-(\varepsilon - 1)} \tag{3.17}$$

Since $\frac{R_L}{R_H}$ is a constant, there is a one-to-one relation between the two variables of our empirical specification, shown in (3.16) and (3.17): a reduction in the relative price of a good leads to an increase in the relative number of firms and the share of value added in the sector producing that same good.

3.2.3 Comparative Statics

In this subsection we aim to derive the theoretical predictions that will drive our empirical analysis. From equations (3.16) and (3.17) we see that the key variable to consider is P_H^*, i.e. the price of sector H in terms of the numeraire price of sector L. Its derivative with respect to the start-up cost δ is:

$$\frac{\partial P_H^*}{\partial \delta} = \frac{\frac{A}{1+\alpha}\frac{\Delta\pi}{B}\left(R_L - R_H\right)}{\left[(A-\delta) + \frac{A}{1+\alpha}\left(\frac{\Delta\pi}{B}R_H - 1\right)\right]^2} > 0$$

Recall that sector H has a high need of external finance, so that it yields a lower output level per unit of investment: $(R_L - R_H) > 0$. Consequently, a reduction in the fixed cost δ leads to a reduction in the relative price of good H. The direct implication is that a reduction in δ leads to an increase in the relative share of value added and number of firms in the sector with the high need of external finance:

$$\frac{d\left(\frac{n_H^*}{n_L^*}\right)}{d\delta} = \frac{\partial\left(\frac{n_H^*}{n_L^*}\right)}{\partial P_H^*}\frac{\partial P_H^*}{\partial \delta} < 0$$

$$\frac{d\left(\frac{P_H C_H}{C_L}\right)}{d\delta} = \frac{\partial\left(\frac{P_H C_H}{C_L}\right)}{\partial P_H^*}\frac{\partial P_H^*}{\partial \delta} < 0$$

due to

$$\frac{\partial\left(\frac{n_H^*}{n_L^*}\right)}{\partial P_H^*} < 0; \quad \frac{\partial\left(\frac{P_H C_H}{C_L}\right)}{\partial P_H^*} < 0.$$

This is the theoretical prediction we will test in our empirical analysis. Intuitively, sector H experiences a relatively greater improvement in economies of scale, and thereby in productivity, after a reduction in the start-up cost δ. This leads to a reduction in its relative price, which in turn shifts more production towards this sector. In other words, a reduction in the fixed cost increases entrepreneurial net worth, which is especially beneficial in sector that needs to borrow more.

3.3 Data and Empirical Methodology

3.3.1 Data

In this subsection we will describe the two main regressors of our analysis: sectoral external finance and country-level bureaucratic start-up costs.

External finance at the sectoral level is measured as the mismatch between investment and cashflow averaged over the 1980-89 period from U.S. publicly traded firms in COMPUSTAT data (from Kroszner et al., 2007). This measure was first used by RZ and is based on the assumption that an industry's external finance dependence has a time-invariant technological aspect that allows to construct a sectoral ranking that is valid across countries. For example, the project's gestation period or the requirement for continuing investment differ substantially across sectors. The index is constructed using publicly listed U.S. firms; under the assumption that they face minor frictions in accessing external capital, the *actual* and the *desired* amount of investment and external finance should be similar, which allows us to ameliorate any supply side effects. In other words, by looking at the large listed firms in the most financially developed country, their level of external finance is supposed to mainly capture the demand component.[15]

Our country-level measures of bureaucratic start-up costs come from the *Doing Business (2006)* dataset constructed by the World Bank for the year 2005. Between the moment in which an agent decides to start a company and the moment in which production takes place, it seems plausible to argue that at least three country-level bureaucratic costs must be incurred: (1) cost of starting up a business, including the completion of inscriptions for the company and obtaining the permission to open a brand; (2) cost of registering the property on which the production plant will be built; (3) cost of dealing

[15]See Table (3.3) for some sectoral measures. For a more detailed explanation on how this variable was constructed, see also RZ.

with licenses while building the production plant.

The literature has mainly focused on the first cost when proxying for bureaucratic frictions to start a business. But just having a company legally registered is clearly not the full representation of all the bureaucracy that has to be incurred before running a business. It is usually also necessary to buy the land on which production will take place, and finally also build the production plant.[16] All the measures are defined as a percentage of income per capita. The final data for these three variables is listed in Tables (3.1)-(3.2).

Our proxy for the first measure of start-up cost is a combination of the cost of procedures and the cost of time to start a business according to the definition of the *Doing Business* report. These include *"...all procedures that are officially required for an entrepreneur to start up and formally operate an industrial or commercial business. These include obtaining all necessary licenses and permits and completing any required notifications, verifications or inscriptions for the company and employees with relevant authorities."* The cost of procedures alone ignores the opportunity cost of the entrepreneur's time and the foregone profits associated with bureaucratic delay. To address this concern, the constructed cost measure adds up the official expenses and an estimate of the value of the entrepreneur's time, valuing his or her time at the country's per capita income per working day, where we assume that there are 250 working days per year, as Djankov et al (2002) assume in the construction of a similar variable. Even within OECD countries we find significant differences: while an entrepreneur in Spain needs to follow 10 different procedures, pay $3,500 in fees and wait 47 days to acquire the necessary permits, an entrepreneur in Canada can finish the process in 3 days by paying $280 in fees and completing 2 procedures.

The *cost (of procedures and time) of registering property*[17] on which the firm will

[16]What I am capturing with the second and the third costs are *only* the bureaucratic costs of dealing with these two additional activities.

[17]It includes *"...the full sequence of procedures necessary when a business purchases land and a building to transfer the property title from the seller to the buyer so that the buyer can use the property."*

be locating its production plant and the *cost (of procedures and time) of dealing with licenses while building the production plant*[18] will also be considered by potential producers. Once these three pre-production bureaucratic costs are added up, we again find large differences between countries: while the bureaucratic cost raises to almost 9 times the GDP per capita in Belgium or Greece, it is only around 2 times the GDP per capita in Denmark or Norway.

Due to data availability, the main disadvantage of these start-up cost measures is that they are posterior in time to the period covered by our dependent variables. This may raise concerns of endogeneity in our specifications in case start-up costs evolved in response to economic performance. However, RZ and Claessens and Laeven (2003) argue that measures of institutional frameworks are quite stable over time. In any case, we will also address this concern by using instrumental variables.

3.3.2 Empirical Methodology

The methodology simultaneously exploits cross-country variation in costs to start a business and cross-industry variation in the level of external finance dependence. In this generalised difference-in-difference approach, exogenous industry-country variation on these two indicators provides identification to assess the differential impact of business start-up costs on industries that vary in their degree of dependence on external sources of finance. This approach has the appealing characteristic that it allows identification even after controlling for both country and industry characteristics, which will be captured by the fixed effects. Therefore, this methodology will be less subject to criticism about an omitted variable bias or model specification. Additionally, problems of collinearity will arise less often, since for two variables to be very correlated we would need *both*

[18] It includes *"...all procedures required for a business in the construction industry to build a standardized warehouse as an example of dealing with licenses. These procedures include obtaining all necessary licenses and permits, receiving all required inspections and completing all required notifications and submitting the relevant documents (for example, building plans and site maps) to the authorities."*.

a high correlation in the sectoral component *together with* a high correlation in the country-level component.

Our first dependent variable is (the log of) the share of a sector's value added over total manufacturing value added in 1990. Our second dependent variable is (the log of) the number of establishments in each of these sectors in 1990.[19] The data comes from the Industrial Statistics of the United Nations Industrial Development Organization (INDSTAT 3 Revision 2 database). The econometric specification is:

$$\ln(\varepsilon + DepVar_{ij}) = \sum_{i=1}^{m} f_i + \sum_{j=1}^{n} f_j + \beta * ExtFin_j * StartUp_i + u_{ij} \qquad (3.18)$$

where $DepVar_{ij}$ stands for dependent variable (i.e. either share of value added or number of establishments) in sector j of country i, $ExtFin_j$ is RZ's sectoral external finance dependence variable, and $StartUp_i$ is the country-level business start-up cost. f_i and f_j are country and sector fixed effects, respectively. The disturbance term u_{ij} is clustered at the country-level, so that standard errors are corrected for a correlation between observations belonging to the same country. ε takes the value of 0.1, which is non-distorsionary given the values that our dependent variable takes[20]. We limit the sample to countries with value added data for at least 10 sectors.

Based on our theoretical prior, we expect β to be negative: a reduction in the country-level start-up cost will increase the share of value added and the number of establishments especially in sectors that have a high dependence on external sources of finance. We use three different econometric specifications: Ordinary Least Squares estimator, Tobit estimator, and Poisson Maximum Likelihood estimator, as suggested by Santos-Silva and Tenreyro (2005).

[19]Since the external finance variable is constructed for the 1980s, using the dependent variable for 1990 reduces endogeneity concerns.

[20]We have also defined ε as 0.01 and 0.001 and the results remain unchanged.

Econometric specifications in the spirit of gravity equations are usually log-linearised and estimated by OLS[21], but this practice might be inappropriate for a number of reasons. First, log-linearisation is unfeasible when the dependent variable can be zero. (This problem is often solved by adding a small positive number to all observations before taking logs, in our case $\varepsilon = 0.1$.) Second, as Santos-Silva and Tenreyro pointed out, under heteroskedasticity, the expected value of the log-linearised error will in general be correlated with the regressors, and OLS will therefore be inconsistent. The reason is that the non-linear transformation changes the properties of the error term, as the conditional expectation of $\ln u_{ij}$ depends on the shape of the conditional distribution of the error term u_{ij}. For these reasons, it might be necessary to additionally estimate the equations in its non-linear form. Santos-Silva and Tenreyro propose the Poisson pseudo-maximum likelihood estimator for this task. This estimator turns out to be consistent under very weak assumptions (mainly that the model is well specified). It also provides a natural way to deal with zero values, as it can be implemented without needing to do any logarithmic transformation.

3.4 Empirical Results

In Tables (3.4)-(3.11) we present the results for the share of value added as our dependent variable, while Tables (3.12)-(3.19) repeat exactly the same regressions with the number of establishments.

[21]Both of our dependent variables are truncated at zero. We both observe zeroes and positive values for both share of value added and number of establishments. Therefore, I link the interpretation of our regression to the gravity equation.

3.4.1 Share of Value Added

Benchmark specification

Table (3.4) presents results for the full sample of countries without including any additional control variables beyond the base specification. Our three measures of start-up costs are: (i) *narrow measure*: only includes the bureaucratic cost of opening a business, see columns (1), (4), and (7); (ii) *intermediate measure*: additionally includes the cost of registering property on which the production plant has to be built, see columns (2), (5), and (8); (iii) *broad measure*: on top of it, also incorporates the cost of dealing with licenses while this production plant is built, see columns (3), (6), and (9).[22] In columns (1)-(3) we use the OLS estimator. Columns (4)-(6) use the Poisson estimator, while columns (7)-(9) show results under the Tobit estimator. Columns (1), (4), and (7) include our *narrow measure* of start-up cost; columns (2), (5), and (8) present results with the *intermediate measure*, while columns (3), (6), and (9) incorporate the *broad measure* of start-up costs.

Already at this initial stage of analysis, we find that the coefficient interacting the country-level start-up cost with the sectoral external finance dependence is negative and statistically significant at the 1% level. In terms of coefficient values, they are similar across econometric specifications for each given measure of start-up costs.

In Table (3.5), we rerun the same regressions after eliminating the low income countries based on the World Bank Classification in 1987[23]. The coefficient of our interaction term remains significant at the 1% level, and we observe a slight increase in the coefficient values for all three measures of country-level start-up costs.

[22]To control for outliers, (i) for our narrow measure, we only include countries satisfying start-up cost<3*GDP per capita; (ii) for our intermediate measure, we only include countries satisfying start-up cost<12*GDP per capita; (iii) for our broad measure, we only include countries satisfying start-up cost<25*GDP per capita.

[23]The low income countries in our sample are Bangladesh, China, Central African Republic, India, Indonesia, Kenya, Malawi, Pakistan, Sri Lanka, Tanzania.

Main control variables

In Table (3.6), we focus on the full sample and control for a series of alternative theoretical channels previously proposed by the literature. Our first three control variables come from the Heckscher-Ohlin framework. *First,* our interaction term of interest might simply be capturing that more capital-abundant countries tend to specialise in capital-intensive industries. Our country-level measure of capital abundancy is the capital-labour ratio in 1980 from Caselli and Feyrer (2007), while our sector-level measure of capital-intensity comes from Braun (2003). It is determined as the median of the ratio of gross fixed capital formation to value added in the U.S. for the 1986-1995 period in each industry using UNIDO data.

Second, it might be reasonable to think that countries with low start-up costs are also the ones with high levels of education. We control for the fact that more skill-intensive sectors have a greater share of value added in countries with greater abundance of human capital. Our sectoral variable on human capital intensity comes from Ciccone and Papaioannou (2006). It is the average years of schooling of workers for each industry in the U.S. in 1980. Our country-level measure of human capital is taken from Barro and Lee (1993), which compute the average years of schooling in the population over 25 years of age in 1980.

And *third,* following Braun (2003), countries abundant in natural resources might be specialising in sectors that intensively use these resources. Our measure for country-level natural resource abundance comes from the World Bank's *Expanding the Measure of Wealth* publication, and includes minerals and fossil fuels, timber, nontimber forest benefits, cropland, and pastureland, net of what is labeled as protected areas. The natural resource intensity is a dummy variable that takes a value of 1 for the following industries (and 0 otherwise): wood products, except furniture; paper and products; petroleum refineries; miscellaneous petroleum and coal products; other nonmetallic mineral products; iron and steel; and nonferrous metals.

We control for two further interaction terms proposed by the more recent literature: (i) RZ and Fisman and Love (2004) find that sectors in greater need of external finance will grow faster and have greater shares of value added in countries with well-functioning financial markets, respectively. Our proxy of financial development is the ratio of domestic credit to GDP in 1980, while our sectoral variable comes from Kroszner et al. (2007) and is exactly the one we interact with start-up costs; (ii) Braun (2003) finds that countries with well-functioning financial markets tend to specialise in sectors that have a small fraction of tangible assets. Our proxy of financial development is the same as before, while our sectoral variable comes directly from Braun (2003) and is measured as the median tangibility of assets (i.e. net property, plant, and equipment over book value of assets) at the sectoral level using COMPUSTAT's annual industrial files for the 1986-1995 period.

After including these five control variables, the coefficient on our interaction term is still always significant at the 1% level: countries with lower start-up costs tend to specialise in sectors with greater need of external sources of capital. The results hold across all econometric specifications for all our measures of start-up costs. Among the control variables, the human capital interaction and the measure proposed by RZ seem to be the most relevant ones from a statistical viewpoint.

On the economic significance of our results, let us use the broad measure of start-up costs for an illustration. A move from Colombia, the country at the 75th percentile of the distribution of start-up costs, to Thailand, the country at the 25th percentile, would increase the difference between the share of value added in Transport Equipment (75th percentile of external finance) and in Beverages (25th percentile in external finance) by 14.5%[24]. For example, if in Colombia the share of value added in Transport Equipment was 2% greater than the share in Beverages, this difference would increase to 2.29% in Thailand. For comparison purposes, the median sectoral share of value added is around

[24]We obtain this number from: 0.059*(9.84-4.11)*(0.47-0.04) being approximately 0.145, i.e. 14.5%. We obtain similar results with the two other measures of start-up costs.

2%.

In Table (3.7), we present the results of Table (3.6) without including the low income countries. Two regularities arise again: our coefficients of interest become larger in absolute value and are still significant at the 1% level. An interesting feature is that the control variable regarding the tangibility of assets becomes statistically significant at the 5% level in all regressions.

Additional control variables

In Table (3.8)-(3.10) we add additional controls into our main specification. *First,* we control for labour market regulation. We expect countries with little bureaucratic frictions to have well-developed labour market institutions. Therefore, we interact the rigidity of employment index from the *Doing Business* database with the sectoral measure of external finance dependence.

Second, we control for property rights enforcement, as another measure of regulation we might be capturing with our start-up cost. The cost of enforcing contracts is the log of the number of days to resolve a payment dispute through courts and comes from Djankov et al. (2007). It is the cost of procedures and time of the judicial system in resolving a commercial dispute. We again interact this country-level measure with RZ's sectoral measure of external finance.[25]

Third, based on Braun (2003), we also control for the results not being driven by either differences in preferences or the general level of economic development. It is possible for differences in preferences to affect the composition of manufactures. Sectoral external finance might be correlated with the income elasticity of demand for the goods produced by each sector. As Braun (2003), we compute the GDP elasticity of each industry's value added based on the change in their real value added in the U.S.

[25]There is a priori no theoretical foundation for interacting these country-level terms with sectoral external finance. We still decide to use these terms, because the goal is rather to verify that our measure of bureaucratic start-up costs is not proxying for neither labour market regulations nor property rights enforcement.

between 1974 and 1994, and interact it with the log of GDP per capita.[26]

Fourth, a concern has been raised that RZ's sectoral variable might not only be capturing external finance dependence, but also sectoral growth opportunities in the U.S. during the 1980s (see Fisman and Love, 2004). Under this alternative interpretation, our interaction term of interest would partly be capturing that sectors with high growth opportunities have a greater share of value added in countries with low start-up costs. Even though this concern is more problematic when having the *growth rate* as the dependent variable[27], we still interact the country-level measure of start-up costs with the direct U.S. sectoral growth opportunities at the sectoral level. The variable captures the sectoral growth in capital during that period and comes from Ciccone and Papaioannou (2006). It was constructed with data from the NBER Manufacturing Productivity Database (Bartelsman and Gray, 1996). Our interaction term, which now should *only* be capturing the external finance component remains significant at the 1% level.

In Table (3.8) we only focus on OLS regressions. Again, we work with all three measures of start-up costs and we add one control at a time. Our result is robust to the inclusion of these additonal control variables and the economic significance is very similar to the one in previous tables. The control variable on the GDP elasticity is statistically significant at the 5% level in all three specifications, thereby showing that the level of economic significance can help to explain the manufacturing production pattern of our sample of countries.

In Table (3.9) we present results for our preferred Poisson estimator. All three measures of start-up costs remain statistically significant throughout all specifications, as is also the control variable suggested by Braun (2003). In the same spirit as in previous regressions with the Poisson estimator, both the interaction term introduced by RZ and also the human capital interaction are always very significant.

Finally, in Table (3.10) we use the Tobit estimator. Our interaction term is always

[26]For a more extended motivation for this interaction term, see Braun (2003).

[27]Recall that our dependent variables are in levels and not in differences.

significant at the 1% level, and the main change with respect to the previous two tables is that the property rights enforcement interaction is significant at the 5% level.

Instrumental Variables

The cost of starting a business, calculated for the first time for the year 2005, in the *Doing Business* dataset could be the outcome of the production structure of the economy, thereby leading to endogeneity concerns. Under this channel, countries that have historically been specialising in sectors with a high demand for external sources of finance might have put more effort into reducing their cost of starting new businesses.

To make sure that this mechanism is not the driver of our results, in Table (3.11) I instrument the cost of opening a business by two measures which are related to the legal environment and that have been used previously by the literature[28]: (i) legal origin; (ii) rule of law. A variety of papers by Djankov, LaPorta, and Shleifer argue that a country's legal origin is a historically predetermined variable with long-lasting effects on regulation policies. Legal origin can therefore be seen as an exogenous variable in our framework. Additionally, the rule of law variable is an assessment of the law and order tradition in the country. Both measures can be found in LaPorta et al. (1998).[29]

In columns (1)-(3) we include the standard controls from Table (3.6) with the full sample; again, we look at what happens with all three measures of start-up costs. In columns (4)-(6) we redo the same regressions but eliminate the low income countries based on the World Bank classification. Our interaction term of interest remains statistically significant at the 1% or 5% level, while the financial interaction term proposed by RZ loses significance. In terms of economic significance, the coefficients are slightly higher than in Table (3.6).

[28]For the choice of the same or similar instrumental variables, see RZ, Guiso et al. (2004), Ciccone and Papaioannou (2006).

[29]The instruments satisfy the test of overidentifying restrictions based on Wooldridge's (1995) robust score test. The results of this test are available upon request.

3.4.2 Number of Establishments

We found that reductions in country-level start-up costs lead to shifts in the share of value added towards sectors with high levels of external finance. Increases in value added can either occur through the intensive margin (growth of already existing firms) or the extensive margin (creation of new firms). Given that start-up costs especially affect smaller and younger firms, we will analyse whether a significant part of the increase in value added is due to the extensive margin. In Tables (3.12)-(3.19) we rerun all the regressions with the number of establishments as our dependent variables. To avoid becoming repetitive, the following subsections briefly highlight the main differences with respect to the previous regressions.

Benchmark specification

Table (3.12) show the results for the full sample of countries without any additional control variables beyond the base specification. The coefficient interacting country-level start-up costs and sectoral external finance is negative and statistically significant at the 1% level for all regressions. Interestingly, the coefficient values across all specifications are similar to the ones we obtained in Table (3.4), the equivalent table with share of value added as the dependent variable.

Table (3.13) repeats the same exercise without including the low income countries based on the World Bank Classification in 1987. The coefficient remains statistically significant mostly at the 1% level, and we again observe a slight increase in the absolute value of the coefficient.

Main control variables

Table (3.14) presents full sample results including the same five control variables that were already introduced in Table (3.6). The statistical significance of our results remains unchanged across econometric specifications and measures of start-up costs.

Regarding the economic significance of our results with number of establishments as the dependent variable, a move from Colombia, the country at the 75th percentile of start-up costs, to Thailand, the country at the 25th percentile, would increase the difference between the number of establishments in Transport Equipment (75th percentile in external finance) and in Beverages (25th percentile) by 22%.[30]

In Table (3.15) we repeat the regressions of Table (3.14) after eliminating the low income countries and the statistically significance is unchanged. The coefficient values become larger in absolute value.

Additional control variables

In Tables (3.16)-(3.18) we also include the additional control variables into our main specification. These controls were already described in length for Tables (3.8)-(3.10). The results again are robust to the inclusion of these control variables. Our coefficient of interest is mostly significant at the 1% level.

Interestingly, three of the additional control variables are very significant across the different measures of country-level start-up costs: columns (4) to (6) tell us that countries with better property rights enforcement have relatively more establishments in sectors with greater demand of external capital. In columns (7) to (9) our interaction term on the level of economic development and sectoral elasticity of demand is again very significant. Columns (10) to (12) reveal that countries with low start-up costs tend to have relatively more establishments in sectors with high sectoral growth opportunities.

Instrumental variables

In Table (3.19) we instrument our start-up cost measures with legal origin and rule of law, as we did in Table (3.11). Columns (1)-(3) present results for the full sample with the five main controls, while columns (4)-(6) eliminate low income countries. The

[30]We obtain this number from: 0.091*(9.84-4.11)*(0.47-0.04) being approximately 0.22, i.e. 22%.

coefficient interacting start-up costs and external finance remains statistically significant at the 1% or 5% level, and the magnitude of the economic significance increases considerably.[31]

3.5 Conclusion

In this paper we attempt to understand the effects of a reduction in the bureaucratic start-up cost to open a new business on an economy's industrial production structure. We introduce a theoretical framework based on Holmstrom and Tirole (1997) that gives us the following theoretical prediction: reductions in country-level bureaucratic start-up costs for entrepreneurs shift the production structure towards sectors with greater need of external sources of capital. We confirm our hypothesis in the empirical part of the paper: using different econometric specifications and various measures of start-up costs, we find that countries with lower bureaucratic costs tend to have a greater share of value added and number of establishments in sectors that rely more on external sources of capital.

Understanding how these different barriers to entrepreneurship (i.e. costly access to external capital and bureaucratic start-up costs) interact can be helpful not only to better assess how entrepreneurs react to changing incentives, but also to identify the winners and losers of deregulation processes that are occurring at a large scale in many countries.

[31] The instruments again satisfy the test of overidentifying restrictions proposed by Wooldridge (1995). The results are available upon request.

Bibliography

[1] Acs, Zoltan J., and D. B. Audretsch (1988). "Innovation in Large and Small Firms: An Empirical Analysis". American Economic Review, 78, pp. 678-690

[2] Acs, Zoltan J., and D. B. Audretsch (1990). *Innovation and Small Firms* (Cambridge; MA; MIT Press, 1990)

[3] Aghion, P., T. Fally, and S. Scarpetta (2007). "Credit Constraints as a Barrier to Entry and Post-entry Growth of Firms: Lessons from Firm-level Cross Country Panel Data". Economic Policy, vol. 22, no. 52, pp. 731-779

[4] Bartelsman, E. J., and W. Gray (1996). "The NBER Productivity Manufacturing Database". National Bureau of Economic Research (Cambridge, MA). Technical Working Paper No. 205, October. (Downloadable from http://www.nber.org/nberces/nbprod96.htm)

[5] Barro, R.J. and J. Lee (1993): "International Comparisons of Educational Attainment". Journal of Monetary Economics, 32, 363-394.

[6] Bernanke, B, and M. Gertler (1989). "Agency Costs, Net Worth, and Business Fluctuations." American Economic Review, March 1989, 79(1), pp.14-31

[7] Blackburn, K. and V. Hung (1998). "A Theory of Financial Intermediation and Growth". Economica, 65, pp. 107-24.

[8] Beck, T., A. Demirguc-Kunt, and R. Levine (2003). "Law, Endowments, and Finance". Journal of Financial Economics, 70(2), pp.137-81.

[9] Boyd, J. and E. Prescott (1986). "Financial Intermediary-Coalitions". Journal of Economic Theory, 3, pp. 211-232.

[10] Braun, M. (2003). "Financial Contractability and Asset Hardness". Mimeo, Harvard University.

[11] Caselli, F. and J. Feyrer (2007). "The Marginal Product of Capital". Quarterly Journal of Economics, 122(2), pp. 535-568.

[12] Ciccone, A. and E. Papaioannou (2007). "Red Tape and Delayed Entry". Journal of the European Economic Association. vol. 5, issue 2-3, pp. 444-58.

[13] Ciccone, A. and E. Papaioannou (2006). "Adjustment to Target Capital, Finance, and Growth". Work in progress.

[14] Claessens, S. and L. Laeven (2003). "Financial Development, Property Rights, and Growth". The Journal of Finance, 58(6), pp. 2401-36.

[15] Deidda, L. (2006). "Interaction Between Economic and Financial Development". Journal of Monetary Economics, 53, pp. 233-48.

[16] Demirguc-Kunt, A. and V. Maksimovic (1996). "Financial Constraints, Uses of Funds and Firm Growth: An International Comparison". World Bank Working Paper Series, 1996.

[17] Demirguc-Kunt, A. and V. Maksimovic (1998). "Law, Finance, and Firm Growth". Journal of Finance, 53, 2107-37.

[18] Djankov, S., R. LaPorta, F. Lopez-de-Silanes, and A. Shleifer (2002). "The Regulation of Entry". Quarterly Journal of Economics, 117(1), 1-37.

[19] Djankov, S., C. McLiesh, and A. Shleifer (2007). "Private Credit in 129 countries". Journal of Financial Economics, 84, pp. 299-329

[20] Doing Business (2006). The World Bank Group. http://www.doingbusiness.org

[21] Fisman, R. and I. Love (2004). "Financial Development and Growth in the Short- and Long-Run". Work in progress.

[22] Fisman, R. and V. Sarria-Allende (2004). "Regulation of Entry and the Distortion of Industrial Organization". National Bureau of Economic Research (Cambridge, MA), Working Paper No. 10929, November.

[23] Fonseca, R. and N. Utrero (2003). "Financial Development, Labor And Market Regulations And Growth". Universitat Autonoma de Barcelona, Document de Treball num. 05/3.

[24] Fonseca, R. and N. Utrero (2004). "Do Market Regulation and Financial Imperfections Affect Firm Size? New Empirical Evidence". Centre For Studies in Economics and Finance, Working Paper No. 119.

[25] Goldsmith, R.W. (1969). *Financial Structure and Development*, New Haven: Yale University Press

[26] Greenwood, J. and B. Jovanovic (1990). "Financial Development, Growth, and the Distribution of Income". Journal of Political Economy, 98, pp. 1076-107.

[27] Guiso, L., T. Jappelli, M. Padula, and M. Pagano (2004). "Financial Market Integration and Economic Growth in the EU". Economic Policy, October 2004, pp. 523-577.

[28] Hermes, N. and R. Lensink (eds.) (1996), *Financial Development and Economic Growth: Theory and Experiences from Developing Countries*, London: Routledge.

[29] Holmstrom, B. and J. Tirole (1997). "Financial Intermediation, Loanable Funds, and the Real Sector". Quarterly Journal of Economics, vol. 112 (3), p. 663-91.

[30] King, R.G. and R. Levine (1993a). "Finance and Growth: Schumpeter might be right." Quarterly Journal of Economics, 109, pp. 717-37.

[31] King, R.G. and R. Levine (1993b). "Finance, Entrepreneurship, and Growth: theory and evidence". Journal of Monetary Economics, 32, vol. pp. 513-42.

[32] Kiyotaki, N. and J. Moore (1997). "Credit Cycles". Journal of Political Economy, vol. 105(2), pp. 211-48

[33] Kiyotaki, N. (1998). "Credit and Business Cycles". The Japanese Economic Review, vol. 49, no. 1

[34] Klapper, L., L. Laeven and R. Rajan (2006). "Entry Regulation as a Barrier to Entrepreneurship". Journal of Financial Economics, 82, pp. 591-629.

[35] Kroszner, R.S., L. Laeven, and D. Klingebiel (2007). "Banking Crises, Financial Dependence, and Growth". Journal of Financial Economics, 84, pp. 187-228

[36] Krugman, P. R. (1979). "Increasing Returns, Monopolistic Competition, and International Trade," Journal of International Economics, 9, pp. 469-479.

[37] Kumar, K. B., R. G. Rajan, and L. Zingales (2002). "What Determines Firm Size?". Work in Progress.

[38] LaPorta, R., F. Lopez-de-Silanes, A. Shleifer, and R.W. Vishny (1997). "Legal Determinants of External Finance". Journal of Finance, 52, pp. 1131-1150.

[39] LaPorta, R., F. Lopez-de-Silanes, A. Shleifer, and R.W. Vishny (1998). "Law and Finance". Journal of Political Economy 106, pp. 1113-1155.

[40] Levine, R. (1997). "Financial Development and Growth". Journal of Economic Literature, 35, pp. 688-726.

[41] Levine, R. and S. Zervos (1998). "Stock Markets and Economic Growth". American Economic Review, June 1998, 88(3), pp. 537-58.

[42] Levine, R., N. Loayza, and T. Beck (2000). "Financial Intermediation and Growth: Causality and Causes". Journal of Monetary Economics, 46, pp. 31-77

[43] McKinnon, R.I. (1973). *Money and Capital in Economic Development*, Washington DC: The Brookings Institution.

[44] Mishkin (2007), F. S. "Globalization and Financial Development" (Central Bank Articles and Speeches, http://www.bis.org/review/r070502f.pdf). Remark to the New Perspectives on Financial Globalization Conference, International Monetary Fund, Washington DC, 26 April 2007

[45] Myers, S. C. and N. S. Majluf (1984). "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not have". Journal of Financial Economics, 13 (2), pp. 187-221.

[46] Rajan, R. and L. Zingales (1998). "Financial Dependence and Growth". American Economic Review 88, p. 559-86.

[47] Rousseau, P., and P. Wachtel (1998). "Financial Intermediation and Economic Performance: Historical Evidence From Five Industrialized Countries". Journal of Money, Credit, and Banking, 30, pp. 658-678

[48] Saint-Paul, G. (1992). "Technological Choice, Financial Markets and Economic Development". European Economic Review, 36, pp. 763-81.

[49] Santos-Silva, J.M.C. and S. Tenreyro (2005). "The Log of Gravity". The Review of Economics and Statistics, vol. 88, no. 4, pp. 641-658.

[50] Shaw, E. S. (1973). *Financial Deepening in Economic Development.* New York: Oxford University Press, 1973.

[51] Stiglitz, J. and A. Weiss (1981). "Credit Rationing in Markets with Imperfect Information". American Economic Review, 71 (3), pp. 393-410.

[52] United National Statistical Division (1993). Industrial Statistics Yearbook, Vol. 1. New York: United Nations.

[53] Wooldridge, J. M. (1995). "Score diagnostics for linear models estimated by two stage least squares" in Advances in Econometrics and Qualitative Economics: Essays in Honor of Professor C. Phillips. Oxford: Blackwell.

[54] World Bank, 1997. Expanding the Measure of Wealth: Indicators of Environmentally Sustainable Development. World Bank, Washington.

[55] Zilibotti, F. (1994). "Endogenous Growth and Intermediation in an 'Archipielago' Economy". Economic Journal, 103, pp. 462-73.

Description of variables

1. *ExtFin:* Industry-dependence on external finance. The median of the ratio of capital expenditure minus cashflow to capital expenditure for U.S. firms averaged over the 1980-1989 period. Source: Kroszner et al. (2007).

2. *Start-up:* Country-level start-up cost. Combination of the cost of time and procedures to start a business, according to the definition of the *Doing Business* report. We assume 250 working days per year and data is for 2005. Source: World Bank-Doing Business (2006).

3. *Start-up+Reg:* In addition to the cost of *Start-up* previously described, we add the cost of time and procedures of registering property, according to the definition of the *Doing Business* report. We assume 250 working days per year and data is for 2005. Source: World Bank-Doing Business (2006).

4. *Start-up+Reg+Deal:* In addition to the cost of *Start-up+Reg* previously described, we add the cost of time and procedures of dealing with licenses while building the production plant, according to the definition of the *Doing Business* report. We assume 250 working days per year and data is for 2005. Source: World Bank-Doing Business (2006).

5. *Number of establishments:* Log number of establishments in industry i in country n in 1990. Source: Industrial Statistics of the United Nations Industrial Development Organization (INDSTAT 3 Revision 2 database).

6. *Share of value added:* Share of value added in industry i in country n over total value added in country n in 1990. Source: Industrial Statistics of the United Nations Industrial Development Organization (INDSTAT 3 Revision 2 database).

7. *Labour rigidity:* Rigidity of employment index. It is defined as the average of the difficulty of hiring a new worker, restrictions on expanding or contracting the

number of working hours, and difficulty and expense of dismissing a redundant worker. Source: World Bank-Doing Business (2006).

8. *Enforcement (Days):* Number of days to resolve a payment dispute through courts. Source: Djankov et al. (2007).

9. *HK intensity:* Average years of schooling of workers for each industry in the U.S. in 1980. Source: Ciccone and Papaioannou (2006).

10. *Total years of schooling:* Average years of schooling in the population over 25 years of age in 1980. Source: Barro and Lee (1993).

11. *Growth opp.:* Annual change of log real capital stock in each industry in the U.S. in 1980-1989 period. Source: Ciccone and Papaioannou (2006).

12. *GDP:* GDP per capita in 1980. Source: World Bank-World Development Indicators

13. *Private Credit:* Domestic credit to the private sector over GDP in 1980. Source: IMF-IFS

14. *Legal Origin:* Colonial origin of a country's legal system. Source: Rajan and Zingales (1998)

15. *Rule of Law:* Index for the efficiency and integrity of the legal system produced by Business International Corporation (a country-risk rating agency). Source: Rajan and Zingales (1998)

16. *KL ratio:* Capital-labour ratio in 1980. Source: Caselli and Feyrer (2007)

17. *K intensity:* Median of the ratio of gross fixed capital formation to value added in the U.S. for the 1986-1995 period in each industry using UNIDO data. Source: Braun (2003)

18. *Nat. Res.:* Minerals and fossil fuels, timber, nontimber forrest benefits, cropland and pastureland, net of what is labeled as protected areas. Source: World Bank's *Expanding the Measure of Wealth (1997),* Braun (2003)

19. *NatRes_intensity:* Dummy variable that takes a value of 1 for the following industries (and 0 otherwise): wood products, except furniture; paper and products; petroleum refineries; miscellaneous petroleum and coal products; other nonmetallic mineral products; irin and steel; and nonferrous metals. Source: Braun (2003)

20. *Tangibility:* Median tangibility of assets (i.e. net property, plant, and equipment over book value of assets) at the sectoral level using Compustat's annual industrial files for the 1986-1995 period. Source: Braun (2003)

21. *Elasticity of demand:* GDP elasticity of each industry's value added based on the change in their real value added in the U.S. between 1974 and 1994. Source: Braun (2003)

Figure 3.1. Correlation between External Finance and Start-up Cost (Start-up, Registering Property, Dealing with Licenses). beta = -1.1. t-stat = -22.15

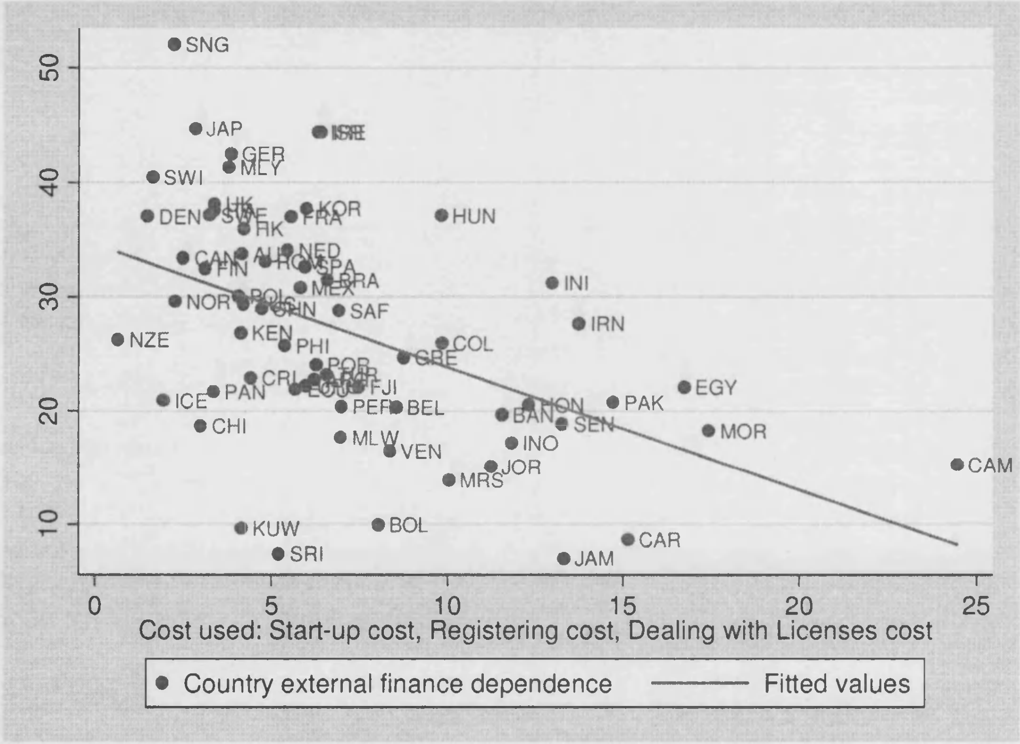

Figure 3.2. Correlation between External Finance and Start-up Cost
(Start-up, Registering Property). beta = -1.6. t-stat = -18.83

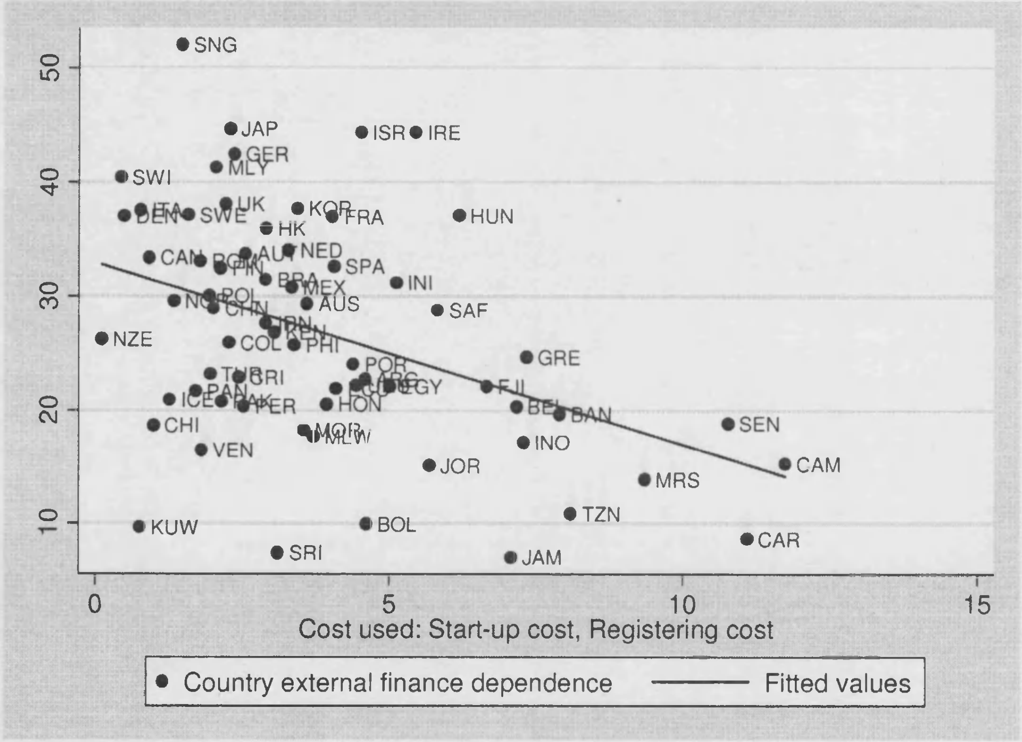

Figure 3.3. Correlation between External Finance and Start-up Cost (Start-up). beta = -11.4. t-stat = -28.2

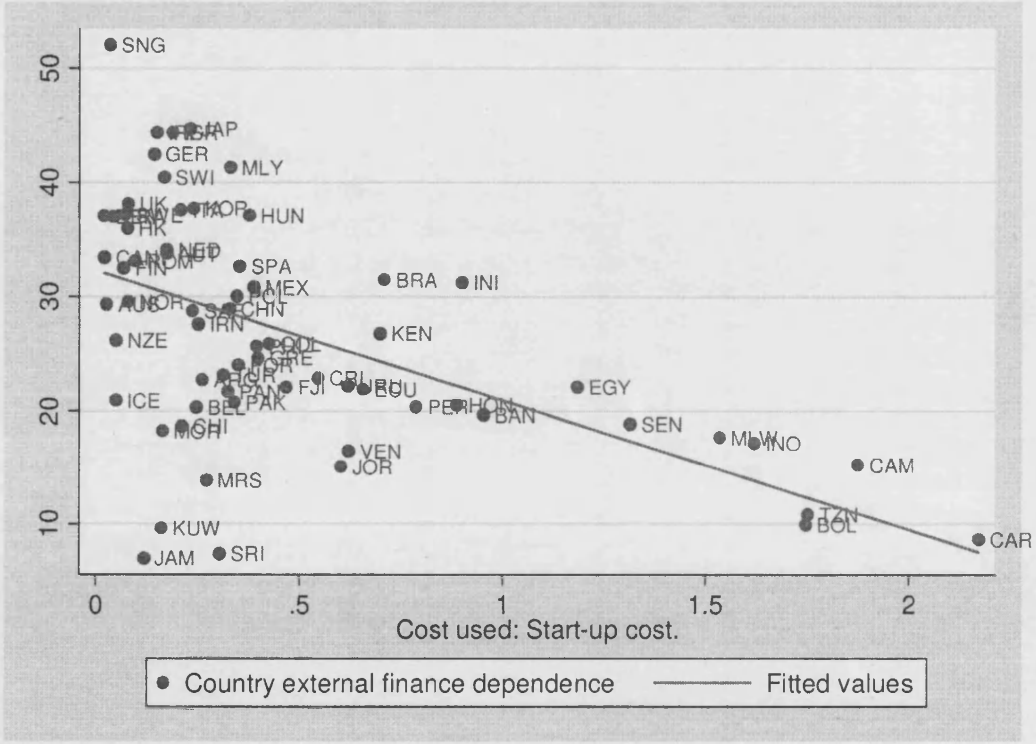

Table 3.1. Country-level start-up costs, measured as a percentage of income per capita (Part 1)

Country	StartUp Cost	Country	StartUp Cost & Registering Property	Country	StartUp Cost & Registering Property & Dealing with Licenses
Denmark	0.020	New Zealand	0.108	New Zealand	0.661
Canada	0.021	Switzerland	0.431	Denmark	1.481
Australia	0.027	Denmark	0.488	Switzerland	1.631
Singapore	0.035	Kuwait	0.762	Iceland	1.929
France	0.044	Italy	0.767	Singapore	2.227
Iceland	0.049	Canada	0.911	Norway	2.260
New Zealand	0.050	Chile	0.985	Mongolia	2.308
Finland	0.068	Iceland	1.265	Canada	2.489
Sweden	0.071	Norway	1.333	Japan	2.832
Hong Kong	0.078	Mongolia	1.336	Chile	3.001
Norway	0.079	Bulgaria	1.450	Finland	3.110
UK	0.079	Singapore	1.471	Sweden	3.239
Romania	0.097	Sweden	1.579	Panama	3.355
Jamaica	0.119	Panama	1.700	UK	3.375
Mongolia	0.142	Romania	1.777	Italy	3.376
Germany	0.143	Venezuela	1.803	Malaysia	3.782
Ireland	0.149	Poland	1.934	Germany	3.845
Tunisia	0.156	Turkey	1.949	Poland	4.053
Kuwait	0.162	China	2.006	Thailand	4.112
Morocco	0.164	Malaysia	2.051	Kenya	4.120
Switzerland	0.167	Finland	2.124	Austria	4.147
Austria	0.173	Pakistan	2.136	Kuwait	4.147
Netherlands	0.174	UK	2.213	Australia	4.204
Israel	0.189	Colombia	2.267	Hong Kong	4.215
Thailand	0.193	Japan	2.287	Costa Rica	4.413
Italy	0.209	Germany	2.357	China	4.718
Chile	0.211	Costa Rica	2.430	Romania	4.818
Bulgaria	0.224	Albania	2.463	Slovenia	5.020
Japan	0.231	Peru	2.520	Sri Lanka	5.214
South Africa	0.238	Austria	2.551	Philippines	5.375
Korea	0.240	Iran	2.895	Netherlands	5.445
Belgium	0.247	Brazil	2.897	Bulgaria	5.549
Iran	0.251	Slovenia	2.905	France	5.549
Argentina	0.262	Hong Kong	2.910	Ecuador	5.687
Mauritius	0.272	Kenya	3.040	Mexico	5.812
Sri Lanka	0.304	Sri Lanka	3.106	Spain	5.932
Turkey	0.313	Netherlands	3.282	Uruguay	5.967
Zambia	0.321	Mexico	3.334	Korea	6.000

Table 1 reports values for the cost and time of starting up a business, where it is assumed that a person works 250 days per year. The data comes from WB-WDI Doing Business for 2005. The first measure only includes the cost of procedures and time of opening a business. The second measure adds to it the cost of procedures and time of registering property. Finally, the third measure adds to it the cost of procedures and time of dealing with licenses while building a production plant. We restrict our sample of countries to StartUp<3*GDP per capita, StartUp & Registering Property<12*GDP per capita, and StartUp & Registering Property & Dealing with Licenses<25*GDP per capita in order to control for outliers. Therefore, we eliminate Syria for StartUp * Registering Property, and we eliminate Tanzania and Burundi for StartUp & Registering Property & Dealing with Licenses.

Table 3.2. Country-level start-up costs, measured as a percentage of income per capita (Part 2)

Country	StartUp Cost	Country	StartUp Cost & Registering Property	Country	StartUp Cost & Registering Property & Dealing with Licenses
Panama	0.324	Thailand	3.351	Albania	6.113
China	0.328	Philippines	3.377	Argentina	6.219
Malaysia	0.329	Korea	3.434	Portugal	6.267
Croatia	0.330	Tunisia	3.434	Israel	6.326
Pakistan	0.340	Morocco	3.542	Ireland	6.411
Slovenia	0.341	Australia	3.597	Turkey	6.564
Poland	0.346	Malawi	3.708	Brazil	6.581
Portugal	0.350	Nepal	3.891	South Africa	6.909
Spain	0.353	Honduras	3.933	Malawi	6.975
Hungary	0.376	France	4.026	Peru	6.987
Mexico	0.388	Spain	4.053	Tunisia	7.450
Philippines	0.395	Ecuador	4.091	Fiji	7.481
Greece	0.398	Portugal	4.382	Nepal	7.626
Colombia	0.425	Uruguay	4.433	Bolivia	8.046
Fiji	0.468	Israel	4.515	Venezuela	8.379
Albania	0.475	Argentina	4.588	Belgium	8.552
Syria	0.533	Bolivia	4.616	Greece	8.763
Costa Rica	0.546	Egypt	5.007	Hungary	9.831
Jordan	0.603	India	5.119	Colombia	9.840
Uruguay	0.619	Zambia	5.401	Mauritius	10.057
Venezuela	0.621	Ireland	5.451	Jordan	11.242
Ecuador	0.657	Jordan	5.691	Bangladesh	11.555
Kenya	0.698	South Africa	5.830	Indonesia	11.834
Brazil	0.709	Hungary	6.188	Honduras	12.325
Ethiopia	0.779	Ethiopia	6.203	India	12.984
Nepal	0.783	Croatia	6.654	Senegal	13.270
Peru	0.788	Fiji	6.660	Jamaica	13.314
Honduras	0.889	Jamaica	7.085	Iran	13.747
India	0.901	Belgium	7.175	Pakistan	14.715
Bangladesh	0.954	Indonesia	7.289	Egypt	16.730
Egypt	1.185	Greece	7.340	Morocco	17.438
Senegal	1.315	Bangladesh	7.906	Syria	20.003
Malawi	1.536	Tanzania	8.097	Croatia	20.133
Indonesia	1.621	Mauritius	9.362	Zambia	22.773
Bolivia	1.748	Senegal	10.771	Ethiopia	24.200
Tanzania	1.753	Cameroon	11.748	Cameroon	24.466
Cameroon	1.876	Burundi	12.005	Tanzania	50.451
Burundi	2.179	Syria	15.869	Burundi	120.618

Table 1 reports values for the cost and time of starting up a business, where it is assumed that a person works 250 days per year. The data comes from WB-WDI Doing Business for 2005. The first measure only includes the cost of procedures and time of opening a business. The second measure adds to it the cost of procedures and time of registering property. Finally, the third measure adds to it the cost of procedures and time of dealing with licenses while building a production plant. We restrict our sample of countries to StartUp<3*GDP per capita, StartUp & Registering Property<12*GDP per capita, and StartUp & Registering Property & Dealing with Licenses<25*GDP per capita in order to control for outliers. Therefore, we eliminate Syria for StartUp * Registering Property, and we eliminate Tanzania and Burundi for StartUp & Registering Property & Dealing with Licenses.

Table 3.3. Main industry-level variables

ISIC	Industry Name	External Finance	Human K intensity	K intensity	Natural Resources	Tangibility
314	Tobacco	-0.45	11.51	0.02	0	0.22
361	Pottery, china, earthenware	-0.15	11.24	0.05	0	0.07
323	Leather products	-0.14	10.14	0.03	0	0.09
324	Footwear, except rubber or plastic	-0.08	10.26	0.02	0	0.12
372	Non-ferrous metals	0.01	11.55	0.10	1	0.38
322	Wearing apparel, except footwear	0.03	10.19	0.02	0	0.13
353	Petroleum refineries	0.04	13.20	0.20	1	0.67
369	Other non-metallic products	0.06	11.66	0.07	1	0.42
313	Beverages	0.08	11.97	0.06	0	0.28
371	Iron and steel	0.09	11.43	0.10	1	0.46
311	Food Products	0.14	11.26	0.06	0	0.38
341	Paper and products	0.17	11.69	0.13	1	0.56
321	Textiles	0.19	10.40	0.07	0	0.37
342	Printing and publishing	0.20	12.79	0.05	0	0.30
355	Rubber products	0.23	11.73	0.07	0	0.38
332	Furniture, except metal	0.24	10.76	0.04	0	0.26
381	Fabricated metal products	0.24	11.58	0.05	0	0.28
351	Industrial chemicals	0.25	12.70	0.12	0	0.41
331	Wood products, except furniture	0.28	10.79	0.07	1	0.38
354	Misc. petroleum and coal products	0.33	11.92	0.07	1	0.30
384	Transport equipment	0.36	12.35	0.07	0	0.25
390	Other manufactured products	0.47	11.35	0.04	0	0.19
362	Glass and products	0.53	11.48	0.09	0	0.33
382	Machinery, except electrical	0.60	12.27	0.06	0	0.18
352	Other chemicals	0.75	13.03	0.06	0	0.20
383	Machinery, electric	0.95	12.36	0.08	0	0.21
385	Prof and scient equipment	0.96	12.52	0.05	0	0.15
356	Plastic products	1.14	11.68	0.09	0	0.34

Table 2 reports values for each 3-digit ISIC manufacturing industry for human capital intensity (Human K intensity), capital growth (Capital growth), sales growth (Sales growth), value added growth (VA growth), and external-finance dependence (External Finance). These measures are all based on U.S. data.

Table 3.4. Share in value added. Full sample

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Share VA)			Share VA			ln(ε+Share VA)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.894			-1.207			-1.006		
	[0.120]***			[0.148]***			[0.125]***		
Start-up+Reg * ExtFin		-0.130			-0.154			-0.126	
		[0.034]***			[0.039]***			[0.024]***	
Start-up+Reg+Deal * ExtFin			-0.088			-0.092			-0.083
			[0.016]***			[0.020]***			[0.011]***
Observations	1946	1909	1892	1946	1909	1892	1946	1909	1892
R-squared	0.54	0.52	0.53						
Pseudo R-squared				0.33	0.31	0.32	0.26	0.26	0.26

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.5. Share in value added. No low income countries.

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Share VA)			Share VA			ln(ε+Share VA)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.997			-1.429			-1.178		
	[0.179]***			[0.202]***			[0.157]***		
Start-up+Reg * ExtFin		-0.110			-0.129			-0.106	
		[0.042]**			[0.046]***			[0.026]***	
Start-up+Reg+Deal * ExtFin			-0.087			-0.092			-0.079
			[0.017]***			[0.022]***			[0.012]***
Observations	1682	1645	1655	1682	1645	1655	1682	1645	1655
R-squared	0.54	0.53	0.54						
Pseudo R-squared				0.34	0.32	0.33	0.27	0.26	0.27

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.6. Share in value added. Full sample

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Share VA)			Share VA			ln(ε+Share VA)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.572			-0.798			-0.623		
	[0.149]***			[0.191]***			[0.152]***		
Start-up+Reg * ExtFin		-0.087			-0.114			-0.095	
		[0.032]***			[0.034]***			[0.026]***	
Start-up+Reg+Deal * ExtFin			-0.059			-0.063			-0.060
			[0.017]***			[0.020]***			[0.015]***
KL ratio * K-intensity	0.014	0.014	0.016	-0.017	-0.018	-0.017	0.003	0.004	0.005
	[0.024]	[0.024]	[0.024]	[0.023]	[0.023]	[0.023]	[0.019]	[0.019]	[0.019]
Schooling * HK intensity	0.036	0.039	0.035	0.039	0.042	0.040	0.041	0.044	0.041
	[0.017]**	[0.017]**	[0.017]**	[0.013]***	[0.014]***	[0.014]***	[0.011]***	[0.011]***	[0.011]***
Nat. Res. * NatRes-intens.	0.014	0.014	0.015	0.014	0.014	0.014	0.017	0.016	0.017
	[0.008]*	[0.008]*	[0.008]*	[0.008]*	[0.009]	[0.008]*	[0.006]***	[0.006]**	[0.006]***
Private Credit * ExtFin	0.985	1.249	1.026	1.091	1.369	1.222	0.961	1.222	1.007
	[0.414]**	[0.367]***	[0.392]**	[0.388]***	[0.344]***	[0.365]***	[0.302]***	[0.285]***	[0.300]***
Private Credit * Tangibility	-1.131	-1.168	-1.153	-1.879	-1.771	-1.798	-1.359	-1.360	-1.370
	[0.807]	[0.801]	[0.806]	[1.038]*	[1.039]*	[1.041]*	[0.797]*	[0.801]*	[0.800]*
Observations	1433	1433	1433	1433	1433	1433	1433	1433	1433
R-squared	0.57	0.57	0.57						
Pseudo R-squared				0.33	0.33	0.32	0.3	0.3	0.3

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.7. Share in value added. No low income countries.

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Share VA)			Share VA			ln(ε+Share VA)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.732			-1.109			-0.839		
	[0.185]***			[0.230]***			[0.180]***		
Start-up+Reg * ExtFin		-0.087			-0.116			-0.099	
		[0.039]**			[0.039]***			[0.028]***	
Start-up+Reg+Deal * ExtFin			-0.065			-0.075			-0.064
			[0.018]***			[0.023]***			[0.016]***
KL ratio * K-intensity	-0.010	-0.010	-0.008	-0.036	-0.039	-0.037	-0.016	-0.016	-0.014
	[0.024]	[0.024]	[0.024]	[0.025]	[0.025]	[0.025]	[0.019]	[0.019]	[0.019]
Schooling * HK intensity	0.034	0.039	0.033	0.031	0.036	0.033	0.037	0.042	0.038
	[0.021]	[0.021]*	[0.021]	[0.015]**	[0.015]**	[0.015]**	[0.012]***	[0.012]***	[0.012]***
Nat. Res. * NatRes-intens.	0.010	0.009	0.010	0.011	0.010	0.010	0.013	0.012	0.013
	[0.007]	[0.007]	[0.007]	[0.008]	[0.009]	[0.008]	[0.006]**	[0.006]*	[0.006]**
Private Credit * ExtFin	0.815	1.112	0.820	0.929	1.229	1.008	0.722	1.030	0.768
	[0.422]*	[0.371]***	[0.403]**	[0.391]**	[0.338]***	[0.361]***	[0.297]**	[0.286]***	[0.300]**
Private Credit * Tangibility	-1.574	-1.613	-1.579	-2.498	-2.440	-2.454	-1.862	-1.867	-1.880
	[0.744]**	[0.743]**	[0.750]**	[1.119]**	[1.144]**	[1.130]**	[0.800]**	[0.807]**	[0.804]**
Observations	1196	1196	1196	1196	1196	1196	1196	1196	1196
R-squared	0.60	0.59	0.60						
Pseudo R-squared				0.34	0.33	0.33	0.32	0.31	0.31

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.8. Share in value added. Full sample. Additional controls. Ordinary Least Squares.

Dependent variable: ln(ε+Share VA)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Start-up * ExtFin	-0.497			-0.597			-0.597			-0.659		
	[0.165]***			[0.157]***			[0.150]***			[0.145]***		
Start-up+Reg * ExtFin		-0.071			-0.090			-0.090			-0.103	
		[0.035]**			[0.034]**			[0.032]***			[0.031]***	
Start-up+Reg+Deal * ExtFin			-0.051			-0.056			-0.062			-0.064
			[0.016]***			[0.017]***			[0.017]***			[0.018]***
KL ratio * K-intensity	0.015	0.015	0.016	0.008	0.008	0.009	0.000	0.001	0.001	0.007	0.010	0.012
	[0.024]	[0.024]	[0.024]	[0.024]	[0.024]	[0.024]	[0.024]	[0.023]	[0.023]	[0.023]	[0.023]	[0.023]
Schooling * HK intensity	0.035	0.037	0.033	0.037	0.038	0.035	0.053	0.055	0.052	0.044	0.044	0.039
	[0.017]**	[0.017]**	[0.017]*	[0.018]**	[0.018]**	[0.018]*	[0.018]***	[0.018]***	[0.018]***	[0.017]**	[0.016]***	[0.017]**
Nat. Res. * NatRes-intens.	0.015	0.014	0.015	0.013	0.013	0.013	0.013	0.013	0.013	0.015	0.014	0.015
	[0.008]*	[0.008]*	[0.008]*	[0.008]*	[0.008]	[0.008]*	[0.007]*	[0.008]	[0.008]*	[0.008]*	[0.008]*	[0.008]*
Private Credit * ExtFin	1.049	1.289	1.082	0.980	1.185	0.983	1.127	1.399	1.167	0.984	1.245	1.024
	[0.392]***	[0.354]***	[0.371]***	[0.453]**	[0.412]***	[0.426]**	[0.425]**	[0.379]***	[0.400]***	[0.409]**	[0.364]***	[0.389]**
Private Credit * Tangibility	-1.124	-1.156	-1.139	-1.279	-1.304	-1.293	-1.086	-1.126	-1.107	-1.143	-1.170	-1.156
	[0.805]	[0.801]	[0.804]	[0.791]	[0.785]	[0.789]	[0.805]	[0.799]	[0.804]	[0.803]	[0.796]	[0.804]
Labour rigidity * ExtFin	-0.005	-0.006	-0.006									
	[0.005]	[0.005]	[0.004]									
Ln(Enforcement Days) * ExtFin				0.028	-0.065	-0.051						
				[0.168]	[0.163]	[0.161]						
Ln(GDPpc) * Elasticity							-0.107	-0.105	-0.110			
							[0.049]**	[0.050]**	[0.050]**			
Start-up * Growth opp.										4.154		
										[3.888]		
Start-up+Reg * Growth opp.											0.700	
											[0.801]	
Start-up+Reg+Deal * Growth opp.												0.239
												[0.436]
Observations	1433	1433	1433	1406	1406	1406	1433	1433	1433	1433	1433	1433
R-squared	0.57	0.57	0.57	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.57	0.57

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.9. Share in value added. Full sample. Additional controls. Poisson estimator.

Dependent variable: Share VA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Start-up * ExtFin	-0.721			-0.867			-0.728			-0.794		
	[0.207]***			[0.207]***			[0.189]***			[0.189]***		
Start-up+Reg * ExtFin		-0.097			-0.100			-0.108			-0.114	
		[0.038]**			[0.035]***			[0.032]***			[0.034]***	
Start-up+Reg+Deal * ExtFin			-0.053			-0.058			-0.060			-0.063
			[0.018]***			[0.020]***			[0.019]***			[0.021]***
KL ratio * K-intensity	-0.017	-0.018	-0.017	-0.020	-0.021	-0.020	-0.037	-0.040	-0.039	-0.018	-0.019	-0.015
	[0.023]	[0.023]	[0.023]	[0.024]	[0.024]	[0.024]	[0.025]	[0.025]	[0.025]	[0.024]	[0.023]	[0.024]
Schooling * HK intensity	0.038	0.041	0.039	0.039	0.041	0.040	0.055	0.059	0.058	0.040	0.042	0.039
	[0.013]***	[0.014]***	[0.013]***	[0.014]***	[0.014]***	[0.014]***	[0.016]***	[0.017]***	[0.017]***	[0.015]***	[0.015]***	[0.015]**
Nat. Res. * NatRes-intens.	0.015	0.015	0.015	0.013	0.013	0.013	0.012	0.011	0.012	0.014	0.014	0.014
	[0.008]*	[0.009]*	[0.008]*	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]*	[0.009]	[0.009]*
Private Credit * ExtFin	1.133	1.391	1.267	0.969	1.204	1.060	1.245	1.513	1.372	1.094	1.369	1.214
	[0.390]***	[0.358]***	[0.376]***	[0.411]**	[0.394]***	[0.400]***	[0.410]***	[0.360]***	[0.384]***	[0.386]***	[0.342]***	[0.361]***
Private Credit * Tangibility	-1.850	-1.740	-1.760	-2.220	-2.121	-2.141	-1.797	-1.669	-1.694	-1.892	-1.773	-1.781
	[1.047]*	[1.045]*	[1.047]*	[1.014]**	[1.021]**	[1.024]**	[1.048]*	[1.046]	[1.048]	[1.046]*	[1.046]*	[1.047]*
Labour rigidity * ExtFin	-0.006	-0.007	-0.008									
	[0.005]	[0.005]	[0.005]									
Ln(Enforcement Days) * ExtFin				0.035	-0.110	-0.093						
				[0.162]	[0.160]	[0.159]						
Ln(GDPpc) * Elasticity							-0.132	-0.143	-0.145			
							[0.045]***	[0.048]***	[0.049]***			
Start-up * Growth opp.										0.675		
										[2.811]		
Start-up+Reg * Growth opp.											0.022	
											[0.566]	
Start-up+Reg+Deal * Growth opp.												-0.125
												[0.343]
Observations	1433	1433	1433	1406	1406	1406	1433	1433	1433	1433	1433	1433
Pseudo R-squared	0.34	0.34	0.34	0.35	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.10. Share in value added. Full sample. Additional controls. Tobit estimator.

Dependent variable: ln(ε+Share VA)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Start-up * ExtFin	-0.547			-0.658			-0.633			-0.636		
	[0.162]***			[0.164]***			[0.151]***			[0.155]***		
Start-up+Reg * ExtFin		-0.080			-0.098			-0.097			-0.101	
		[0.028]***			[0.028]***			[0.026]***			[0.027]***	
Start-up+Reg+Deal * ExtFin			-0.053			-0.058			-0.063			-0.062
			[0.016]***			[0.016]***			[0.015]***			[0.016]***
KL ratio * K-intensity	0.004	0.004	0.005	-0.001	-0.000	0.001	-0.009	-0.009	-0.008	0.002	0.001	0.003
	[0.019]	[0.019]	[0.019]	[0.018]	[0.018]	[0.018]	[0.019]	[0.019]	[0.019]	[0.019]	[0.019]	[0.019]
Schooling * HK intensity	0.040	0.042	0.039	0.041	0.043	0.040	0.054	0.057	0.054	0.042	0.046	0.042
	[0.011]***	[0.011]***	[0.011]***	[0.011]***	[0.011]***	[0.011]***	[0.012]***	[0.012]***	[0.012]***	[0.012]***	[0.012]***	[0.012]***
Nat. Res. * NatRes-intens.	0.017	0.017	0.017	0.015	0.015	0.015	0.015	0.015	0.015	0.017	0.016	0.017
	[0.006]***	[0.006]***	[0.006]***	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]***	[0.006]**	[0.006]***
Private Credit * ExtFin	1.026	1.266	1.069	0.934	1.145	0.945	1.083	1.349	1.127	0.961	1.223	1.008
	[0.306]***	[0.287]***	[0.303]***	[0.306]***	[0.297]***	[0.308]***	[0.305]***	[0.289]***	[0.303]***	[0.302]***	[0.285]***	[0.300]***
Private Credit * Tangibility	-1.333	-1.328	-1.333	-1.551	-1.538	-1.547	-1.300	-1.302	-1.307	-1.365	-1.366	-1.374
	[0.798]*	[0.802]*	[0.801]*	[0.788]**	[0.794]*	[0.793]*	[0.793]	[0.798]	[0.796]	[0.797]*	[0.800]*	[0.800]*
Labour rigidity * ExtFin	-0.005	-0.006	-0.007									
	[0.004]	[0.004]	[0.004]*									
Ln(Enforcement Days) * ExtFin				0.022	-0.081	-0.066						
				[0.092]	[0.086]	[0.087]						
Ln(GDPpc) * Elasticity							-0.096	-0.095	-0.100			
							[0.036]***	[0.036]***	[0.036]***			
Start-up * Growth opp.										0.924		
										[2.166]		
Start-up+Reg * Growth opp.											0.359	
											[0.401]	
Start-up+Reg+Deal * Growth opp.												0.087
												[0.227]
Observations	1433	1433	1433	1406	1406	1406	1433	1433	1433	1433	1433	1433
Pseudo R-squared	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.11. Share in value added. Full sample & No low income countries. Instrumental variables.
Dependent variable: ln(ε+Share VA)

	(1)	(2)	(3)	(4)	(5)	(6)
Start-up * ExtFin	-2.254			-1.738		
	[0.917]**			[0.536]***		
Start-up+Reg * ExtFin		-0.464			-0.315	
		[0.199]**			[0.126]**	
Start-up+Reg+Deal * ExtFin			-0.181			-0.154
			[0.069]***			[0.050]***
KL ratio * K-intensity	-0.026	-0.026	-0.023	-0.053	-0.053	-0.051
	[0.025]	[0.024]	[0.025]	[0.024]**	[0.024]**	[0.024]**
Schooling * HK intensity	0.001	0.010	0.004	0.010	0.017	0.009
	[0.022]	[0.022]	[0.023]	[0.023]	[0.024]	[0.023]
Nat. Res. * NatRes-intens.	0.015	0.015	0.015	0.012	0.011	0.012
	[0.008]**	[0.008]*	[0.008]**	[0.007]*	[0.008]	[0.007]*
Private Credit * ExtFin	-0.160	0.950	0.222	0.119	0.859	0.208
	[0.822]	[0.454]**	[0.532]	[0.644]	[0.391]**	[0.527]
Private Credit * Tangibility	-1.374	-1.449	-1.419	-1.603	-1.672	-1.640
	[0.780]*	[0.770]*	[0.781]*	[0.801]**	[0.798]**	[0.802]**
Observations	972	972	972	836	836	836
R-squared	0.62	0.59	0.62	0.67	0.65	0.66

Note: The share of value added is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. The Pseudo R-squared available for the Poisson estimator is calculated without clustering. ε=0.1.

Table 3.12. Number of establishments. Full sample

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Establishments)			Establishments			ln(ε+Establishments)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.863			-1.991			-0.833		
	[0.193]***			[0.434]***			[0.124]***		
Start-up+Reg * ExtFin		-0.134			-0.281			-0.134	
		[0.046]***			[0.062]***			[0.025]***	
Start-up+Reg+Deal * ExtFin			-0.085			-0.159			-0.080
			[0.021]***			[0.040]***			[0.012]***
Observations	1732	1705	1677	1732	1705	1677	1732	1705	1677
R-squared	0.86	0.85	0.85						
Pseudo R-squared				0.92	0.92	0.91	0.45	0.44	0.45

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.

Table 3.13. Number of establishments. No low income countries.

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Establishments)			Establishments			ln(ε+Establishments)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.943			-3.052			-0.897		
	[0.288]***			[0.806]***			[0.149]***		
Start-up+Reg * ExtFin		-0.123			-0.305			-0.124	
		[0.056]**			[0.094]***			[0.027]***	
Start-up+Reg+Deal * ExtFin			-0.084			-0.199			-0.079
			[0.022]***			[0.058]***			[0.012]***
Observations	1503	1476	1476	1503	1476	1476	1503	1476	1476
R-squared	0.85	0.84	0.84						
Pseudo R-squared				0.92	0.92	0.91	0.45	0.44	0.44

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.

Table 3.14. Number of establishments. Full sample

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Establishments)			Establishments			ln(ε+Establishments)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.663			-1.710			-0.646		
	[0.285]**			[0.451]***			[0.165]***		
Start-up+Reg * ExtFin		-0.112			-0.264			-0.113	
		[0.048]**			[0.055]***			[0.028]***	
Start-up+Reg+Deal * ExtFin			-0.091			-0.152			-0.080
			[0.020]***			[0.049]***			[0.016]***
KL ratio * K-intensity	0.013	0.013	0.015	-0.110	-0.097	-0.104	0.013	0.014	0.015
	[0.027]	[0.027]	[0.027]	[0.038]***	[0.038]**	[0.039]***	[0.020]	[0.020]	[0.020]
Schooling * HK intensity	0.017	0.020	0.014	0.025	0.034	0.028	0.020	0.023	0.018
	[0.022]	[0.021]	[0.022]	[0.024]	[0.024]	[0.023]	[0.012]	[0.012]*	[0.012]
Nat. Res. * NatRes-intens.	0.009	0.009	0.010	0.003	0.002	0.003	0.006	0.006	0.007
	[0.008]	[0.008]	[0.008]	[0.005]	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]
Private Credit * ExtFin	0.821	1.060	0.694	0.563	0.613	0.316	0.787	0.996	0.711
	[0.424]*	[0.354]***	[0.379]*	[0.277]**	[0.294]**	[0.316]	[0.320]**	[0.302]***	[0.315]**
Private Credit * Tangibility	-0.814	-0.840	-0.792	0.720	0.628	0.688	-0.796	-0.813	-0.777
	[1.205]	[1.195]	[1.202]	[0.615]	[0.627]	[0.620]	[0.815]	[0.814]	[0.812]
Observations	1238	1238	1238	1238	1238	1238	1238	1238	1238
R-squared	0.88	0.88	0.88						
Pseudo R-squared				0.92	0.92	0.91	0.47	0.46	0.48

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.

Table 3.15. Number of establishments. No low income countries.

Estimator:	OLS			Poisson			Tobit		
Dependent variable:	ln(ε+Establishments)			Establishments			ln(ε+Establishments)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Start-up * ExtFin	-0.740			-2.381			-0.724		
	[0.415]*			[0.845]***			[0.182]***		
Start-up+Reg * ExtFin		-0.116			-0.264			-0.116	
		[0.056]**			[0.072]***			[0.028]***	
Start-up+Reg+Deal * ExtFin			-0.098			-0.181			-0.086
			[0.020]***			[0.058]***			[0.016]***
KL ratio * K-intensity	0.004	0.004	0.007	0.005	0.020	0.016	0.004	0.004	0.007
	[0.032]	[0.032]	[0.031]	[0.078]	[0.081]	[0.083]	[0.020]	[0.020]	[0.020]
Schooling * HK intensity	0.028	0.031	0.022	0.065	0.077	0.070	0.030	0.033	0.026
	[0.027]	[0.026]	[0.027]	[0.022]***	[0.021]***	[0.021]***	[0.013]**	[0.013]***	[0.013]**
Nat. Res. * NatRes-intens.	0.006	0.006	0.008	0.006	0.006	0.006	0.004	0.004	0.005
	[0.008]	[0.008]	[0.008]	[0.005]	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]
Private Credit * ExtFin	0.529	0.749	0.287	0.684	0.550	0.298	0.489	0.701	0.325
	[0.413]	[0.338]**	[0.381]	[0.287]**	[0.309]*	[0.314]	[0.313]	[0.300]**	[0.314]
Private Credit * Tangibility	-1.118	-1.145	-1.079	0.933	0.986	0.949	-1.141	-1.172	-1.124
	[1.243]	[1.231]	[1.238]	[0.678]	[0.676]	[0.682]	[0.814]	[0.813]	[0.810]
Observations	1037	1037	1037	1037	1037	1037	1037	1037	1037
R-squared	0.87	0.87	0.88						
Pseudo R-squared				0.93	0.93	0.93	0.48	0.48	0.48

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.

Table 3.16. Number of establishments. Full sample. Additional controls. Ordinary Least Squares

Dependent variable: ln(ε+Establishments)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Start-up * ExtFin	-0.626			-0.461			-0.751			-0.893		
	[0.307]**			[0.309]			[0.295]**			[0.278]***		
Start-up+Reg * ExtFin		-0.103			-0.119			-0.121			-0.146	
		[0.050]**			[0.052]**			[0.048]**			[0.039]***	
Start-up+Reg+Deal * ExtFin			-0.087			-0.082			-0.097			-0.107
			[0.019]***			[0.019]***			[0.021]***			[0.018]***
KL ratio * K-intensity	0.013	0.013	0.015	0.011	0.012	0.013	-0.012	-0.011	-0.009	-0.004	0.003	0.005
	[0.027]	[0.027]	[0.027]	[0.028]	[0.027]	[0.027]	[0.030]	[0.030]	[0.029]	[0.029]	[0.029]	[0.028]
Schooling * HK intensity	0.017	0.019	0.013	0.012	0.009	0.006	0.047	0.049	0.044	0.035	0.031	0.024
	[0.022]	[0.021]	[0.022]	[0.023]	[0.022]	[0.022]	[0.023]**	[0.023]**	[0.023]*	[0.020]*	[0.018]	[0.021]
Nat. Res. * NatRes-intens.	0.009	0.009	0.010	0.009	0.009	0.010	0.007	0.007	0.008	0.009	0.009	0.010
	[0.008]	[0.008]	[0.008]	[0.008]	[0.007]	[0.007]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
Private Credit * ExtFin	0.853	1.088	0.724	0.691	0.728	0.463	1.063	1.336	0.960	0.811	1.044	0.678
	[0.422]**	[0.348]***	[0.371]*	[0.403]*	[0.367]*	[0.365]	[0.458]**	[0.400]***	[0.422]**	[0.409]*	[0.344]***	[0.369]*
Private Credit * Tangibility	-0.812	-0.831	-0.785	-0.895	-0.886	-0.855	-0.769	-0.803	-0.752	-0.839	-0.848	-0.808
	[1.206]	[1.198]	[1.203]	[1.221]	[1.218]	[1.220]	[1.189]	[1.178]	[1.184]	[1.197]	[1.182]	[1.193]
Labour rigidity * ExtFin	-0.002	-0.004	-0.004									
	[0.006]	[0.006]	[0.005]									
Ln(Enforcement Days) * ExtFin				-0.328	-0.396	-0.349						
				[0.138]**	[0.114]***	[0.115]***						
Ln(GDPpc) * Elasticity							-0.196	-0.188	-0.193			
							[0.068]***	[0.069]***	[0.068]***			
Start-up * Growth opp.										9.727		
										[3.918]**		
Start-up+Reg * Growth opp.											1.365	
											[0.828]	
Start-up+Reg+Deal * Growth opp.												0.633
												[0.408]
Observations	1238	1238	1238	1211	1211	1211	1238	1238	1238	1238	1238	1238
R-squared	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.

Table 3.17. Number of establishments. Full sample. Additional controls. Poisson estimator.
Dependent variable: Establishments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Start-up * ExtFin	-1.158			-1.532			-1.740			-1.766		
	[0.413]***			[0.468]***			[0.469]***			[0.376]***		
Start-up+Reg * ExtFin		-0.178			-0.267			-0.260			-0.285	
		[0.042]***			[0.053]***			[0.057]***			[0.056]***	
Start-up+Reg+Deal * ExtFin			-0.088			-0.138			-0.151			-0.168
			[0.051]*			[0.050]***			[0.051]***			[0.043]***
KL ratio * K-intensity	-0.103	-0.094	-0.099	-0.109	-0.098	-0.104	-0.108	-0.096	-0.102	-0.112	-0.098	-0.111
	[0.038]***	[0.037]**	[0.039]**	[0.038]***	[0.037]***	[0.038]***	[0.038]***	[0.038]**	[0.039]***	[0.039]***	[0.037]***	[0.040]***
Schooling * HK intensity	0.021	0.027	0.024	0.021	0.025	0.022	0.033	0.038	0.034	0.028	0.036	0.035
	[0.023]	[0.023]	[0.023]	[0.024]	[0.025]	[0.024]	[0.028]	[0.027]	[0.027]	[0.025]	[0.023]	[0.024]
Nat. Res. * NatRes-intens.	0.003	0.002	0.003	0.004	0.003	0.003	0.003	0.001	0.002	0.003	0.001	0.002
	[0.004]	[0.005]	[0.004]	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]
Private Credit * ExtFin	0.496	0.509	0.364	0.152	-0.068	-0.167	0.645	0.660	0.388	0.560	0.606	0.319
	[0.195]**	[0.222]**	[0.257]	[0.291]	[0.335]	[0.329]	[0.311]**	[0.324]**	[0.364]	[0.278]**	[0.292]**	[0.315]
Private Credit * Tangibility	0.710	0.653	0.686	0.731	0.661	0.710	0.796	0.671	0.754	0.719	0.613	0.659
	[0.615]	[0.624]	[0.619]	[0.613]	[0.618]	[0.616]	[0.568]	[0.568]	[0.567]	[0.622]	[0.637]	[0.643]
Labour rigidity * ExtFin	-0.014	-0.015	-0.014									
	[0.006]**	[0.006]**	[0.007]*									
Ln(Enforcement Days) * ExtFin				-0.288	-0.444	-0.346						
				[0.177]	[0.173]**	[0.189]*						
Ln(GDPpc) * Elasticity							-0.050	-0.024	-0.040			
							[0.066]	[0.074]	[0.069]			
Start-up * Growth opp.										3.106		
										[7.433]		
Start-up+Reg * Growth opp.											0.834	
											[0.986]	
Start-up+Reg+Deal * Growth opp.												0.808
												[0.558]
Observations	1238	1238	1238	1211	1211	1211	1238	1238	1238	1238	1238	1238
Pseudo R-squared	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.

Table 3.18. Number of establishments. Full sample. Additional controls. Tobit estimator.

Dependent variable: ln(ε+Establishments)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Start-up * ExtFin	-0.576			-0.439			-0.713			-0.812		
	[0.181]***			[0.179]**			[0.163]***			[0.171]***		
Start-up+Reg * ExtFin		-0.102			-0.124			-0.119			-0.139	
		[0.029]***			[0.029]***			[0.027]***			[0.029]***	
Start-up+Reg+Deal * ExtFin			-0.074			-0.071			-0.084			-0.096
			[0.017]***			[0.016]***			[0.016]***			[0.017]***
KL ratio * K-intensity	0.014	0.014	0.015	0.012	0.014	0.015	-0.010	-0.009	-0.008	-0.000	0.006	0.006
	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]
Schooling * HK intensity	0.019	0.021	0.017	0.015	0.012	0.010	0.049	0.052	0.047	0.034	0.031	0.028
	[0.012]	[0.012]*	[0.012]	[0.012]	[0.012]	[0.012]	[0.014]***	[0.013]***	[0.013]***	[0.013]***	[0.013]**	[0.013]**
Nat. Res. * NatRes-intens.	0.006	0.007	0.007	0.006	0.007	0.007	0.004	0.004	0.004	0.006	0.006	0.006
	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Private Credit * ExtFin	0.846	1.033	0.757	0.648	0.660	0.466	1.031	1.268	0.971	0.790	0.986	0.699
	[0.326]***	[0.303]***	[0.317]**	[0.326]**	[0.313]**	[0.323]	[0.320]***	[0.304]***	[0.316]***	[0.319]**	[0.301]***	[0.314]**
Private Credit * Tangibility	-0.792	-0.801	-0.767	-0.823	-0.810	-0.789	-0.735	-0.758	-0.718	-0.819	-0.820	-0.793
	[0.814]	[0.813]	[0.811]	[0.814]	[0.810]	[0.810]	[0.806]	[0.806]	[0.804]	[0.811]	[0.812]	[0.810]
Labour rigidity * ExtFin	-0.004	-0.006	-0.005									
	[0.004]	[0.004]	[0.004]									
Ln(Enforcement Days) * ExtFin				-0.332	-0.398	-0.359						
				[0.109]***	[0.101]***	[0.102]***						
Ln(GDPpc) * Elasticity							-0.193	-0.187	-0.191			
							[0.038]***	[0.038]***	[0.038]***			
Start-up * Growth opp.										8.147		
										[2.437]***		
Start-up+Reg * Growth opp.											1.125	
											[0.434]***	
Start-up+Reg+Deal * Growth opp.												0.658
												[0.245]***
Observations	1238	1238	1238	1211	1211	1211	1238	1238	1238	1238	1238	1238
Pseudo R-squared	0.47	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.

Table 3.19. Number of establishments. Full sample & No low income countries. Instrumental variables.
Dependent variable: ln(ε+Establishments)

	(1)	(2)	(3)	(4)	(5)	(6)
Start-up * ExtFin	-2.902			-2.318		
	[1.031]***			[0.826]***		
Start-up+Reg * ExtFin		-0.802			-0.608	
		[0.377]**			[0.278]**	
Start-up+Reg+Deal * ExtFin			-0.306			-0.227
			[0.119]**			[0.093]**
KL ratio * K-intensity	0.024	0.025	0.026	0.010	0.013	0.013
	[0.030]	[0.030]	[0.031]	[0.034]	[0.034]	[0.035]
Schooling * HK intensity	0.001	0.004	0.003	0.039	0.042	0.036
	[0.032]	[0.034]	[0.032]	[0.030]	[0.032]	[0.031]
Nat. Res. * NatRes-intens.	0.011	0.014	0.012	0.008	0.010	0.009
	[0.008]	[0.010]	[0.008]	[0.008]	[0.009]	[0.008]
Private Credit * ExtFin	-0.723	0.201	-0.581	-0.295	0.458	-0.412
	[0.790]	[0.746]	[0.610]	[0.537]	[0.440]	[0.495]
Private Credit * Tangibility	-0.980	-0.937	-0.937	-0.622	-0.570	-0.598
	[1.164]	[1.163]	[1.153]	[1.125]	[1.128]	[1.128]
Observations	897	897	897	790	790	790
R-squared	0.85	0.82	0.85	0.87	0.85	0.87

Note: The number of establishments is measured for 1990 and data comes from the UNIDO dataset. Regressions include country and sector fixed effects. Standard errors are clustered by country and reported in parenthesis. *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively. ε=0.1.