Foreign Direct Investment: Analysis of Aggregate Flows

Razin, Assaf and Sadka, Efraim

Princeton University Press

2006

Foreign Direct Investment: Analysis of Aggregate Flows

Assaf Razin[1] Efraim Sadka[2]

August 2006

[1]The Bernard Schwartz Chair in Global Markets, Tel-Aviv University; The Friedman Chair in International Economics, Cornell University; CESifo; CEPR; and NBER.

[2]The Heury Kaufman Chair in International Capital Markets, Tel-Aviv University; CESifo; and IZA.

Contents

List of Illustrations

List of Tables

12. Table 8.4: Bilateral FDI Flows and Selection Equations (Observations on Non-OECD to Non-OECD FDI are included)

13. Table 8.5: The Instrumented Productivity Equation

14. Table 9.1: The Effects of Host and Source Corporate-Tax Rates on FDI

15. Table 10.1: Statutory Corporate Tax Rates in the Enlarged EU, 2003

Preface

Economists tend to favor the free flow of capital across national borders, because it allows capital to seek out the highest rate of return. They also offer several other advantages . First, they reduce the risk faced by owners of capital by allowing them to diversify their lending and investment. Second, the global integration of capital markets can contribute to the spread of best practices in corporate governance, accounting standards, and legal traditions. Third, the global mobility of capital limits the ability of governments to pursue bad policies.

Capital can flow across countries in a variety of ways. One can distinguish among three major ones: foreign direct investment (FDI), foreign portfolio investment and loans. Among all these types, FDI, which involves a lasting interest and control, stands out. The world flows of FDI rose about sevenfold (in current U.S. dollars) over the 1990's; the vast majority is flowed between developed countries, but there are recently increased flows into emerging markets.

This book provides a treatise of the unique features of FDI flows, covering both theory and data. It focuses on the determinants of the aggregate flows of FDI at the source-host country level.

The book is likely to find its main readership among academics, graduate students, and trained policy professionals. The level of analysis is appropriate for an advanced graduate course, and could be accessible to anyone with some graduate training in economics. The book is also relatively self contained, including a special chapter reviewing the econometric techniques used, which means that reader do not necessarily have to consult other reference books.

The scope is particular to the topic studied. As a result, it could find some use as a textbook in a course specially designed to study foreign direct investment. Also, chapters of the book can be assigned as readings in a broader based international finance course.

To the best of our knowledge, there are no other books covering the same subject matter. There has been a great deal of work studying FDI from a micro or trade based perspective, but little has focused on the macroeconomics of FDI. The existing macroeconomic literature, available mostly in research papers (other than a book form), tends to focus on FDI to developing countries. As a result, this book can be expected to fill a niche in the literature on FDI.

In writing this book, we greatly benefitted from previous collaborations. Specifically, chapter two is based on Goldstein and Razin (2006). We thank Itay Goldstein for allowing us to use this work in the book. Part two and Chapter Nine are based, respectively, on Razin, Rubinstein and Sadka (2004

and 2005). We thank Yona Rubinstein for allowing us to use these works in the book. Chapter Eight is based on the unpublished paper of Razin, Sadka and Tong (2005). We thank Hui Tong for this collaboration. Chapter Three is based on our previous research, Razin and Sadka (forthcoming).

Financial support from the Bernard A. Schwartz Program in the Political Economy of Free Markets at Tel-Aviv University is gratefully acknowledged. Part of this book was written while Assaf Razin and Efraim Sadka were visiting the CEFSifo Institute, Munich, which provided us with an excellent research environment and warm hospitality.

We are indebted to two anonymous reviewers of Princeton University Press for many insightful comments and suggestions that improved the quality of the "final product". Special thanks are due to our Ph. D. student, Alon Cohen, who provided an excellent research assistance in the estimation results of Chapter Nine and the simulation results of Chapters Five and Ten.

Chapter 1

Overview

Economists tend to favor the free flow of capital across national borders, because it allows capital to seek out the highest rate of return. Unrestricted capital flows may also offer several advantages, as noted by Feldstein (2000). First, international flows reduce the risk faced by owners of capital by allowing them to diversify their lending and investment. Second, the global integration of capital markets can contribute to the spread of best practices in corporate governance, accounting standards, and legal traditions. Third, the global mobility of capital limits the ability of governments to pursue bad policies.

1.1 Channels of International Capital Flows

Capital can flow across countries in a variety of ways. One can distinguish among three major ones: foreign direct investment (FDI), foreign portfolio investment (FPI) and loans. FDI is defined as an investment involving a long term relationship and reflecting a lasting interest and control of a resident entity in the source country (foreign direct investor or parent firm) in the host country.

In national and international accounting standards, FDI is defined as involving an equity stake of 10% or more. In general, FDI itself has three components: equity capital, intra-firm loans and reinvestment of retained earnings. Because different countries have different recording practices relating to these three components, there arise some measurement problems[1]. Not all countries follow the 10% mark for the definition of FDI. Most countries do indeed report long-term intra-firm loans, but not all countries report short-term loans. Most countries report reinvestment of retained earning only with a considerable lag. One implication of these measurement problems is that FDI inflows do not contemporaneously match FDI outflows.

Foreign portfolio investment is different from FDI in that it lacks the element of lasting interest and control. Foreign portfolio investment includes also lending in the form of tradable bonds. The third type of foreign investment is loan, primarily bank loans.

Among these types of foreign investment flows, FDI stands out. The world flows of FDI rose about sevenfold in current U.S. dollars over the 1990's (see Figure 1.1, A and B)[2]. Furthermore, the vast majority of these flows are among OECD countries. FDI flows from OECD to non-OECD countries are also significant (see Table 1.1).[3] Maurice Obtfeld and Alan M. Taylor (2002) make a succinct observation: "A century ago, world income and productivity levels were far less divergent than they are today, so it is all the more remarkable that so much capital was directed to countries at or below the 20 percent and 40 percent income levels (relative to the United States). Today, a much larger fraction of the world's output and population is located in such low-productivity regions, but a smaller share of global foreign investment reaches them."

(Figure 1.1 A&B about here)

(Table 1.1 about here)

The U.N. (2005) annual report on world investment documents how countries are becoming more receptive to FDI. Table 1.2, which refers to the years 1991-2004, shows that the vast majority of changes in laws and regulations pertaining to investment were more favorable to FDI. An exception is developing countries which introduced some laws and regulations intended to protect some natural resources (especially in the energy field) against "foreign intruders"[4]. The report also indicates that countries are cooperating with each other in designing pro-FDI bilateral policies: "The number of bilateral investment treaties (BITs) and double taxation treaties (DTTs) reached 2,392 and 2,559 respectively, in 2004, with developing countries concluding more such treaties with other developing countries."

(Table 1.2 about here)

This book focuses on the unique features of FDI, vis-a-vis other types of capital flows.

1.2 Micro Level Studies

Studies of FDI can essentially be divided into two main categories: micro level (industrial organization and international trade) studies and macro-finance studies. Initially, the literature that explained FDI in microeconomics terms focused on market imperfections, and on the desire of multinational enterprises to expand their market power; see, for instance, Caves (1971). Subsequent literature centered more on firm-specific advantages, owing to product superiority or cost advantages, stemming from economies of scale, multi-plants economies and advanced technologies, or superior marketing and distribution; see, for instance, Helpman (1984).

A multinational may find it cheaper to expand directly in a foreign country, rather than through trade, in cases where its advantages stem from internal, indivisible assets associated with knowledge and technology.[5] The latter form of FDI is referred to as horizontal FDI. Note therefore that horizontal FDI is a substitute for exports. Brainard (1997) employs a differentiated product framework to provide an empirical support for this hypothesis. Helpman, Melitz and Yeaple (2004) incorporate intraindustry heterogeneity to conclude, among other things, that FDI plays a lesser role in substituting for exports in industries with large productivity dispersion.

However, horizontal FDI is not the only form of FDI. Multinational corporations account for a very significant fraction of world trade flows, with trade in intermediate inputs between divisions of the same firm constituting an important portion of these flows; see, for instance, Hanson, Mataloni and Slaughter (2001). This is referred to as vertical FDI.[6] One of the key determinants of vertical FDI is the abundance of human capital; see Antras (2004) for a comprehensive theoretical and empirical treatise of the various forms of FDI.

In a recent survey, Helpman (2006) observes that between 1990 and 2001 sales by foreign affiliates of multinational corporations expanded much faster than exports of goods and nonfactor services. He also points out that the fast expansion of trade in services has been accompanied by fast-growing trade in inputs. Furthermore: "...the growth of input trade has taken place both within and across the boundaries of the firm, i.e., as intra-firm and arm's-length trade." In light of these developments, Helpman argues that "the traditional classification of FDI into vertical and horizontal forms has become less meaningful in practice." Indeed, his survey includes some new applications of the theory of the organization of the firm to analyze the patterns of exports, FDI, outsourcing, etc.

1.3 Macro-Finance Studies

FDI combines not only aspects of international trade in goods and services but also aspects of international financial flows. The macro-finance literature attempts to analyze the composition of aggregate international flows into FDI, FPI and bank loans, as well as the breakdown of the aggregate flow of FDI according to either modes of entry or modes of finance. As with respect to the modes of entry, FDI can be made either at the greenfield stage or in the form of purchasing ongoing firms (Mergers and Acquisitions - M&A). U.N. (2005) observes that "the choice of mode is influenced by industry - specific factors. For example, greenfield investment is more likely to be used as a mode of entry in industries in which technological skills and production technology are key. The choice may also be influenced by institutional, cultural and transaction cost factors, in particular, the attitude towards takeovers, conditions in capital markets, liberalization policies, privatization, regional integration, currency risks and the role played by intermediaries (e.g. investment bankers) actively seeking acquisition opportunities and taking initiatives in making deals."

As for the modes of finance, there is a distinction between equity capital, intra-company loans and reinvestment of retained earnings. Figure 1.2 [which reproduces Figure 1.4 of U.N. (2005)] describes the relative share of these three modes of finance over the last decade. The lion's share of FDI is financed through equity capital, 60%-70%. The share of intra-firm loans has risen in the 1990s but has declined sharply in the 2000s. This decline is due mainly to repatriation of such loans by multinationals in developed economies. The third mode of finance, reinvestment of retained earnings, seems to exhibit a mirror image pattern to the intra-firms loans.

(Figure 1.2 about here)

The macro-finance literature on FDI started with studies examining the effects of exchange rates on FDI. These studies focused on the positive effects of an exchange rate depreciation in the host country on FDI inflows. A real exchange rate depreciation lowers the cost of production and investment in the host country, thereby raising the profitability of foreign direct investment[7]. The wealth effect is another channel through which a depreciation of the real exchange rate could raise FDI. By raising the relative wealth of foreign firms, a depreciation of the real exchange rate could make it easier for these firms to use the retained earnings to finance investment abroad, or to post a collateral in borrowing from domestic lenders in the host country capital market; see, for instance, Froot and Stein (1991).

Later macroeconomic studies emphasize the effect of FDI on long-run economic growth and cyclical fluctuations. A comprehensive study by Bosworth and Collins (1999) provides evidence on the effect of capital inflows on domestic investment for 58 developing countries during 1978-95.[8] The sample covers nearly all of Latin America and Asia, as well as many countries in Africa. They find that an increase of a dollar in the volume of capital inflows is associated with an increase in domestic investment of about 50 cents. (In the regression, both capital inflows and domestic investment are expressed as percentages of GDP). This result, however, masks significant differences among different types of inflows. FDI appears to bring about a one-for-one increase in domestic investment; there is virtually no discernible relationship between portfolio inflows and investment (little or no impact); and the impact of loans falls between those of the other two. These results hold both for the 58-country sample and for a subset of 18 emerging markets. Boresztein, De Gregorio, and Lee (1998) find that FDI increases economic growth when the level of education in the host country - a measure of its absorptive capacity - is high.[9] Similarly, Razin (2004) finds strong evidence for the dominant positive effect of FDI (relative to other forms of foreign investments) on domestic investment and growth.

The macroeconomic-finance literature also notes that foreign direct investment (FDI) has proved to be resilient during financial crises. For instance, in East Asian countries, such investment was remarkably stable during the global financial crises of 1997-98. In sharp contrast, other forms of private capital flows - portfolio equity and debt flows, and particularly short-term flows - were subject to large reversals during the same period; see

Dadush, Dasgupta, and Ratha (2000), Lipsey (2001), Loungani and Razin (2001), and Razin and Sadka (2003). The resilience of FDI during financial crisis was also evident during the Mexican crisis of 1994-95 and the Latin American debt crisis of the 1980s.[10]

1.4 Scope and Purpose

Foreign direct investment is a form of international capital flows. It may play an important role in the general allocation of world capital across countries. It is often pictured, together with other forms of capital flows, as shifting capital from rich, capital-abundant economies to poor, capital-scarce economies, so as to close the gap between the rates of return to capital, and enhance the efficiency of the world-wide stock of capital. This is the neo-classical paradigm. This general portrayal of international capital flows may indeed pertain to FDI flows from developed countries to developing countries. The latter are almost all net recipients of FDI. Even in this case, multinational FDI investors bring not only scarce capital to the host developing countries but also superior technologies and new industries.

However, the neo-classical portrayal of international capital flow is hardly reminiscent of the FDI flows among developed countries, which are much larger that those from developed to developing countries. Although *net* aggregate FDI flows from, or to, a developed country is typically small, the gross flows are quite large (see Table 1.1). As Lipsey (2000) observes: "The flows among the developed countries mainly seem to reshuffle the ownership of productive assets, moving them to owners who want them more than their current owners and who are willing to pay the most for them. Presumably, capital flows move assets from less efficient to more efficient owners, or from owners who are technologically or commercially backward in their industries to firms that are technological leaders. In none of these cases do such flows necessarily change the location of the production, assets, or employment of these industries, though."

In view of this succinct account of FDI flows among developed countries, there arises a question whether FDI plays any useful economic role except the mere shift of asset ownership. Similarly, in many cases FDI to developing countries is also merely a roundtripping of capital. Savers in a developing country which does not have developed and well-functioning saving and financial intermediation institutions export their capital to a location which specializes in exporting back FDI to this country (China and Hong Kong are a notable example). In this case too there arises the same question of whether this roundtripping of capital, which created no *net* import of capital, serves any useful economic role.

The theme advanced in this book views things in a sharply different way. We develop an empirically oriented theory which attributes a meaningful economic consequences and implications to a *two-way* flows of FDI among developed countries. Also, our book assigns a clearly unique role to FDI, as distinct from FPI and other forms of international capital flows. A key hypothesis of this book is that FDI firms are more efficiently managed than other firms. Thus, for instance, Perez-Gonzalez (2005) shows that after a foreign investor establishes a position that is greater than 50% of the firm's shares, the firm's productivity improves significantly. Having an empirically oriented theory enables us to confront its implications with the data.

1.4.1 Bilateral FDI Flows

FDI flows between a pair of countries. Therefore, there may be important country-pair characteristics that drives the flows of FDI between these two countries. For instance, a common language, the geographical distance, the similarity or difference in the legal systems (especially, corporate governance and accounting standards), bilateral trade or monetary agreements, common security arrangements, etc. are all factors that can facilitate or undermine the bilateral flows of FDI. This book studies the determinants of the aggregate flows of FDI between pairs of countries rather than the aggregate flows into a specific country from the rest of the world. Indeed, there are recently rich dataset on bilateral FDI flows, especially on flows that originate from OECD source counties. Needless to say, studies of bilateral FDI flows help us to better understand the aggregate flows in and out of a country.

1.4.2 Roadmap

We start by studying the features that divide foreign investment between FDI and portfolio flows. FDI stands out, relative to other flows, in that FDI investors assume control and management. Therefore, FDI firms are more efficiently managed. This is a *key* hypothesis in the analysis in this book. There are, however, also costs to direct investments. We specify two types of costs. The first type reflects the initial fixed cost that an FDI investor has to incur in order to manage the firm. The second type, endogenously determined, reflects the cost that may be inflicted on a direct investor when she must sell the firm because of some liquidity shock. Because this idiosyncratic shock is unobserved, the market may not be able to distinguish whether the sale is caused by this shock or rather by some negative signal, private to the FDI investor, about the firm's profits; and therefore the sale price is decreased. Thus, foreign investors with a low probability of liquidity shocks (for instance, high-pocket multinationals) select to be foreign direct investors, whereas the other choose portfolio investments.

Having analyzed the formation of foreign direct investors, relative to portfolio investments, we turn to analyze aspects of foreign direct investors in relation to domestic investors. We study a screening mechanism through which foreign direct investors manifest their comparative advantage over domestics investors in eliciting high-productivity firms. We show that this advantage diminishes as corporate transparency is improved; and the flows of FDI fall accordingly.

The existence of fixed setup cost of new investments introduces two margins of FDI decisions. There is an intensive margin of determining the magnitude of the flows of FDI, according to standard marginal productivity conditions, and also an extensive margin of determining whether at all to make a new investment. Country-pair specific shocks may affect these two margins in different ways. Maintaining wages fixed in the host country, a positive productivity shock in this country increases the marginal products of the factors of production (including capital), and has therefore a positive effects on the flows of FDI that are governed by the intensive margin. However, when wages are allowed to adjust, the productivity shock generates an upward pressure on wages which raises the fixed setup costs and discourage FDI through the extensive margin. We formulate these conflicting effects of productivity shock in the host country in a way that allows an econometric application. We also analyze productivity shocks in the source country which may have different effects on mergers and acquisitions (M&A) FDI, and on greenfield FDI.

Datasets on bilateral FDI flows typically include many source-host country observations with zero flows. This, by itself, is somewhat indicative of the existence of an extensive margin with the country-pair heterogeneity of fixed setup costs. In Part Two we explain and illustrate the advantage of employing the Heckman selection bias method (over Tobit and other methods) in empirically studying the determinants of bilateral FDI flows. This is done in a sample of panel data on 24 OECD countries over the period from 1981 to 1998. The data are drawn from the source OECD dataset which reports FDI flows from OECD countries to both OECD and non-OECD countries, as well as FDI flows from non-OECD countries to OECD countries But it does not report FDI flows from non-OECD countries to non-OECD countries. We therefore chose to employ for much of the analysis a panel data on 24 OECD countries over the period 1981 to 1998, for which data on flows in all directions are available.

Part Three analyzes the main empirical studies of the country-pair determinants of FDI. The effects of productivity shocks are investigated in a sample of panel data on 62 countries (29 OECD countries and 33 non-OECD countries) over the period from 1987 to 2000. As there is a large heterogeneity in the productivity shocks between OECD and non-OECD countries, which is useful for analyzing the effects of productivity on FDI flows, we chose to study a larger sample of panel data in this case. We find some evidence in support of the conflicting effects of productivity shocks.

We also investigate the role played by the host and source corporate tax rates on the intensive and extensive margins. We find that the host country tax rate has a negative effect primarily on the intensive margin, whereas the source tax rate has a positive effect mostly on the extensive margin.

Finally, we discuss some policy implications. Specifically, we formulate an international tax competition model to explain the co-existence of a "rich" source country with high capital-income (business and individuals) taxes and public expenditures and a "poor" host country with low capital-income taxes and public expenditures. This phenomenon may be common in the enlarged EU with the new accession countries which, are predominant recipients of FDI from the old member countries. We also analyze the welfare gains from a tax coordination.

Notes

[1]See also Feenstra (1999).

[2]Other forms of foreign investment, such as debt, also increased dramatically, especially to non-OECD countries.

[3]Among non-OECD countries, China with $72 billion is by far the largest recipient of FDI in 2005, surpassed only by the U.K. and the U.S.

[4]For instance, BG and BP of the U.K., Total of France and Repsol of Spain were all expropriated in Bolivia in 2006. A similar fate hit Petrobras in Brazil and Occidental-Petroleum in Ecuador.

[5]For a comprehensive treatise of the role of multinational in International trade see Markusen (2002).

[6]See, for instance, Yeaple (2003) for an empirical investigation of the scope of vertical FDI.

[7]See, for instance, Blonigen (1997).

[8]Note that foreign investment per se is not related one-to-one to domestic investment. As noted by Froot (1991) for the case of FDI, it actually requires neither capital flows nor investment in capacity. Conceptually, FDI is an extension of corporate control over international boundaries: "When Japanese-owned Bridgestone takes control over the US firm Firestone, capital need not flow into the US. US domestic lenders can largely finance the equity purchase. Any borrowing by Bridgestone from foreign-based third parties also does not qualify as FDI (although it would count as an inflow

of portfolio capital into US). And, of course, in such acquisition there is no investment expenditure; merely an international transfer in the title of corporate assets."

[9] Alfaro et al (2004) find that education level, development of local financial markets, and other local conditions play an important role in allowing the positive effect of FDI to materialize.

[10] Furthermore, the flow of FDI may even intensify during financial crises. Krugman (2000) argues that in financial crises foreigners can take advantage of fire sales of assets by liquidity-constrained domestic investors.

Part I

A Theory of FDI with Threshold Barriers

Chapter 2

Foreign Direct Investment and Foreign Portfolio Investment

2.1 Introduction

International equity flows take two major forms: Foreign Direct Investments (FDI) and Foreign Portfolio Investments (FPI). Despite the empirical interest in foreign equity flows, very little work has been done on jointly explaining FDI and FPI in a rigorous analytical framework. In this chapter, we propose such a framework, and provide a model of a trade off between FDI and FPI, which is consistent with the empirical regularity that FDI flows are generally less volatile than FPI flows. For instance, Table 5 in Lipsey (2000) shows that the ratio of standard deviations to means are 1.008 for FDI flows to Europe and 2.102 for FPI flows to Europe in the years 1969-1993.

Our model highlights a key difference between the two types of investment: FDI investors, who exert a higher degree of control and supervision over the management in the domestic firms; and FPI investors, who exert much less control over domestic firms, and delegate decisions to managers, but limit their freedom to make decisions because the managers' agenda may not be always consistent, with that of the owners. Consequently, due to agency problem between managers and owners and "free rider" problem among the firm' owners, portfolio investment projects are managed less efficiently than direct investment projects. As noted succinctly by Oliver Hart (2001), "If the shareholder does something to improve the quality of management, then the benefits will be enjoyed by all shareholders. Unless the shareholder is altruistic, she will ignore this beneficial impact on other shareholders and so will under-invest in the activity of monitoring or improving management." To be more specific, direct investors, who act effectively as managers of their own projects, are more willing to acquire costly information regarding changes in the prospects of their projects than portfolio investors. FDI investors are also more willing to exert costly monitoring on their managers than portfolio investors. Thus, FDI investors manage their projects more efficiently. This effects generates an advantage, with an added value in the capital markets, to direct investment relative to portfolio investments.

As we pointed out above, there are, however, costs to direct investments. We specify two types of costs. The first type, exogenously given in the model, reflects the initial fixed cost that an FDI investor has to incur in order to build and train managerial teams, supervisory boards, etc. which are required to achieve efficient functioning of the domestic project.[1]

The second type of costs, an information-based cost, is derived endogenously in the model. It results from the possibility that investors need to sell their investments before maturity because they face liquidity shocks. In such circumstances, the price they can get will be lower if they have more information on the economic fundamentals of the investment project. This is because when potential buyers know that the seller has more information, they may suspect that the sale results from bad information on the prospects of the investment (rather than from a negative liquidity shocks), and will thus be willing to pay only a lower price. Hence, if they invest directly, the investors bear the cost of getting a lower price if and when they are forced to sell the project before maturity. As the Economist (May 1, 2003) succinctly puts it: "FDI is a lot more difficult to withdraw when times are hard. Investments may have to be sold at a loss, if they can be sold at all." This may create a bias of less illiquidity-prone investors, such as "deep-pocket" multinationals, in favor of FDI.[2] More illiquidity-prone investors, such as institutional investors who are subject to frequent withdrawals, are biased in favor of FPI.[3]

Our model, therefore, describes a key trade off between management efficiency and liquidity. Both sides of this trade off are driven by the effect of asymmetric information, which comes with control. When they invest directly, investors get more information about the fundamentals of the investment, and thereby can manage the project more efficiently, than their portfolio-investors counterparts. However, this also generates a "lemons" type problem when they try to sell the investment before maturity (Akerlof (1970)). Therefore, this superior information effect reduces the price they can get when they are forced to sell the project prematurely.

This trade off between efficiency and liquidity has strong roots in existing empirical evidence. The idea that control increases efficiency and the value of the firm, which constitutes one side of the trade off, is supported empirically by two recent papers in the international finance literature. The first paper - by Perez-Gonzalez (2005) - shows that after a foreign investor establishes a position that is greater than 50% of the firm's shares, the firm's productivity, computed using data on future earnings, improves. The second paper - by Chari, Ouimet, and Tesar (2005) - demonstrates the positive response in the stock market to the establishment of control (defined, again, as more than 50% ownership). Because having more than 50% ownership is the ultimate indication for control, these two papers provide clear evidence on the link between control and value, which is a basic premise of this book. It should be noted, however, that large shareholders can achieve affective control in many cases by holding a block that is much smaller than 50% of the firm. This has been noted in the finance literature by Sheifer and Vishny (1986), Bolton and von Thadden (1998), and others. Going back to our basic premise, this implies that the value of the firm may increase with ownership concentration even when the controlling shareholder has a block that is smaller than 50%.

Such evidence is provided by Wruck (1989) and by Hertzel and Smith (1993). This is much in line with our focus on the trade off between FDI and FPI, as many FDI investments exhibit blocks that are much smaller than 50%.

The other side of the trade off - the idea that the sale of shares by control holders generates a larger price impact than a sale by other investors - can be supported by two strands in the financial literature. First, it has been shown that the sale of stocks by large block holders has a bigger downward effect on the price than sales of stocks by other investors; for example, see Mikkelson and Partch (1985), Holthausen, Leftwich, and Mayers (1990), and Chan and Lakonishok (1995). Following the logic above, this result may well apply to the basic premise in this chapter, as large block holders probably have more control over the firms' management. Second, perhaps the best evidence on the price impact of sale in the presence of control can be obtained by looking at what happens when the firm sells its own shares. After all, the firm has ultimate control over its operations, and thus this type of transaction is expected to suffer most from asymmetric information between the seller (firm) and potential buyers. Indeed, the finance literature has documented the large decrease in price following an announcement by the firm that it is going to sell mew equity (a seasoned equity offering, SEO); for example, see Masulis and Korwar (1986) and Korajczyk, Lucas, and MacDonald (1991).

A main implication of the trade off between efficiency and liquidity described in this chapter is that investors with high (low) expected liquidity needs are more likely to choose less (more) control. This is because investors with high expected liquidity needs are affected more by the low sale price associated with control, whereas those with low expected liquidity needs are affected more by the efficiency in management. As a result, in equilibrium, assets under control are less likely to be liquidated prematurely. This is consistent with evidence provided, for example, by Hennart, Kim, and Zeng (1998) in the management literature. This show that international investors are much more likely to exit from joint ventures than from fully owned investments, which clearly exhibit more control. In the context of our chapter, since FDI exhibit more control than FPI, the former is expected to be liquidated less often.[4]

This chapter has some roots in the existing literature. Albuquerque (2003) develops a model aimed at explaining the differences between the volatility of direct investments and the volatility of portfolio investments. His work relies on expropriation risks and the inalienability of direct investments, and thus is different from the information-based mechanism developed here.

Other works in the literature use the asymmetric information hypothesis to address different issues related to FDI. In Froot and Stein (1991), Klein and Rosengren (1994), and Klein, Peek and Rosengren (2002), the hypothesis is that FDI is information intensive, and thus FDI investors, who know more about their investments that outsiders, face a problem in raising resources for their investments. Gordon and Bovenbrg (1996) assume asymmetric information between domestic investors and foreign investors to explain the home bias phenomenon. Razin, Sadka and Yuen (1998) explain the pecking order of international capital flows with a model of asymmetric information. Finally, Razin and Sadka (2003) analyze the gains from FDI when foreign direct investors have superior information on the fundamentals of their investment, relative to foreign direct portfolio investors. Importantly, none of these papers analyzes the effects of asymmetric information on the liquidity of FDI and FPI, which is a major factor in the trade off developed in this chapter.[5]

2.2 The Model

A small economy is faced by a continuum [0,1] of foreign investors. Each investor has an opportunity to invest in one investment project. Investment can occur in two forms. The first form is a direct investment. The second form is a portfolio investment. The difference between the two forms of investment, in our model, is that a direct investor will effectively act like a manager, whereas in case of a portfolio investment, the investor will not be the manager, and the project will be managed by an "outsider". We assume that investors are risk neutral, and thus each investor chooses the form of investment that maximizes her ex-ante expected payoff.

There are three periods of time: 0, 1, and 2. In period 0, each investor decides whether to make a direct investment or a portfolio investment. In period 2, the project matures. The net cash flow from the project is denoted by $R(K, \varepsilon)$, where ε is a random productivity factor that is independently realized for each project in period 1, and K is the level of capital input invested in the project in period 1, after the realization of ε. For tractability we assume that $R(K, \varepsilon)$ takes the special form:

$$R(K, \varepsilon) = (1 + \varepsilon)K - \frac{1}{2}BK^2. \tag{2.1}$$

We assume that ε is distributed between -1 and 1, according to a cumulative distribution function $G(\cdot)$, and a density function $g(\cdot) = G'(\cdot)$. we also assume that $E(\varepsilon) = 0$. The parameter B, that affects negatively the net cash flow from a project, may reflect higher production costs, and/or lower productivity. For brevity we simply refer to it as a production cost parameter.

2.2.1 Management and Efficiency

In period 1, after the realization of the productivity shock, the manager of the project observes ε. Thus, if the investor owns the project as a direct investment, she observes ε, and chooses K, so as to maximize the net cash flow. The chosen level of K is denoted by $K^*(\varepsilon)$, and is given by:

$$K^*(\varepsilon) = \frac{1 + \varepsilon}{B}. \tag{2.2}$$

Thus, the ex-ante expected net cash flow from a direct investment, if it is held until maturity, is given by:

$$E\left(\frac{(1+\varepsilon)\cdot(1+\varepsilon)}{B}-\frac{1}{2}B\left(\frac{1+\varepsilon}{B}\right)^2\right)=\frac{E\left((1+\varepsilon)^2\right)}{2B}. \tag{2.3}$$

In case of a portfolio investment, the owner is not the manager, and thus she does not observe ε. In this case, the manager follows earlier instructions as for the level of K. A possible rationale behind this sequence of firm decisions, whereby the level of capital input K is determined *ex ante*, has to do with a potential agency problem between the owner and the manager (who is responsible for making these decisions). Loosely speaking, the latter is not exclusively interested in the net worth of the firm as in the former. For example, with no explicit instructions at hand, the manager may wish to set K at the highest possible level in order to gain power and financial rewards. As a result, when the owner does not have information about the firm's productivity, she will have to set investment guidelines for the manager (who knows more about ε than she does) so as to protect her own interests.[6]

The ex-ante instruction is chosen by the owner so as to maximize the expected return absent any information on the realization of ε, and is based on the *ex ante* zero mean. Thus, the manager will be instructed to choose $\bar{K}=K^*(0)=\frac{1}{B}$. Then, the ex-ante expected payoff from a portfolio investment, if it is held until maturity, is:

$$E\left(\frac{(1+\varepsilon)}{B}-\frac{1}{2B}\right)=\frac{E\left(1+2\varepsilon\right)}{2B}=\frac{1}{2B}. \tag{2.4}$$

It follows from Jensen's inequity[7] that $E\left[(1+\varepsilon)^2\right]>\left[E\left(1+\varepsilon\right)\right]^2=1.$

Therefore, comparing equation (2.3) with equation (2.4), we see that if the project is held until maturity, it yields a higher payoff as a direct investment than as a portfolio investment. This result reflects *efficiency* that results from a hands-on management style in the case of a direct investment.

There are, however, costs to direct investments. We specify two types of costs. The first type, reflects the fixed initial cost that an FDI investor has to incur in order to acquire the expertise to manage the project directly. We denote this cost, which is exogenously given in the model, by C. We simply assume that an investor who chooses FDI over FPI has to pay the fixed cost at time 0. We refer to this cost as an FDI cost.

The second type, an information-based cost, is derived endogenously in the model. It results from the possibility of liquidity shocks occurring in period 1.

2.2.2 Liquidity Shocks and Resale Prices

In period 1, before the value of ε is observed, the owner of the project might get a liquidity shock. With the realization of a liquidity shock, the investor is forced to sell the project immediately, that is, in period 1.[8]

We denote by λ the probability of liquidity shocks. We assume that there are two types of foreign investors. A proportion of one-half of the investors has high expected liquidity needs, and the remaining proportion has low expected liquidity needs. Formally, we assume that the first type of investors ("investors of type "H") face a liquidity need with probability λ_H, whereas the second type (type "L") face a liquidity need with probability λ_L. For simplicity, we assume that $1 > \lambda_H > \frac{1}{2} > \lambda_L > 0$, and that $\lambda_H + \lambda_L = 1$.[9] Investors know their type ex ante, but this is their own private information.

There is, however, also a possibility that an investor will liquidate a project in period 1 even if there is no liquidity shock. This can happen if and only if the initial investor observes a relatively low realization of ε. In such a case she does have superior information over the potential buyer, and can exploit it. Because portfolio investors do not observe ε in period 1, only direct investors sell their investment project at that time when a liquidity shock is absent. Because all kinds of sales occur simultaneously in period 1, buyers do not know the reason for a sale of any individual project. They know, however, whether the investment project is sold by a direct investor or by a portfolio investor. Because only direct investment projects are sold due to low productivity shocks, the price that direct investors can get when they try to sell the project in period 1 will be lower than the price obtained by portfolio investors. This generates a cost of the second type to FDI.

To evaluate this cost, we now derive the price that a direct investor gets if she sells the project in period 1. The pice is equal to the expected value of the project from the point of view of the potential buyer, given that the buyer knows that the owner is trying to sell, and given that she does not know the reason for the sale. We denote the maximum level of ε, under which the direct investor is selling the project in absence of a liquidity shock, by $\underline{\varepsilon}_D$. Also, we denote by λ_D the probability, as perceived by the market, that an

FDI investor gets a liquidity shock. Both $\underline{\varepsilon}_D$ and λ_D will be endogenously determined in equilibrium. Given that the FDI owner sells her projects, the buyer thinks that with probability $(1-\lambda_D)G(\underline{\varepsilon}_D)$ the owner is selling the project due to a low realization of ε, and with probability λ_D that she sells the projects because of a liquidity shock.

If the project is sold due to a liquidity shock, that is, before the initial owner observes ε (recall that liquidity shocks are realized before productivity shocks), the value of ε is not recorded in the firms before the sale. Therefore, the buyer does not know the value of ε. However, if the project is sold for low-profitability reasons, the owner will know the value of ε after the sale.[10]

Using Bayes' rule, the period 1 price that the direct investor gets for the project is given by:

$$P_{1,D} = \frac{(1-\lambda_D)\int_{-1}^{\underline{\varepsilon}_D} \frac{(1+\varepsilon)^2}{2A} g(\varepsilon)d\varepsilon + \lambda_D \int_{-1}^{1} \frac{1+2\varepsilon}{2A} g(\varepsilon)d\varepsilon}{(1-\lambda_D)\, G(\underline{\varepsilon}_D) + \lambda_D}. \tag{2.5}$$

The initial owner, in turn, sets the threshold level $\underline{\varepsilon}_D$, such that given $P_{1,D}$, when observing $\underline{\varepsilon}_D$, she is indifferent between selling, or not selling, the project. This yields the following equation:

$$P_{1,P} = \frac{(1+\underline{\varepsilon}_D)^2}{2B}. \tag{2.6}$$

Thus, equations (2.5) and (2.6) simultaneously determine $P_{1,D}$ and $\underline{\varepsilon}_D$ as functions of the market-perceived probability λ_D, denoted by $\underline{\varepsilon}_D(\lambda_D)$ and $P_{1,D}(\lambda_D)$, respectively.

Note that $\underline{\varepsilon}_D(\lambda_D)$ and $P_{1,D}(\lambda_D)$ are increasing in λ_D: when λ_D is high, the buyer thinks that the probability that an early sale results from a liquidity shock (and not from a bad realization of the productivity parameter) is also high. Consequently, the resale price of the project in period 1 is high as well. This means that FDI investors sell their projects more often (that is, under a higher threshold $\underline{\varepsilon}_D$). An implication is that investors have a greater incentive to choose FDI in period 0 when the market participants think that investors with high liquidity needs choose FDI. This externality plays an important role in the next section where we derive the equilibrium allocation and market prices.

Note also that $\underline{\varepsilon}_D$ is always below 0, and consequently $P_{1,D}$ is always below $\frac{1}{2B}$. This feature plays an important role in the comparison between the resale price of FDI and the resale price of portfolio investments. To conduct this comparison, let us characterize the resale price of a portfolio investment project. Essentially, when a portfolio investor sells the projects in period 1, everybody knows she does it because of a liquidity shock. Thus, the price she gets for the project is given by:

$$P_{1,P} = \int_{-1}^{1} \frac{1+2\varepsilon}{2B} g(\varepsilon) d\varepsilon = \frac{1}{2B}. \tag{2.7}$$

Now, we can see that the resale price of a direct investment in period 1 is always lower than the resale price of a portfolio investment in that period. The intuition is that if a direct investor prematurely sells the investment project, the market price must reflect the possibility that the sale originates from inside information on low prospects of this investment project.

We can now summarize the essential trade-off between FDI and FPI. A benefit of a direct investment is that it enables the investor to manage the project more efficiently. This increases the return that she gets in case she does not have to sell the project prematurely. However, if a foreign investor ex-ante chooses to hold the project as a direct investment, but sells the project prematurely, she gets a relatively low price. This is because potential buyers perceive that with some probability the project is sold due to negative inside information about the prospects of the investment. Thus, the additional information associated with a direct investment is not necessarily beneficial. In addition, investing directly entails a fixed cost C. With such trade off between FDI and FPI in mind, investors choose the type of investment that maximizes their ex ante expected net cash flow. We now turn to study this choice.

2.3 Ex-Ante Choice between FDI and FPI

2.3.1 Expected Value of FDI

With probability λ_i ($i = H, L$), an investor of type i gets a liquidity shock, and sells the project in period 1. (Note that this probability can be different from λ_D, the probability perceived by the market.) The market price is:

$$P_{1,D}(\lambda_D) = \frac{(1+\underline{\varepsilon}_D(\lambda_D))^2}{2B}.$$

With probability $1-\lambda_i$, the investor does not get a liquidity shock. She sells the project if the realization of ε is below $\underline{\varepsilon}_D(\lambda_D)$, but she does not sell it if the realization of ε is above $\underline{\varepsilon}_D(\lambda_D)$. Recall that $\underline{\varepsilon}_D(\lambda_D)$ is determined by equations (2.5) and (2.6). Therefore, the expected payoff, in the state of no liquidity shock is

$$\int_{-1}^{\underline{\varepsilon}_D(\lambda_D)} \frac{(1+\underline{\varepsilon}_D(\lambda_D))^2}{2B} g(\varepsilon)d\varepsilon + \int_{\underline{\varepsilon}_D(\lambda_D)}^{1} \frac{(1+\varepsilon)^2}{2B} g(\varepsilon)d\varepsilon.$$

In addition, a direct investor has to incur a fixed cost of C. Thus, the ex-ante expected net cash flow for a direct investor, as a function of λ_i, and A, is given by:

$$\begin{aligned} EV_{Direct}(\lambda_i,\lambda_D,B) = \quad & (1-\lambda_i)\left[\int_{-1}^{\underline{\varepsilon}_D(\lambda_D)} \tfrac{(1+\underline{\varepsilon}_D(\lambda_D))^2}{2B} g(\varepsilon)d\varepsilon + \int_{\underline{\varepsilon}_D(\lambda_D)}^{1} \tfrac{(1+\varepsilon)^2}{2B} g(\varepsilon)d\varepsilon\right] \\ & +\lambda_i \tfrac{(1+\underline{\varepsilon}_D(\lambda_D))^2}{2B} - C. \end{aligned} \tag{2.8}$$

2.3.2 Expected Value of FPI

When the investor holds the investment as a portfolio investment, with probability λ_i, she receives a liquidity shock, and sells the project in period 1. Then, the selling price is:

$$P_{1,P} = \frac{1}{2B}.$$

With probability $1 - \lambda_i$, the investor does nor receive a liquidity shock. Then, her expected net cash flow is:

$$\frac{E\left(1+2\varepsilon\right)}{2B} = \frac{1}{2B}.$$

Therefore, the ex ante expected net cash flow from a portfolio investment is given by:

$$EV_{Portfolio}(B) = \frac{1}{2B}. \tag{2.9}$$

2.3.3 FDI and FPI

We denote the difference between the expected value of FDI and the expected value of FPI by:

$$Diff\left(\lambda_i, \lambda_D, B\right) \equiv EV_{Direct}\left(\lambda_i, \lambda_D, B\right) - EV_{Portfolio}(B). \tag{2.10}$$

Then, investor i will choose FDI when $Diff\left(\lambda_i, \lambda_D, B\right) > 0$; will choose FPI when $Diff\left(\lambda_i, \lambda_D, B\right) < 0$; and will be indifferent between the two (that is, may choose either FDI or FPI) when $Diff\left(\lambda_i, \lambda_D, B\right) = 0$.

The choice between FDI and FPI is governed by the parameters B and

C. Investor i is more likely to choose FDI when:

(i) The FDI cost (C) is lower.

(ii) The productivity cost (B) is lower.

(iii) The probability of getting a liquidity shock (λ_i) is lower.

(iv) The market-perceived probability λ_D of a liquidity shock for FDI investors is higher.

The result in Part (i) is expected: investors are less likely to choose FDI when the fixed cost they have to incur in order to set the direct investment up is higher. Part (ii) says that when the production cost is higher, investors are less likely to choose FDI. The intuition behind this result is that when the production cost increases, the overall profitability of investment projects decreases, and this makes it less beneficial to incur the additional fixed cost associated with FDI. Part (iii) means that investors with lower ex ante liquidity needs are more likely to choose direct investments. This is because these investors expect to benefit more from the long-term efficiency associates with FDI, and to suffer less from the lower short-term price of this form of investment. Finally, Part (iv) states that when the probability λ_D that is assessed by the market to a liquidity shock of FDI investors increases, investors are more likely to choose FDI. The intuition is related to the fact that the resale price of FDI increases in λ_D. This makes direct investments more attractive relative to portfolio investments.

2.4 Market Equilibrium

2.4.1 The Allocation of Investors between FDI and FPI

So far, we analyzed the (partial) equilibrium choice of the two types of investors between the two types of investments, given the market-perceived operability λ_D.To complete the description of equilibrium, it remains to specify how λ_D is determined. Assuming that rational expectations hold in the market, λ_D has to be consistent with the equilibrium choice of investors between FDI and FPI. thus, it is given by the following equation:

$$\lambda_D = \frac{\lambda_H \lambda_{H,FDI} + \lambda_L \lambda_{L,FDI}}{\lambda_{H,FDI} + \lambda_{L,FDI}}, \tag{2.11}$$

where $\lambda_{H,FDI}$ is the proportion of λ_H investors who choose FDI in equilibrium and $\lambda_{L,FDI}$ is the proportion of λ_L investors who choose FDI in equilibrium.[11]

Note that there cannot be an equilibrium where some λ_H investors choose FDI, while some λ_L investors choose FPI. Thus, only five cases can potentially be observed in equilibrium. These are summarized as follows:

Case 1: All λ_H and λ_L investors choose FDI.

Case 2: All λ_L investors choose FDI; λ_H investors split between FDI and FPI.

Case 3: All λ_L investors choose FDI; λ_H investors choose FPI.

Case 4: All λ_L investors split between FDI and FPI; all λ_H investors choose FPI.

Case 5: All λ_H and λ_L investors choose FPI.

In describing the equilibrium outcomes below, we will often refer to these cases. It is worth noting that as we move from Case 1 to Case 5, the amount of FDI in the economy decreases, while the amount of FPI increases. Note also that only in cases 2, 3, and 4, FDI and FPI coexist in the economy. Also, among these, Case 3 exhibits the largest difference between expected liquidity needs for a representative FDI investor and those for a representative FPI investor.

Figure 2.1 provides a full characterization of the equilibrium allocation of investors between FDI and FPI as a function of two parameters, λ_H and B. The value of λ_H reflects the probability that investors with high expected liquidity needs will get a liquidity shock. Since we assumed that $\lambda_H + \lambda_L = 1$, we know that the value of λ_H also indirectly determines the value of λ_L (which reflects the probability that investors with low expected liquidity needs will get a liquidity shock). Thus, our interpretation is that an increase in λ_H reflects an increase in the heterogeneity across investors. The derivation of Figure 2.1 is relegated to the Appendix 2A.

(Figure 1.1 A&B about here)

Several features of Figure 2.1 are worth elaborating on. First, if FDI and FPI coexist in equilibrium, then the expected liquidity needs of FDI investors are lower, on average, than the expected liquidity needs of FPI investors. As noted above, the only possible cases in equilibrium, where FDI and FPI coexist, are cases 2, 3, and 4. In all these cases, liquidity shocks are more common among FPI investors than among FDI investors. Investors with high expected liquidity needs care less about the long-term efficiency of FDI, and care more about the short-term price. Thus, they have a higher tendency to invest in FPI. On the other hand, investors with low expected liquidity needs tend to prefer FDI. This result is consistent with the casual observation that FDI investors are often large and stable multinational corporations with low expected liquidity needs, whereas FPI investors (such as global mutual funds) are, on average, more vulnerable to liquidity shocks. This result contributes to the high withdrawal ratio of FPI relative to FDI, which can account for the empirically-observed higher volatility of net FPI inflows.

Second, as the production cost parameter (B) increases, there will be more FPI and less FDI in equilibrium. As the level of B, which represents the cost of production in the host country, increases, equilibrium outcomes change from Case 1, via Cases 2 and 3, to Case 5 - that is, they gradually exhibit more FPI and less FDI. Since B represents the cost of production, we expect developed countries to have higher levels of B. Thus, our model predicts that developed countries will attract more FPI, whereas developing countries will attract more FDI. This is indeed consistent with empirical evidence. Developed countries have higher costs of production, and thus lower profitability of investment projects. Thus, in these countries, it is less beneficial to pay the fixed costs associated with establishing an FDI investment. Furthermore, when a foreign direct investor from a developed country acquires a firm in a developing country, she may transfer her TFP in the source country to the new firm, thereby reducing the productivity cost B. This strengthens the relative attractivity of developing countries for FDI.

Third, as the liquidity need heterogeneity among investors increases, a separating equilibrium - with a large difference between the withdrawal rate of FPI and the withdrawal rate of FDI - becomes more likely. When $B < B^*$, an increase in λ_H shifts the equilibrium outcome from Case 1, which is a pooling equilibrium, to Case 3, which is a separating equilibrium with a large difference between the withdrawal rates of the two types of investment. When $B > B^*$, an increase in λ_H shifts the equilibrium outcome from Case 5, which is a pooling equilibrium, to Case 3. The implication is that a high level of liquidity need heterogeneity among investors causes them to be attracted to different types of investment, and leads to observed differences in withdrawal rates and volatility between FDI and FPI.

Fourth, there is a region of the fundamentals (B, λ_H, C) with multiple equilibria. Multiple equilibria exist when $B < B^*$ and $\lambda_H^*(B) < \lambda_H < \lambda_H^{**}(B)$. In this region, Case 1, Case 2 and Case 3 are possible equilibria. The reason for the multiplicity is the existence of externalities among λ_H investors. A λ_H investor benefits from having other investors of her type investing in the same type of investment. This is because, then, when she tries to sell the project, the price will not be that low since the market knows that the sale is very likely to be driven by a liquidity shock. As a result, when all λ_H investors invest in portfolio investments, an individual λ_H investor would like to do the same thing in order to avoid the low price of a direct investment in case she needs to sell (given that he needs to sell quite often). Similarly, when all λ_H investors invest in direct investments, an individual λ_H investor would like to invest in a direct investment as well. This multiplicity may generate jumps from an equilibrium with a lot of direct investments to an equilibrium with much less direct investments. This may explain why some countries have more direct investments than other countries with similar characteristics, and why some periods of time are characterized by more direct investments than others. The existence of multiple equilibria also generates interesting welfare implications that will be discussed below.

2.4.2 The Probability of Early Withdrawals

Our analysis thus far showed that whenever the two types of investments coexist in equilibrium, portfolio investors will be more likely than direct investors to get a liquidity shock that forces them to sell their investments in the short term. This, however, does not necessarily imply that FPI's are being liquidated more often than FDI's in equilibrium. This is because, in our model, FDI's are being liquidated, not only because of a liquidity shock, but sometimes due to a low realization of ε.

To see this formally, consider Case 3 as n equilibrium. This is the equilibrium where all λ_H investors choose FPI and λ_L investors choose FDI. Thus, it exhibits the largest difference between the expected liquidity needs of FPI investors and those of FDI investors. In this equilibrium, the probability of an early withdrawal of FPI is λ_H, whereas the probability of an early withdrawal of FDI is $\lambda_L + (1-\lambda_L)G\left(\underline{\varepsilon}_D\left(\lambda_L\right)\right)$. Analyzing the difference between these two expressions, we can see that there are two opposite effects. On the one hand, since $\lambda_H > \lambda_L$, the probability of an early withdrawal that is driven by a liquidity shock is greater for FPI than for FDI. But, on the other hand, there is a probability of $(1-\lambda_L)G\left(\underline{\varepsilon}_D\left(\lambda_L\right)\right)$ that an FDI will be sold in period 1 due to a low realization of ε. This possibility does not exist with an FPI. The condition, under which portfolio investments are being liquidated more often in period 1 than direct investments, is then:

$$\lambda_H > \lambda_L + (1-\lambda_L)G\left(\underline{\varepsilon}_D\left(\lambda_L\right)\right). \tag{2.12}$$

Since the left hand side of this condition increases in λ_H and the right hand side decreases in λ_H (recall that $\lambda_H = 1 - \lambda_L$), this condition implies that portfolio investments will be liquidated more often in the short term as long as λ_H is high enough, or, in other words, as long as the heterogeneity among investors is sufficiently strong.

2.4.3 Welfare Analysis

Our model has interesting welfare implications for the region of parameters with multiple equilibria - that is, when $B < B^*$ and $\lambda_H^*(B) < \lambda_H < \lambda_H^{**}(B)$.

As a starting point, we analyze foreign investors' welfare in this region of parameters, given the current framework. then, we use the result to study the implications for the welfare of residents of the host country which is the main focus of our welfare analysis.

When $B < B^*$ and $\lambda_H^*(B) < \lambda_H < \lambda_H^{**}(B)$, our model has three equilibria: Case 1, Case 2, and Case 3. Our analysis shows that when these three equilibria are possible, Case 1 represents a Pareto improvement over the other two equilibria. To see this, note that under Case 3, λ_H investors choose FPI and gain an expected payoff of $\frac{1}{2B}$. Similarly, under Case 2, they are indifferent between FDI and FPI, and thus also gain an expected payoff of $\frac{1}{2B}$. Under Case 1, however, they choose FDI and gain an expected payoff of $EV_{Direct}(\lambda_H, \frac{1}{2}, B)$. We know that this payoff is greater than $\frac{1}{2B}$, since λ_H investors chose to get it rather than to invest in FPI and get $\frac{1}{2B}$. Thus, λ_H investors are better off under Case 1, where they benefit from the higher efficiency of FDI. When other equilibria occur in this range of parameters, it is because of a coordination failure: λ_H investors choose not to invest in FDI because they believe other λ_H investors will not invest in FDI, and thus will reduce the expected value of this type of investment.

As for λ_L investors, in all three equilibria they choose FDI. Under Case 3, their expected payoff is $EV_{Direct}(\lambda_L, \lambda_L, B)$; under Case 1, it is $EV_{Direct}(\lambda_H, \frac{1}{2}, B)$; while under Case 2, it gets a value between $EV_{Direct}(\lambda_L, \lambda_L, B)$ and $EV_{Direct}(\lambda_H, \frac{1}{2}, B)$. Since $EV_{Direct}(\lambda_i, \lambda_D, B)$ is increasing in λ_D, we know that λ_L investors are better off in Case 1. The reason is that under Case 1, all λ_H investors choose direct investments, and thus the price of direct investments in period 1 is higher.

With these results in mind, let us address the differences in welfare from the point of view of the residents of the host country. Up to this point, the residents of the host country did not have an explicit role in our model. A natural way to introduce them is to assume that they own the domestic project initially, and sell them to foreign investors. Consider a representative host country resident who owns one project in period 0 and behaves competitively. At this time, she sells the projects to a foreign investor. After the sales have taken place, the events in the model are exactly the same as we described before: in period 0, foreign investors choose the form of investment, and in period 1 they make a decision on whether to sell their investments or not. Given this structure, the welfare analysis from the point of view of the host-country representative resident boils down to analyzing the price that she gets for her project in period 0.

In period 0, there are two types of foreign investors buying the investment projects from the residents of the host country: λ_H investors and λ_L investors. Since the type of each investor is not observable, in a competitive equilibrium, the price of projects in period 0 will be determined by the lowest between the value that is incurred to λ_H investors and the value that is incurred to λ_L investors from holding the project. In our model, this is always the value that is incurred to λ_H investors. Thus, λ_L investors capture some of the rent due to their ability to maintain the project for a long time, and λ_H investors do not capture any rent.

The price that host-country resident gets for the project in period 0 will then be $\frac{1}{2B}$ when either Case 2 or Case 3 is the realized equilibrium, and $EV_{Direct}(\lambda_H, \frac{1}{2}, B)$ when Case 1 is the realized equilibrium. As we showed above, in the region where all three equilibria are possible, the first expression is lower than the second one, meaning that the domestic resident gets higher price when Case 1 is the realized equilibrium. This suggests that the host country may benefit from encouraging more investments to be in the form of FDI.

2.5 Conclusion

The model we developed in this chapter describes an information-based trade off between direct investments and portfolio investments. In the model, direct investors are more informed about the fundamentals of their projects. This information enables them to manage their projects more efficiently. However, it also creates an asymmetric-information problem in case they need to sell their projects permanently, and reduces the price they can get in that case. As a result, investors, who know they are more likely to get a liquidity shock that forces them to sell early, are more likely to choose portfolio investments, whereas investors, who know they are less likely to get a liquidity shock, are more likely to choose direct investments.

The model generates several results that are consistent with empirical evidence. First, developed economies attract larger shares of FPI than developing economies. This may happen because the high labor costs in developed economies make the projects there less profitable, and thus make it less beneficial to incur the fixed costs associated with FDI. Moreover, the high transparency in developed economies makes FPI there more efficient. Second, because investors with high expected liquidity needs are attracted to FPI, while those with low expected liquidity needs are attracted to FDI, our model can account for the high observed withdrawal rates of FPI relative to FDI, which also contribute to a high volatility of the former relative to the latter. Third, developed economies with high levels of transparency are expected to have smaller differences between the withdrawal ratios of FPI and those of FDI. This is because the high efficiency of FPI in those economies attracts more investors with low expected liquidity needs to FPI, and prevents complete separation in equilibrium between investors with low expected liquidity needs and those with high expected liquidity needs.

2A. Appendix

A key in the characterization of the equilibrium allocation of investors between FDI and FPI is the threshold value B^*, which is defined by the following equation:

$$Diff(\frac{1}{2}, \frac{1}{2}, B^*) = 0. \tag{2.A.1}$$

Proposition: *(i) for any $B < B^*$, there exist $\lambda_H^*(B)$ and $\lambda_H^{**}(B)$, where $\frac{1}{2} < \lambda_H^*(B) < \lambda_H^{**}(B) < 1$ and both $\lambda_H^*(B)$ and $\lambda_H^{**}(B)$ are strictly decreasing in B. Then, when $\frac{1}{2} < \lambda_H < \lambda_H^*(B)$, the only possible equilibrium is Case 1; when $\lambda_H^*(B) < \lambda_H < \lambda_H^{**}(B)$, the possible equilibria are Case 1, Case 2, and Case 3; and when $\lambda_H^{**}(B) < \lambda_H < 1$, the only possible equilibrium is Case 3.*

(ii) For any $B > B^*$, there exists $\lambda_H^{***}(B)$, where $\frac{1}{2} < \lambda_H^{***}(B) \leq 1$ and $\lambda_H^{***}(B)$ is strictly increasing in B. Then, when $\frac{1}{2} < \lambda_H < \lambda_H^{***}(B)$, the only possible equilibrium is Case 5; and when $\lambda_H^{***}(B) < \lambda_H < 1$, the only possible equilibrium is Case 3.[12]

Proof:

We start by defining the condition for each case to be an equilibrium. We base these conditions on: The equilibrium choice of agents between FDI and FPI, as defined in Section 2.3; the equilibrium value of λ_D, as defined in Section 2.4 (including the off-equilibrium assumptions); and the properties of the function $D(\lambda_i, \lambda_D, B)$, which is defined by:

$$D(\lambda_i, \lambda_D, B) = (1-\lambda_i)\begin{bmatrix} \int_{-1}^{\underline{\varepsilon}_D(\lambda_D)} \left(1+\underline{\varepsilon}_D(\lambda_D)\right)^2 g(\varepsilon)d\varepsilon \\ + \int_{\underline{\varepsilon}_D(\lambda_D)}^{1} (1+\varepsilon)^2 g(\varepsilon)d\varepsilon \end{bmatrix} + \lambda_i \cdot \left(1+\underline{\varepsilon}_D(\lambda_D)\right)^2 - 2BC - 1. \tag{2.A.2}$$

Then, the decision of investors between FDI and FPI depends on the sign of $D(\lambda_i, \lambda_D, B)$. An increase (decrease) in $D(\lambda_i, \lambda_D, B)$ makes it more likely that the investor will choose FDI (FPI). We now show that the signs of the derivative of $D(\lambda_i, \lambda_D, B)$ with respect to the different parameters support the four parts of the proposition.

(i)

$$\frac{\partial D(\lambda_i, \lambda_D, B)}{\partial C} = -2B < 0.$$

(ii)

$$\frac{\partial D(\lambda_i, \lambda_D, B)}{\partial B} = -2C < 0.$$

(iii)

$$\begin{aligned}\tfrac{\partial D(\lambda_i,\lambda_D,B)}{\partial \lambda_i} &= (1+\underline{\varepsilon}_D(\lambda_D))^2 - \textstyle\int_{-1}^{\underline{\varepsilon}_D(\lambda_D)} (1+\underline{\varepsilon}_D(\lambda_D))^2\, g(\varepsilon)d\varepsilon - \int_{\underline{\varepsilon}_D(\lambda_D)}^{1} (1+\varepsilon)^2 g(\varepsilon)d\varepsilon \\ &= \textstyle\int_{\underline{\varepsilon}_D(\lambda_D)}^{1} \left[(1+\underline{\varepsilon}_D(\lambda_D))^2 - (1+\varepsilon)^2\right] g(\varepsilon)d\varepsilon < 0.\end{aligned}$$

(iv)

$$\frac{\partial D(\lambda_i, \lambda_D, B)}{\partial \lambda_D} = 2\frac{\partial \underline{\varepsilon}_D(\lambda_D)}{\partial \lambda_D} (1+\underline{\varepsilon}_D(\lambda_D)) \left[(1-\lambda_i)\, G\left(\underline{\varepsilon}_D(\lambda_D)\right) + \lambda_i\right] > 0,$$

because $\frac{\partial \underline{\varepsilon}_D(\lambda_D)}{\partial \lambda_D} > 0$.

Case 1 is an equilibrium iff $D(\lambda_i, \lambda_D, B) \geq 0$. Case 2 is an equilibrium iff $D(\lambda_i, \lambda_D, B) \leq 0$ and $D(\lambda_H, \frac{1}{2}, B) \geq 0$. Case 3 is an equilibrium iff $D(\lambda_H, \lambda_L, B) \leq 0$ and $D(\lambda_L, \lambda_L, B) \geq 0$. Case 4 is an equilibrium iff $D(\lambda_L, \lambda_L, B) = 0$. Case 5 is an equilibrium iff $D(\lambda_L, \lambda_L, B) \leq 0$.

Now, we define the thresholds $\lambda_H^*(B)$, $\lambda_H^{**}(B)$, and $\lambda_H^{***}(B)$ that are included in the proposition. Threshold $\lambda_H^*(B)$ is defined by the equation

$D(\lambda_H^*(B), \lambda_L^*(B), B) = 0$ (here, $\lambda_L^*(B) \equiv 1 - \lambda_H^*(B)$). Threshold $\lambda_H^{**}(B)$ is defined by: $D(\lambda_H^{**}(B), \frac{1}{2}, B) = 0$. Finally, threshold $\lambda_H^{***}(B)$ is defined by: $D(\lambda_L^{***}(B), \lambda_L^{***}(B), B) = 0$ (here, $\lambda_L^{***}(B) \equiv 1 - \lambda_H^{***}(B)$).

Now, we characterize these thresholds as functions of B. The illustration provided in Figure 2.2 can help in understanding this part of the proof. First, note that $\lambda_H^*(B^*) = \lambda_H^{**}(B^*) = \lambda_H^{***}(B^*) = \frac{1}{2}$. (Recall that B^* is defined in (2.A.1).) Then, by the properties of $D(\lambda_i, \lambda_D, B)$, we know that both $\lambda_H^*(B)$ and $\lambda_H^{**}(B)$ are decreasing in B, and that $\lambda_H^{**}(B) > \lambda_H^*(B)$. Also, by examining the function $D(\lambda_i, \lambda_D, B)$ we can tell that $\lambda_H^{**}(B) < 1$. Finally, since $D(\lambda, \lambda, B)$ is decreasing in λ (this is shown at the end of this proof) and in B, $\lambda_H^{***}(B)$ is increasing in B.

Using the equilibrium conditions, we can now specify when each equilibrium will occur relative to the thresholds defined above. This specification relies on the properties of the function $D(\lambda_i, \lambda_D, B)$, and on the property that $D(\lambda, \lambda, B)$ is decreasing in λ, which will be shown below. Case 1 is an equilibrium iff $\lambda_H \leq \lambda_H^{**}(B)$. Case 2 is an equilibrium iff $\lambda_H^*(B) \leq \lambda_H \leq \lambda_H^{**}(B)$. Case 3 is an equilibrium iff $\lambda_H \geq \lambda_H^*(B)$ and $\lambda_H \geq \lambda_H^{***}(B)$. Case 4 is an equilibrium iff $\lambda_H = \lambda_H^{***}(B)$. Case 5 is an equilibrium iff $\lambda_H \leq \lambda_H^{***}(B)$. This leads to the characterization of equilibrium outcomes stated in the proposition.

To complete the proof, we need to show that $D(\lambda, \lambda, B)$ is decreasing in λ. We know that:

$$D(\lambda,\lambda,B) = (1-\lambda)\begin{bmatrix} \int_{-1}^{\underline{\varepsilon}_D(\lambda)} (1+\underline{\varepsilon}_D(\lambda))^2 g(\varepsilon)d\varepsilon \\ +\int_{\underline{\varepsilon}_D(\lambda)}^{1}(1+\varepsilon)^2 g(\varepsilon)d\varepsilon \end{bmatrix} + \lambda\cdot(1+\underline{\varepsilon}_D(\lambda))^2 - 2BC - 1.$$

Thus,

$$\begin{aligned}\frac{\partial D(\lambda,\lambda,B)}{\partial\lambda} &= \int_{\underline{\varepsilon}_D(\lambda)}^{1}\left[(1+\underline{\varepsilon}_D(\lambda))^2-(1+\varepsilon)^2\right]g(\varepsilon)d\varepsilon \\ &\quad +2\tfrac{\partial\underline{\varepsilon}_D(\lambda)}{\partial\lambda}\left(1+\underline{\varepsilon}_D(\lambda)\right)\left[(1-\lambda)\,G(\underline{\varepsilon}_D(\lambda))+\lambda\right].\end{aligned}$$

Plugging in the expression for $\frac{\partial\underline{\varepsilon}_D(\lambda)}{\partial\lambda}$, as is implied by equations (2.5) and (2.6), we get:

$$\begin{aligned}\frac{\partial D(\lambda,\lambda,B)}{\partial\lambda} &= \int_{\underline{\varepsilon}_D(\lambda)}^{1}\left[(1+\underline{\varepsilon}_D(\lambda))^2-(1+\varepsilon)^2\right]g(\varepsilon)d\varepsilon \\ &\quad +\left(1-(1+\underline{\varepsilon}_D(\lambda))^2\right)+\int_{-1}^{\underline{\varepsilon}_D(\lambda)}\left[(1+\underline{\varepsilon}_D(\lambda))^2-(1+\varepsilon)^2\right]g(\varepsilon)d\varepsilon \\ &= 1-\int_{-1}^{1}(1+\varepsilon)^2 g(\varepsilon)d\varepsilon.\end{aligned}$$

Notes

[1]Note that in subsequent chapters this cost may consist of both host and source country labor cost.

[2]Albuquerque (2003) provides another explanation as to why FDI may be sold only at a loss. He argues that foreign direct investors typically transfer some knowledge into the firm they acquire. This also makes it hard to withdraw capital as the owner cannot guarantee that these intangible resources will remain with the firm.

[3]See Klein, Peek and Rosengren (2002) for evidence on the accessibility of multinationals to bank credits.

[4]Regarding *net* foreign equity flows, the volatility of FDI is, in general, much smaller than that of FPI.

Using World Bank data on 111 countries, Albuquerque (2003) shows that 89% of the countries in his sample have lower coefficient of variation of net FDI inflows than that of other net inflows. A related set of evidence suggests that FDI has proven to be much more resilient during financial crises, and thus contributes to the stability of the host country; see Chuhan, Perez-Quiros and Popper (1996), Frankel and Rose (1996), Lipsey (2001), and Sarno and Taylor (1999). Moreover, empirical analysis has established that the differences in volatility between FPI and FDI flows are much smaller for developed economies than for developing economies.

Lipsey (2000) shows that the ratio of FDI's volatility to other long-term flows' volatility is 0.59 in Latin America, 0.74 in South East Asia, 0.86 in Europe, and 0.88 in the US. Thus, the differences in volatility between net FDI inflows and other types of net inflows are smaller in developed economies.

[5]Although we write this chapter in the context of international capital flows, we believe the mechanism we suggest here is more general, and can serve to analyze the trade off between direct investments and portfolio investments, or between management efficiency and liquidity, in other contexts. In a related paper, Bolton and von-Thadden (1998) analyze a trade off between direct investments and portfolio investments. This model, however, is not based on the differences in information that each one of these investments provides. Kahn and Winton (1998) and Maug (1998) study models where the information held by institutional investors does not always improve the value of the firm, as institutional investors might use this information to make trading profits instead of to improve firm performance. These models do not look, however, at the decision of the investors on whether to acquire information when they might get liquidity shocks. This chapter also touches on other issues that have been discussed in the finance literature. Admati and Pfleiderer (1991) discuss the incentive of traders to reveal the fact that they are trading for liquidity reasons and not because of bad information. Admati and Pfleiderer (1998) and Foster and Viswanathan (990) point to the existence of externalities among traders who trade for liquidity reasons.

[6]The argument, according to which the manager wishes to make larger investments and build an empire is common in the corporate finance liter-

ature (see Jensen (1986)). In such a case, if the owner cannot verify the information that the manger had at the time of the decision, she will not be able to prove that the manager acted to maximize his own objective function. As a result, a contract that instructs the manger to maximize the value of the firm given his information will not be enforceable.

The agency problem is not modelled explicitly here because we want to focus instead on its implications for the trade off between direct investments and portfolio investments. What we do, however, capture in our model is the spirit of the agency problem, and the inefficiency associated with the fact that the owner of the project is not the manager.

[7]Jensen's inequity states that $F\left[E(\tilde{x})\right] < E\left[F(\tilde{x})\right]$, if F is strictly convex and $\tilde{x}$ is a non-degenerate random variable.

[8]This feature of the model is similar to the preference-shock assumption made by Diamond and Dybvig (1983): an investor who is subject to a liquidity shock derives her utility only from period-1 consumption. If, however, she is subject to a liquidity shock, she derives her utility from period-2 consumption. As a result, an investor who is subject to a liquidity shock is forced to sell the project in period 1, because she cannot afford to wait and collect the payoff from the project in period 2.

[9]Note that our results hold in a more general setting, that is, for any $\lambda_H > \lambda_L$.The assumption that $\lambda_H + \lambda_L = 1$ allows us to change the difference between λ_H and λ_L by changing only one of these two parameters.

[10]Note that this is just a technical assumption regarding the procedures of

the sales. It does not qualitatively affect the results.

[11]It should be noted that if all investors choose FPI in equilibrium, λ_D cannot be defined by the above equation. This is because in such an equilibrium, investors are not expected to choose FDI at all. Thus, we need to make an off-equilibrium assumptions to determine λ_D in case that an investor diverges from that equilibrium and chooses to hold a direct investment. Since λ_L investors have greater incentives to hold direct investments than λ_H investors, we assume that, in an equilibrium where all investors choose FPI, if an investor diverges and invests in FDI, the market assesses a probability of λ_L to the event that this investor had a liquidity shock. Note that this off-equilibrium assumption is not important for our results.

[12]For brevity, we do not characterize here the equilibrium outcomes for the specific values: $B = B^*$; $\lambda_H = \lambda_H^*$; $\lambda_H = \lambda_H^{**}$; $\lambda_H = \lambda_H^{***}$.

Chapter 3

Foreign versus Domestic Direct Investment: Cream-Skimming

3.1 Introduction

In the preceding chapter we attempted to explain the allocation of foreign investors between FDI and FPI. We highlighted a key difference between the two types of investments. In the case of FDI, both ownership and control of a firm is acquired, whereas in the case of FPI control is not achieved. The acquisition of control entails FDI investors an efficiency gain, but at a fixed cost of becoming an FDI investors and a variable information-based cost associated with liquidity needs. A balance between the benefit and cost of FDI generates an equilibrium assignment of investors to foreign direct investments and foreign portfolio investments, depending on some liquidity need characteristics of the investors.

The focus of this book is FDI. Having analyzed the endogenous formation of this type of investments, relative to portfolio investments, we turn from now on to analyze other key aspects of foreign direct investors in relation to domestic investors. In doing so, we abstract from the liquidity aspect and take the formation of investor types to be exogenous.

In this chapter we analyze a screening mechanism through which foreign direct investors manifest their comparative advantage over domestic investors in gaining control over domestic firms.

We develop a simple model in which the industry specialization in the source country provides a comparative advantage to potential foreign direct investors in eliciting good investment opportunities in the host country, relative to domestic investors. The advantage stems from the ability of FDI investors to apply better industry-specific, micro-management standards (an "intangible capital"). The advantage of FDI investors in their cream-skimming skills is less pronounced when corporate transparency and capital market institutions are of high quality; in which case FDI inflows are less abundant.[1]

This chapter also suggests that the gains from FDI to the host country are reflected in a more efficient size of the stock of domestic capital and its allocation across firms. Domestic firms that are controlled by FDI investors are typically the "cream" (high-productivity firms). The magnitude of these non-traditional gains from trade that arise in our setup depends crucially (and inversely) on the degree of competition among potential FDI investors over the domestic firms. The non-traditional gains can vanish entirely if there is no such competition. Also, FDI inflows could make the size of the aggregate stock of domestic capital larger than otherwise (under some plausible assumptions).

This result is consistent with some recent empirical evidence. For instance, Borenzstein, De Gregorio and Lee (1998) and Bosworth and Collins (1999) provide such evidence for a sample of developing countries during the period 1978-1995. More recently, in a sample of developing countries, Razin (2005) finds that the effect of FDI inflows on domestic investment is significantly larger than other forms of foreign flows of capital.[2] He also provides evidence that FDI inflows promote efficiency: The effect of FDI on GDP growth is higher than the effect of other forms of foreign investment. In Part Three we also provide an empirical illustration of some implications of the analysis presented here. It demonstrates how host-country transparency, relative to the source-country transparency, affects bilateral FDI flows from source to host countries; either through the source-host selection channel, or by affecting the intensity of the FDI flows.

3.2 FDI and Skimming High-Productivity Firms

As we abstracted here from the issue of the endogenous formation of FDI (and PFI) investors, we simplify the analysis by switching from the three-period framework of the preceding chapter to a two-period setup. Assume a large number (N) of $ex - ante$ identical domestic firms in an industry. As before, each firm employs capital input (K) in the first period, in order to produce a single composite good in the second period. We assume that capital depreciates at the rate $\delta(< 1)$. Output in the second period is equal to $F(K)(1+\varepsilon)$, where F is a production function, which exhibits diminishing marginal productivity of capital.[3] The term ε is a firm-specific productivity factor. We assume that ε is bounded from below by -1, so that output is always non-negative. For notational ease, we also assume that ε is bounded from above by 1. Suppose that ε is purely idiosyncratic, so that there is no aggregate uncertainty in the model. As before, consumers-investors are well diversified and will thus behave in a risk-neutral way. We denote by G the cumulative distribution function of ε, and by $g = G'$ the corresponding density function.

At the starting point of the agents' decision process, the productivity factor (ε) of each firm is not revealed with full accuracy. Rather, each firm receives a signal ε' about its productivity, which is common knowledge.[4] The true ε of the firm is within an interval of $\pm\beta$ around ε'. Formally, given ε' the true value of ε is distributed according to the distribution of the productivity factor, conditional on its being in the interval $(\varepsilon'-\beta, \varepsilon'+\beta)$. The conditional distribution is therefore:

$$\varphi(\varepsilon/\varepsilon') = \frac{G(\varepsilon) - G(\varepsilon' - \beta)}{G(\varepsilon' + \beta) - G(\varepsilon' - \beta)}. \tag{3.1}$$

This conditional distribution denotes the cumulative distribution function of ε, conditional on the signal ε'. We assume that the signal ε' is distributed according to the distribution function G.

The firm chooses the level of the capital stock (and investment), denoted by $K(\varepsilon')$, after the signal ε' is received, so as to maximize its expected market value, conditional on ε'. This maximized value is:

$$V(\varepsilon') = \int\limits_{\varepsilon'-\beta}^{\varepsilon'+\beta} \left\{ \frac{F[K(\varepsilon')](1+\varepsilon) + (1-\delta)K(\varepsilon')}{1+r} - [K(\varepsilon') - (1-\delta)K_0] \right\} d\varphi(\varepsilon/\varepsilon'), \tag{3.2}$$

where $(1-\delta)K_0$ is the initial stock of capital, and r is the world rate of interest.[5] (Free mobility of debt flows fixes the domestic interest rate at the world rate.) The optimal stock of capital in this case, $K(\varepsilon')$, is implicitly defined by the first-order condition:

$$\int\limits_{\varepsilon'-\beta}^{\varepsilon'+\beta} \left[\frac{F'(K)(1+\varepsilon) + (1+\delta)}{1+r} - 1 \right] d\varphi(\varepsilon/\varepsilon') = 0.$$

This expression can be simplified to:

$$F'[K(\varepsilon')][1 + E(\varepsilon/\varepsilon')] = r + \delta, \tag{3.3}$$

where $E(\varepsilon/\varepsilon')$ is the conditional expected value of the productivity factor, given that this factor lies within the interval $(\varepsilon' - \beta, \varepsilon' + \beta)$, that is:

$$E(\varepsilon/\varepsilon') = \int_{\varepsilon'-\beta}^{\varepsilon'+\beta} \varepsilon d\varphi(\varepsilon/\varepsilon'). \tag{3.4}$$

The level of capital $K(\varepsilon')$ is based on an imprecise signal ε' rather than on the true productivity of the firm. We refer to it as a signal-based optimal stock of capital, to distinguish it from what will be defined later as the productivity-dependent optimal stock of capital, which depends directly on the true productivity of the firm, ϵ.

Suppose that there is a screening (or search) technology, which, at some fixed cost per firm, can elicit the true value of the productivity factor of the firm, ε. A potential buyer can apply the technology after she acquires and then gains control of the domestic firm. We assume that foreign direct investors have a cutting-edge advantage over domestic investors in extracting information about the true value of the firm. If foreign direct investors acquire a domestic firm, they can apply their superior micro-management skills in order to elicit the true value of ε. This advantage stems from some sort of "intangible capital" (specialized knowledge) in this particular industry. The basic idea is that firms get involved in foreign operations in order to exploit this unique advantage that they have accumulated over time in their source country. The advantage is modeled here by specifying a lower screening cost for foreign direct investors than for domestic investors. Formally, the cost per firm for a foreign direct investor is C_F, which is assumed to be lower than C_D, the corresponding cost for a domestic direct investor (i.e., a domestic investor who gains and acquires control of the domestic firm). This sort of comparative advantage is industry based, and is therefore not confined solely to flows of FDI from developed to developing countries. Thus, this sort of comparative advantage can explain also two-way FDI flows among (multi-industry) developed countries.

If the true value of ε were to be known, then the firm would choose an optimal capital stock, denoted by $K^*(\varepsilon)$, according to the marginal productivity condition:

$$F'\left[K^*(\varepsilon)\right](1+\varepsilon) = r + \delta. \tag{3.5}$$

This stock of capital is referred to as the productivity-dependent stock of capital.

Given the signal ε', a potential foreign direct investor knows that the true value of ε must lie between $\varepsilon' - \beta$ and $\varepsilon' + \beta$, and that she will be able to elicit the true value of ε if she purchases the firm, at a cost C_F. Therefore, her gross bid price, given the signal ε', is given by:

$$P(\varepsilon') = \int_{\varepsilon'-\beta}^{\varepsilon'+\beta} \left\{ \frac{F[K^*(\varepsilon)](1+\varepsilon) + (1-\delta)K^*(\varepsilon)}{1+r} - [K^*(\varepsilon) - (1-\delta)K_0] \right\} d\varphi(\varepsilon/\varepsilon'). \tag{3.6}$$

Her net bid price is then $P(\varepsilon') - C_F$. Because C_F is smaller than C_D, the bid price of the foreign direct investor is higher than that of the domestic investor. When foreign direct investors are abundant, competition among them will indeed drive up the price of the domestic firms to $P(\varepsilon') - C_F$.[6]

Given the signal ε', the value of information to the FDI investor (that is, the value of eliciting the true productivity of the firm) is $P(\varepsilon') - V(\varepsilon') > 0$. To see that indeed $P(\varepsilon') > V(\varepsilon')$, note that in calculating $P(\varepsilon')$ it is assumed that K is optimally adjusted for the true value of the productivity, whereas in calculating $V(\varepsilon')$ it is required that $K(\varepsilon')$ is fixed across all firms in the $\beta-$interval around ε'. The associated cost of gaining the information is denoted by C_F. In order to incur this cost, the value of information must exceed this cost. Naturally, one would expect the value of information to rise with ε'. This is because, given the signal ε', the deviations of the productivity-independent signal-based $K(\varepsilon')$ over the interval $(\varepsilon' - \beta, \varepsilon' + \beta)$, from the productivity-dependent $K^*(\varepsilon)$ over this interval, and consequently the deviations of $F(K^*(\varepsilon'))$ from $F(K(\varepsilon'))$ over this interval, are magnified by the productivity factor $1+\varepsilon$. We therefore assume indeed that $P(\varepsilon') - V(\varepsilon')$ rises with ε'.[7] Hence, there exists a cutoff level of the signal, denoted by $\varepsilon_0^{'}$, such that for all $\varepsilon' < \varepsilon_0^{'}$, the bid-ask price differential $P(\varepsilon^{'}) - C_F - V(\varepsilon^{'})$ is negative and, similarly, for all $\varepsilon' > \varepsilon_0'$, the bid-ask price differential is positive. Thus, all the firms that receive a low-productivity signal will be retained by the original (domestic) owners, whereas all the firms that receive a high-productivity signal will be acquired by foreign direct investors, who manage to outbid their domestic counterparts concerning the high-productivity firms. The cutoff level of the signal depends on the screening cost C, and is defined implicitly by:

$$P\left[\varepsilon_0'(C)\right] - C = V\left[\varepsilon_0'(C)\right]. \tag{3.7}$$

With FDI investors who can do the screening at a cost C_F per firm, the cutoff level of the signal is a function of C_F , denoted by $\varepsilon_{0F}' \equiv \varepsilon_0'(C_F)$.

The assumption that $P(\varepsilon') - V(\varepsilon')$ rises with ε' implies also that as the screening cost (C_F) of the FDI investors falls, the cutoff productivity level (that is, ε_{0F}') declines as well. This means that with a fall in C_F, more firms will be acquired by FDI investors. Therefore, a lower screening cost of FDI investors gives rise to a larger volume of FDI inflows.[8] By the same token, as the signal becomes more accurate (that is, as β becomes smaller), the benefit of the screening technology, which is $P(\varepsilon')-V(\varepsilon')$, declines. We may interpret a more accurate signal as an improvement in corporate transparency. The advantage of FDI investors in their cream-skimming skills is less pronounced when host-country corporate transparency improves,[9] and FDI inflows are expected to be less abundant. After the signals are revealed, then a firm with a signal ε', below ε_{0F}', actually adjusts its capital stock to the signal-dependent, productivity-independent level $K(\varepsilon')$. But a firm, which receives a signal ε' above ε_{0F}', expects to adjust its capital stock to a productivity-dependent level $K^*(\varepsilon)$ with a cumulative distribution function $\varphi(\varepsilon/\varepsilon')$. The expected value of its capital stock, denoted by $E[K^*(\varepsilon)/\varepsilon']$ is given by:

$$E[K^*(\varepsilon)/\varepsilon'] = \int_{\varepsilon'-\beta}^{\varepsilon'+\beta} K^*(\varepsilon) d\varphi(\varepsilon/\varepsilon'). \tag{3.8}$$

Thus, the total expected value of the stock of capital (before signals are revealed) is:

$$K^F = \int_{-1}^{\varepsilon'_{0F}} K(\varepsilon')dG(\varepsilon') + \int_{\varepsilon'_{0F}}^{1} E[(K^*(\varepsilon)/\varepsilon']dG(\varepsilon'). \tag{3.9}$$

This is our measure of the size of domestic capital.

3.3 FPI Inflows Versus FDI Inflows

To understand the unique role played by FDI, suppose now that instead of FDI inflows there are only FPI inflows. Note that portfolio investors (whether foreign or domestic) do not acquire control and management. That is, they do not incur the screening cost and do not elicit the true productivity parameter ε. In this case, only domestic direct investors acquire and gain control of the firms with the high-productivity signals. Domestic and FPI investors will acquire all the other firms with low-productivity signals. The cutoff level of the signal in this case is $\varepsilon'_{0D} \equiv \varepsilon'_0(C_D)$. Because $C_D > C_F$, it follows that $\varepsilon'_{0F} < \varepsilon'_{0D}$ [see equation (3.7) and recall that $P(\varepsilon') - V(\varepsilon')$ is increasing in ε', by assumption]. Thus, the difference in investment in capacity between the two regimes lies only in the range of signals between ε'_{0F} and ε'_{0D}. The capital stock of a firm with a signal below ε'_{0F} is the same in the two regimes. The expected capital stock of a firm with a signal above ε'_{0D} will also be the same in the two regimes. But a firm, which receives a signal ε' in-between these two cutoff levels, will invest a signal-dependent, productivity-independent $K(\varepsilon')$ in the foreign portfolio-investment regime, compared to a productivity-dependent schedule, $K^*(\varepsilon)$, with a cumulative distribution $\varphi(\varepsilon/\varepsilon')$, in the FDI regime. Naturally, the latter is more efficient, in the sense that it yields a higher expected return.[10]

3.3.1 Gains to the Host Country

The economic gains from FDI, relative to FPI inflows, consist of the efficiency of investment and the lower screening cost of FDI investors. Note that because the same world interest rate, r, prevails in the home country in the two regimes, it follows that the gains from FDI relative to FPI in our case do not include the traditional gains from opening up the domestic capital market to foreign capital inflows. (Evidently, the traditional gains are present also in the portfolio regime.) In the FDI regime the firms with signals above the cutoff signal ε'_{0F} are screened; whereas in the FPI regime a smaller set of firms, namely only the firms with signals above ε'_{0D} are screened (recall that $\varepsilon'_{0D} > \varepsilon'_{0F}$). Therefore, the gains to the host country stemming from the efficiency of investment is:

$$GAIN_E = \int_{\varepsilon'_{0F}}^{\varepsilon'_{0D}} [P(\varepsilon') - C_F - V(\varepsilon')]dG(\varepsilon'). \tag{3.10}$$

In addition, for the firms that are screened in the two regimes (that is, the firms with signals above $\varepsilon_{0D}^{'}$), the screening cost is lower under the FDI regime than under the portfolio flow regime. This gives rise to further gains from FDI, which are

$$GAIN_C = (C_D - C_F)[1 - G(\varepsilon_{0D}^{'})]. \tag{3.11}$$

Observe that the entire gain, attributable to the lower screening cost of FDI investors, is captured by the host country because of the assumed **perfect competition among the FDI investors over the domestic firms**. Competition among FDI investors must drive up the price they pay for a domestic firm to their net bid-price [that is, $P(\varepsilon') - C_F$], which exceeds the ask-price of the domestic owners [that is, $V(\varepsilon')$]; except for the cutoff firm (for which the bid price and ask price are equal to each other). Thus, the total gain to the host country from FDI is

$$\begin{aligned} GAIN_E + GAIN_C \ &= \int_{\varepsilon_{OF}^{'}}^{\varepsilon_{0D}^{'}} [P(\varepsilon') - C_F - V(\varepsilon')]dG(\varepsilon') \\ &\quad + (C_D - C_F)[1 - G(\varepsilon_{0D}^{'})]. \end{aligned} \tag{3.12}$$

Note, however, that in the extreme opposite case of a monopoly, the single FDI investor will never offer a price for a domestic firm above the price that will be offered by domestic investors, which is $P(\varepsilon') - C_D$, as long as this price is above, or equal, to the ask price of the domestic owner, which is $V(\varepsilon')$.

Thus, the price at which the foreign direct investor buys a domestic firm with a signal ε' is $Max[P(\varepsilon') - C_D, V(\varepsilon')]$. Because $P(\varepsilon'_{0D}) - C_D = V(\varepsilon'_{0D})$, it follows that $P(\varepsilon') - C_D < V(\varepsilon')$ in the interval $(\varepsilon'_{0F}, \varepsilon'_{0D})$. This means that in this interval the domestic firms are purchased by the foreign direct investor at the ask price $V(\varepsilon')$. Hence, the efficiency gain of investment, $GAIN_E$, vanishes. Similarly, firms in the interval $[\varepsilon'_{0D}, 1]$ must be purchased at the price $P(\varepsilon') - C_D$ [rather than $P(\varepsilon') - C_F$ in the competitive case]. Hence, $GAIN_C$ vanishes as well. Thus, as expected, the entire gain from FDI accrues to the single FDI investor. Evidently, this is an extreme case. If there is an additional domestic input, say labor, the host country still gains, even in the case of a single FDI investor, through infra-marginal gains to domestic labor. However, these gains are sharply smaller than what they could have been in the case of competitive FDI investors.

To retain some of the gains of FDI, a possible remedy for the host country is to impose some sort of a floor on the sale prices of domestic firms. Another partial remedy for the host country is to impose a (source-based) capital gains tax on FDI investors. In the intermediate case of imperfect competition among a few FDI investors, but not a strict monopoly, the gains from FDI are split between the host country and the FDI investors.

3.3.2 The Size of Investment in Capacity in the Host Country

We have already established that the allocation of the capital stock (its aggregate level and distribution over firms) is more efficient in the FDI regime than in the portfolio regime. Is the capital stock also larger in the FDI regime than in the FPI regime? Recall that the fundamental difference between the two regimes is the screening cost C. Therefore, rephrasing the question one can ask whether a decline in the search cost increases the aggregate stock of capital. In order to answer this question, we write the aggregate stock of capital as a function of C, as follows [see equation (3.9)]:

$$\bar{K}(C) = \int_{-1}^{\varepsilon_0^{'}(C)} K(\varepsilon')dG(\varepsilon') + \int_{\varepsilon_0^{'}(C)}^{1} E[K^*(\varepsilon)/\varepsilon'[dG(\varepsilon'), \tag{3.13}$$

where, $\varepsilon_0^{'}(C), K(\varepsilon')$ and $E[K^*(\varepsilon)/\varepsilon']$ are defined by equations (3.7), (3.3) and (3.8), respectively.

Now, differentiate $\bar{K}(C)$ with respect to C, to get:

$$\frac{d\bar{K}(C)}{dC} = \{K\left[\varepsilon_0^{'}(C)\right] - E\left[K^*(\varepsilon)/\varepsilon_0^{'}(C)\right]\}g\left[\varepsilon_0^{'}(C)\right]\frac{d\varepsilon_0^{'}(C)}{dC}. \tag{3.14}$$

From equations (3.3) and (3.5) we can conclude that

$$K\left[\varepsilon_0^{'}(C)\right] = H\left\{E\left[\varepsilon/\varepsilon_0^{'}(C)\right]\right\}, \tag{3.15}$$

and

$$K^*(\varepsilon) = H(\varepsilon),$$

where the function $H(\cdot)$ is defined by

$$H(x) = (F')^{-1}\left(\frac{r+\delta}{1-x}\right),$$

and where the function $(F')^{-1}$ denotes the inverse of F'. Thus, we can rewrite equation (3.14) as:

$$\frac{d\bar{K}}{dC} = (H\left\{E\left[\varepsilon/\varepsilon_0^{'}(C)\right]\right\} - E\left[H(\varepsilon)/\varepsilon_0^{'}(C)\right])g\left[\varepsilon_0^{'}(C)\right]\frac{d\varepsilon_0^{'}(C)}{dC}. \quad (3.16)$$

If the function $H(\cdot)$ is convex, then it follows from Jensen's inequality that $d\bar{K}/dC$ is negative (because $d\varepsilon_0^{'}/dC > 0$). Indeed, one may plausibly assume that H is convex (for instance, this is the case with a Cobb-Douglas production function), in which case $d\bar{K}/dC < 0$. That is: The size of investment in capacity is larger under the regime of FDI inflows than under the regime of FPI inflows.

3.4 Conclusion

In this chapter, we develop a model in which foreign direct investors are better equipped to cream skim domestic projects than their direct domestic and portfolio counterparts, due to rich experience in the skimming of "good" firms. Employing this advantage, foreign direct investors are able to outbid direct domestic and portfolio investors for the good firms. Better hands-on management standards, which characterize FDI investors, entails a cutting-edge advantage over portfolio investors in reacting in real time to a changing business environment. This feature is naturally more pronounced in high-productivity firms, resulting in the acquisition of high-productivity firms by FDI investors. This mechanism applies both to mergers and acquisitions and to green-field investments. The productivity signal, though, is likely to be coarser in the latter, conveying less information about the true productivity. Thus, the advantage of the FDI investors over their domestic direct investors counterparts is even more pronounced in the case of green-field investments than in mergers and acquisition.

We emphasize that FDI as distinct from FPI investment with respect to the quality of monitoring management. Foreign direct investors, by definition, acquire some significant control over the firm they invest in, whereas portfolio investors, plagued by free-rider problems, have no control. Consequently, foreign direct investors can apply hands-on management (or micro-management) standards that would enable them to react in real time to changing economic environments. This feature may stem from "intangible capital" accumulated through a specialization by the foreign direct investors in a certain niche.[11] Indeed, there is some micro evidence in support of this hypothesis. For example, Djankov and Hoekman (2000) report that foreign direct investors pick the high-productivity firms in transition economies. Similarly, Griffith and Simpson (2003) find that foreign-owned manufacturing establishments in Britain, over the period 1980 to 1996, have significantly higher labor productivity than those that remain under domestic ownership.[12] Britain is not an exception as a developed country attracting FDI. In fact, the vast majority of FDI flows among developed countries. The model in this chapter is relevant when a developed country may specialize in certain niches and gain cutting-edge advantages in these niches over another developed country. The opposite may apply in some other niches.

In Razin and Sadka (2005), we provide some illustrative evidence on the effects of transparency on FDI flows. Transparency is proxied by accounting standards, creditors rights, etc.; see La Porta et al (1998). Consistent with the theoretical predictions of this chapter, less transparency in the host country, relative to the source country, encourages FDI.

Notes

[1]See also Wei (2000), Razin and Sadka (2003), and Albuquerque (2003).

[2]See also Chapter One.

[3]Here we specify separately the output (gross revenue) and the cost of the firm rather than just the net cash flow, as in Chapter Two.

[4]One can think of this signal as sort of encapsulated information about the firm's productivity, which is reported by its up-to-date financial statements.

[5]Because of the assumption that there is a single composite good, which serves both for investment and for consumption, it is irrelevant whether the optimal K is above or below $(1-\delta)K_0$.

[6]Presumably, foreign direct investors bring specific skills that entail them a cutting-edge screening advantage. Therefore, these skilled investors may not be abundant. The case of less than perfect competition among foreign direct investors is analyzed later in this chapter.

[7]We conjecture that this can be proved for a wide range of distribution and production functions. However, we were unable to characterize this range of functions. In any event, this assumption is irrelevant to our main point that FDI investors choose a more efficient level of capital than domestic investors. In case this assumption does not hold, that is $P(\varepsilon') - V(\varepsilon')$ is not monotonically increasing in ε', then the FDI investors do not necessarily acquire only the top productivity firms.

[8]We refer to the sum of the acquisition price of the firm and the investment

in its capacity (that is financed by the FDI owner) as FDI inflows.

[9]Indeed, these results also hold in Burstein's (2003) example, albeit with a different stochastic specification.

[10]We have assumed that the only advantage of FDI investors over direct domestic investors lies in the search/screening cost. Naturally, if we were to assume that FDI investors can also obtain better information about the true ε (we have assumed that both can accurately elicit ε), then the difference between the two regimes may expand to the entire range of $[-1, 1]$ of signals.

[11]See Gopinath (2004) for a different application of a search model for a study of FDI flows into developing economies.

[12]In addition, labor productivity improves faster over time and faster with age in foreign-owned establishments. Other studies found that this phenomenon is accounted for by the greater capital intensity of multinationals. For an overview see Lipsey (2002).

Chapter 4

FDI Flows with Endogenous Domestic Wages: Heterogenous Firms

4.1 Introduction

So far we focused on the host country as a recipient of FDI and other flows in a partial equilibrium context in which domestic input prices are fixed. Specifically, labor inputs and wages were fixed and therefore ignored. Naturally, the flows of FDI affect the domestic stock of capital. Therefore, even though the return to capital may be tied through capital mobility to the world rate of interest, wages may still vary. This chapter addresses the general-equilibrium interaction between wages and FDI. We cast this interaction in a bilateral source-host country setup.

4.2 Wage Determination

Consider a pair of countries, "host" and "source", in a world of free capital mobility which fixes the world rate of interest, denoted by r. We will now describe the host country, whose economic variables will be subscripted by "H". The description of the source country is similar with a subscript "S". Variables with neither an H nor S subscript are identical for the two countries. There is a representative industry whose product serves both for consumption and investment. As before, firms last for two periods. In the first period there exists a continuum of N_H firms which differ from each other by a productivity factor ε. The number N_H of firms (or entrepreneurs) is *fixed*. We refer to a firm which has a productivity factor of ε as an $\varepsilon-$firm. The cumulative distribution function of ε is denoted by $G(.)$, with a density function $g(.)$.

As before, we assume for simplicity that the initial net capital stock of each firm is the same and denote it by $(1-\delta)K_H^0$, where δ is the rate of physical depreciation. If an $\varepsilon-$firm invests I in the first period, it augments its capital stock to $K=(1-\delta)K_H^0+I$, and its gross output in the second period will be $A_H F(K,L)(1+\varepsilon)$, where L is the labor input (in effective units) and A_H is a country (H) - specific productivity parameter. Note that ε is firm-specific, whereas A_H is country-specific.

As before, we assume that there exists a fixed setup cost of investment, C_H, which is the same for all firms (that is, independent of ε). We assume that, due to some (suppressed) fixed factor, F is strictly concave, exhibiting diminishing returns to scale in K and L. Note that the average cost curve of the firm is U-shaped, so that perfect competition, which we assume, can prevail.[1] Consider an ε-firm which does invest in the first period an amount $I = K - (1-\delta)K_H^0$ in order to augment its stock of capital to K. Its present value becomes $V^+(A_H, K_H^0, \varepsilon, w_H) - C_H$, where

$$V^+(A_H, K_H^0, \varepsilon, w_H) = \max_{(K,L)} \left\{ \frac{A_H F(K,L)(1+\varepsilon) - wL + K}{1+r} - (K - K_H^0) \right\}, \tag{4.1}$$

and where we assume for notational simplicity that $\delta = 0$.

The demands of such a firm for K and L are denoted by $K^+(A_H, \varepsilon, w_H)$ and $L^+(A_H, \varepsilon, w_H)$. They are given by the marginal productivity conditions:

$$A_H F_K(K,L)(1+\varepsilon) = r, \tag{4.2}$$

and

$$A_H F_L(K,L)(1+\varepsilon) = w_H, \tag{4.3}$$

where F_K and F_L denote the partial derivatives of F with respect to K and L, respectively. As before, we assume that ε is bounded from below by -1, so that output is always nonnegative; and bounded from above by one.

Note, however, that an ε-firm may chose not to invest at all [that is, to stick to its existing stock of capital (K_H^O)] and avoid the lumpy setup cost C_H. Naturally, a firm with a low ε may not find it worthwhile to incur the setup cost C_H. In this case, its present value is:

$$V^{-}(A_H, K_H^0, \varepsilon, w_H) = \max_L \left\{ \frac{A_H F(K_H^0, L)(1+\varepsilon) - w_H L + K_H^0}{1+r} \right\}. \tag{4.4}$$

The labor demand of such a firm, denoted by $L^{-}(A_H, K_H^O, \varepsilon, w_H)$, is defined by:

$$A_H F_L(K_H^O, L)(1+\varepsilon) = w_H. \tag{4.5}$$

It is straightforward to show that $(\partial V^{+}/\partial\varepsilon) - (\partial V^{-}/\partial\varepsilon) > 0$ (see Appendix 4A.1). Therefore, there exists a cutoff level of ε, denoted by ε_0, such that an ε-firm will make a new investment, if $\varepsilon > \varepsilon_0$. This cutoff level of ε depends on A_H, C_H, K_H^0, and w_H. We write the cutoff ε as $\varepsilon_0(A_H, C_H, K_H^0, w_H)$. It is defined implicitly by:

$$V^{+}(A_H, K_H^0, \varepsilon_0, w_H) - C_H = V^{-}(A_H, K_H^0, \varepsilon_0, w_H). \tag{4.6}$$

That is, the cutoff productivity level is the level at which the firm is just indifferent between making a new investment and incuring the setup cost or sticking to its existing capital stock.

The wage rate w_H is determined in equilibrium by a clearance in the labor market. We assume that labor is confined within national borders. Denoting the country's endowment of labor in effective units by $\tilde{L}_H^0$, we have the following labor market-clearing equation:

$$\begin{aligned} N_H &\int_{-1}^{\varepsilon_0(A_H,C_H,K_H^0,w_H)} L^-(A_H, K_H^0, \varepsilon, w_H) g(\varepsilon) d\varepsilon + \\ N_H &\int_{\varepsilon_0(A_H,C_H,K_H^0,w_H)}^{1} L^+(A_H, \varepsilon, w_H) g(\varepsilon) d\varepsilon = \tilde{L}_H^0. \end{aligned} \tag{4.7}$$

Dividing the latter equation through by N_H, yields:

$$\begin{aligned} &\int_{-1}^{\varepsilon_0(A_H,C_H,K_H^0,w_H)} L^-(A_H, K_H^0, \varepsilon, w_H) g(\varepsilon) d\varepsilon \\ + &\int_{\varepsilon_0(A_H,C_H,K_H^0,w_H)}^{1} L^+(A_H, \varepsilon, w_H) g(\varepsilon) d\varepsilon = L_H^0, \end{aligned} \tag{4.8}$$

where $L_H^0 \equiv \tilde{L}_H^0 / N_H$ is the effective labor per firm.

Note that no similar market-clearing equation is specified for capital, because we assume that capital is freely mobile internationally, and its rate of return is equalized internationally. The same description with the subscript "S" replacing "H" holds for the source country.

Note that differences in labor abundance between the two countries are manifested in the wage differences. To see this, suppose that the two countries are identical, except that effective labor per firm is more abundant in the host country than in the source country, that is: $L_H^0 > L_S^0$. Note also that the number of firms in the economy is also a measure of the abundance of entrepreneurship. Thus, the abundance (respectively, scarcity) of labor is also relative to the scarcity (respectively, abundance) of entrepreneurship. If wages were equal in the two countries, then effective labor demand per firm were equal and the market-clearing condition [equation (4.7)] could not hold for both countries. Because of the diminishing marginal product of labor, it follows that the wage in the relatively labor-abundant country is lower than in the relatively labor-scarce country, that is: $w_H < w_S$[2]. Thus, equal returns to capital (through capital mobility coexist with unequal wages[3]. This provide a complementary reconciliation of the Lucas (1990) paradox of why capital does not flow from rich to poor countries.[4]

4.3 M&A and Greenfield Investments

One may think of FDI as the investment of source-country entrepreneurs in the acquisition of host-country firms. Suppose that the source-country entrepreneurs are endowed with some "intangible" capital, or known-how, stemming from their specialization or expertise in the industry at hand. As before, we model this comparative advantage by assuming that the setup cost of investment in the host country, when investment is done by source-country entrepreneurs (FDI investors) is only C_H^*, which is below C_H (the setup cost of investment when carried out by the host country direct investors). As before, this cost advantage implies that the foreign investors can bid up the direct investors of the host country in the purchase of the investing firms in the host country. Each such firm [that is, each firm whose ε is above $\varepsilon_0(A_H, C_H^*, K_H^0, w_H)$] is purchased at its market value, which is $V^+(A_H, K_H^0, \varepsilon, w_H) - C_H^*$. This essentially assumes that competition among the foreign direct investors shift all the gains from their lower setup cost to the host-country original owners of the firm. The new owners also invest an amount $K^+(A_H, \varepsilon, w_H) - K_H^0$ in the firm. Thus, the amount of foreign direct investment made in an $\varepsilon-$firm (where $\varepsilon > \varepsilon_0$) is:

$$FDI(A_H, C_H^*, K_H^0, \varepsilon, w_H) = V^+(A_H, K_H^0, \varepsilon, w_H) - C_H^* + K^+(A_H, \varepsilon, w_H) - K_H^0. \tag{4.9}$$

This specification assumes that the setup cost C_H^* is incurred in the source country and does not therefore constitute a part of the definition of FDI. It conforms with the notion that C_H^* represents, for instance, R&D of a new product line carried out by the parent firm in the source country.[5] Aggregate FDI is given by

$$\overline{FDI}(A_H, C_H^*, K_H^0, w_H) = \int_{\varepsilon_0(A_H, C_H^*, K_H^0, w_H)}^{1} FDI(A_H, C_H^*, K_H^0, \varepsilon, w_H) g(\varepsilon) d\varepsilon. \tag{4.10}$$

Suppose first that w_H is fixed. Note that it follows from equation (4.1) that $\partial V^+/\partial K_H^0 = 1$, by the envelope theorem. Therefore, $\partial(FDI)/\partial K_H^0 = 0$, by equation (4.9). Thus, the amount of FDI in a firm whose ε is above ε_0 does not depend on the initial capital stock, K_H^0: an increase of \$ 1 in the initial stock of capital of such a firm increases the value of the firm by 1\$, but decreases the required new investment by the same amount, so that FDI does not change[6]. However, the aggregate amount of FDI diminishes, when the initial stock of capital (K_H^0) rises. This is because fewer firms will make new investment and be purchased by foreign direct investors, that is, the cutoff ε_0 rises, when K_H^0 rises. To see this, differentiate equation (4.10) with respect to K_H^0 to get :

$$\frac{\partial \overline{FDI}}{\partial K_H^0} = -FDI(A_H, C_H^*, K_H^0, \varepsilon_0, w_H)g(\varepsilon_0)\frac{\partial \varepsilon_0}{\partial K_H^0} < 0, \tag{4.11}$$

because $\frac{\partial \varepsilon_0}{\partial K_H^0} > 0$ (see Appendix 4A.1).

Similarly, it follows from equation (4.10) that:

$$\frac{\partial \overline{FDI}}{\partial C_H^*} = (-1)\left[1 - G(\varepsilon_0)\right] - FDI(A_H, C_H^*, K_H^0, \varepsilon_0, w_H)g(\varepsilon_0)\frac{\partial \varepsilon_0}{\partial C_H^*}. \tag{4.12}$$

Note that $\frac{\partial \varepsilon_0}{\partial C_H^*} > 0$; see Appendix 4A.1. Hence, it follows that $\frac{\partial \overline{FDI}}{\partial C_H^*} < 0$.

It also follows from equation (4.10) that

$$\begin{aligned}\frac{\partial \overline{FDI}}{\partial A_H} = & \left[\frac{F(K^+, L^+)}{1+r} + \frac{\partial K^+}{\partial A_H}\right][1 - G(\varepsilon_0)] \\ & -FDI(A_H, C_H^*, K_H^0, \varepsilon_0, w_H) g(\varepsilon_0) \frac{\partial \varepsilon_0}{\partial A_H} > 0,\end{aligned} \tag{4.13}$$

because $\frac{\partial K^+}{\partial A_H} > 0$ and $\frac{\partial \varepsilon_0}{\partial A_H} < 0$; see Appendix 4A.1.

Thus, a lower level of the initial stock of capital in the host country attracts more foreign direct investment. Similarly, a lower level of the setup cost of investment in the host country for the FDI investors from the source country promotes more FDI[7]. Also, a higher country-specific productivity factor in the host country promotes more FDI. These conclusions were drawn under the assumption that the wage (w_H) in the host country is fixed. When it is not fixed, then lower K_H^0 and/or C_H^* attract more FDI and push the wage rate upward, thereby mitigating the initial increase in FDI, but not eliminating it altogether.

Observe that FDI flows constitute only a fraction of the international capital transactions between the host and source countries. In a globalized world capital market, where the world rate of interest is given to our pair of countries, domestic saving and domestic investment are not equal to each other, and FDI is not equal to either saving or investment.

So far, FDI took the form of mergers or acquisitions of existing firms. Consider now the possibility of establishing a new firm (that is, a greenfield FDI, where $K^0 = 0$). Suppose that the newcomer entrepreneur does not know in advance the productivity factor (ε) of the potential firm. The entrepreneur therefore takes G(.) as the cumulative probability distribution of the productivity factor of the new firm. However, we assume that ε is revealed to the entrepreneur, before she decides whether or not to make new investment. The expected value of the new firm is therefore:

$$V(A,C,w) = \int_{-1}^{1} Max\left\{V^{+}(A,\varepsilon,w) - C, 0\right\} g(\varepsilon)d\varepsilon. \tag{4.14}$$

Note that if K^0 is equal to zero, only the firms with an ε high enough to justify a greenfield investment have a positive value. This explains equation (4.14).

Now suppose that greenfield entrepreneurship is in limited supply and capacity. An entrepreneur in a source country (and there is a limited number of them) may have to decide whether to establish a new firm at home (the source country) or abroad (the host country), but not in both. Her decision is naturally determined by where $V(\cdot)$, as defined in equation (4.14), is higher. She will invest in the host country rather than in the source country if, and only if,

$$V(A_H, w_H) - C_H^{*} > V(A_S, w_S) - C_S^{*}. \tag{4.15}$$

Naturally, the lower wage rate in the host country works as a pull factor for that country, that is, it works in the direction of satisfying condition (4.15). Thus, the lower wage rate in the host country attracts greenfield FDI. On the other hand, if the total factor productivity in the source country (namely, A_S) is higher than its counterpart in the host country (namely, A_H), this discourages FDI. Assuming that the wage differential dominates the total factor productivity differential, the host country attracts greenfield FDI from the source country.

Assuming that newcomer entrepreneurs evolve gradually over time and that technology spillover equates total factor productivity, eventually this process may end up with full factor price equalization. Naturally, the capital-labor ratios and $L \equiv \widetilde{L}/N$ are equalized in such long-run steady state. This all happens even though labor is not internationally mobile. The establishment of new firms in the global economy may be an engine for FDI flows by multinationals.

Our two-country model, which generates capital flows from the source to the host country, can be extended in a straightforward manner to explain two-way FDI flows. By assuming more than one industry, the extension allows two-way flows between two rich countries, when each country has a setup cost advantage in a different industry.

4.4 Conclusion

The existence of setup costs of FDI presents the investors with a two-fold decision: whether or where to invest at all, and, if so, how much to invest. Consider some source-host country pair. A comparative advantage of parent companies in the source country with respect to R&D of a new product line or some other setup cost comparative advantage due to specialization, etc. enhance the likelihood of embarking on a new FDI and raises its volume. This is relevant mainly for FDI in form of M&A.

A greenfield investment, which takes the form of establishing a new firm rather than just acquiring and expanding the capital stock of an existing firm, works to close the wage gap between the source and the host countries. This occurs with the evolving over time of new supply of entrepreneurship. The new entrepreneurs must choose whether to keep their talent at home (the source country) or utilize it abroad (at the host country) in establishing a greenfield FDI. The limited supply of entrepreneurs gives rise to a discrete choice about the location of new firms.

We account for these forms of FDI flows in a general-equilibrium context of endogenous wages. This framework gives also rise to a coexistence of equal rates of return to capital across countries, due to free capital mobility, and wage gaps as in Lucas (1990).

Appendix 4A.1: Some Comparable Statics Derivations

(i) The Effect of the Firm's Productivity on its Market Value

An ε-firm will choose to incur the setup cost of a new investment, if and only if

$$V^+(A_H, K_H^0, \varepsilon, w_H) - C_H > V^-(A_H, K_H^0, \varepsilon, w_H).$$

It follows from the envelope theorem that

$$\begin{aligned} \frac{\partial V^+}{\partial \varepsilon} - \frac{\partial V^-}{\partial \varepsilon} = & \frac{A_H}{1+r}\{F[K^+(A_H, \varepsilon, w_H), L^+(A_H, \varepsilon, w_H)] \\ & -F[K_H^0, L^-(A_H, K_H^0, \varepsilon, w_H)]\}. \end{aligned} \tag{4A.1}$$

We assume that $K^+(A_H, \varepsilon, w_H) > K_H^0$[8]. Assuming further that capital and labor are complementary in production (that is, $F_{KL} > 0$), it follows from equation (4.5) that $L^+(A_H, \varepsilon, w_H) > L^-(A_H, K_H^0, \varepsilon, w_H)$. Thus,

$$F[K^+(A_H, \varepsilon, w_H), L^+(A_H, \varepsilon, w_H)] > F[K_H^0, L^-(A_H, K_H^0, \varepsilon, w_H)]. \tag{4A.2}$$

Hence, $\frac{\partial V^+}{\partial \varepsilon} - \frac{\partial V^-}{\partial \varepsilon} > 0$.

(ii) The Effect of the Initial Stock of Capital and the Setup Cost on the Cutoff Productivity Level

It follows from equation (4.6) that

$$\frac{\partial \varepsilon_0}{\partial K_H^0} = \frac{\frac{\partial V^-}{\partial K_H^0} - \frac{\partial V^+}{\partial K_H^0}}{\frac{\partial V^+}{\partial \varepsilon} - \frac{\partial V^-}{\partial \varepsilon}} = \frac{\frac{1}{1+r} + A_H F_K(K_H^0, L^-)(1+\varepsilon_0) - 1}{\frac{A_H F_K(K^+, L^+)}{1+r} - \frac{A_H F_K(K_H^0, L^-)}{1+r}}. \tag{4A.3}$$

Note that $A_H F_K(K_H^0, L^-)(1+\varepsilon_0) > r$, as the firm retains its K_H^0 rather than invest in order to equalize the marginal product of capital to the rate of interest. Note also that the denominator in equation (4A.3) is positive; see condition (4A.2). Thus, $\frac{\partial \varepsilon_0}{\partial K_H^0} > 0$.

Similarly, it follows from equation (4.6) that

$$\frac{\partial \varepsilon_0}{\partial C_H^*} = \frac{-1}{\frac{\partial V^+}{\partial \varepsilon} - \frac{\partial V^-}{\partial \varepsilon}} < 0. \tag{4A.4}$$

(*iii*) The Effect of the Host-Country Specific Productivity Level on FDI

By total differentiation of equations (4.2) and (4.3) with respect to A_H, we find that:

$$\frac{\partial K^+}{\partial A_H} = \frac{-F_K F_{LL} + F_{KL} F_L}{A_H (F_{KK} F_{LL} - F_{KL}^2)} > 0, \tag{4A.5}$$

because the denominator and -F_{LL} are positive by the strict concavity of F, and F_{KL} is positive by the factor-complementary assumption.

It similarly follows from equation (4.6) that

$$sign\left(\frac{\partial \varepsilon_0}{\partial A_H}\right) = sign\left(\frac{\partial V^-}{\partial A_H} - \frac{\partial V^+}{\partial A_H}\right) = sign\left[F(K_H^0, L^-) - F(K^+, L^+)\right], \tag{4A.6}$$

which is negative. Thus, $\frac{\partial \varepsilon_0}{\partial A_H} < 0$.

Appendix 4A.2: Reconciliation of the International Flow Paradox according to Lucas

Lucas (1990) employs a standard concave *constant-returns-to-scale* production function:

$$Y = AF(K, L), \tag{4A.7}$$

where Y is output, K is capital and L is *effective* labor. The latter is used in order to allow for differences in the human capital content of labor between developed and developing countries. The parameter A is a productivity index which may reflect the average level of human capital in the country, *external* to the firm. In addition, A may reflect the stock of public capital (roads and other infrastructure) that is *external* to the firm. In per effective-labor terms, we have:

$$y \equiv Y/L = AF(K/L, 1) \equiv Af(k). \tag{4A.8}$$

The return to capital (r) is:

$$r = Af'(k), \tag{4A.9}$$

whereas the wage per effective unit of labor (w) is:

$$w = A[f(k) - kf'(k)]. \tag{4A.10}$$

Let a variable with an asterisk (*) stand for a rich (developed) country and a variable without an asterisk for a poor (developing) country. The function f is common to all countries. Initially, before there is any capital mobility, the returns to capital is higher in the rich country than in the poor country: $r_0^* < r_0$. But when capital can freely move from the rich to the poor country, then the rates of return are equalized, so that:

$$r^* = A^* f^{*\prime}(k^*) = Af'(k) = r. \tag{4A.11}$$

Lucas essentially assumes that $A^* > A$ (because of a human-capital externality). Hence, it follows from equation (4A.11) that $k^* > k$ (because of a diminishing marginal product of capital). Therefore, employing equation (4A.10), it follows that $w^* > w$.

That is, at equilibrium, workers can earn higher wages (per effective labor) in the rich country than in the poor country, and administrative means (migration quotas) are employed to impede the flow of labor from poor to rich countries. Yet, there is no pressure on capital to flow in the opposite direction, because rates-of-return to capital are equalized.

Notes

[1]With constant returns to scale, the fixed cost will entail diminishing average cost curve, in which case perfect competition cannot be sustained. Were we to assume that entry is free, one could have constant returns to scale at the industry level.

[2]The equilibrium wage gap implies that the host country employs more workers per firm than the source country. Thus, even though the productivity distribution across firms is assumed equal, the source country is effectively more productive in equilibrium.

[3]See also Amiti (2005) who studies the effect of agglomeration on cross-regional wage differences. See also Melitz (2003) for the role of fixed costs in intra-industry reallocations in reaction to industry-specific productivity shocks.

[4]See appendix 4A.2 for a brief description of the paradox posed by Lucas and his reconciliation of it.

[5]Whether we interpret C_H^* as being carried out in the source country or in the host country, and accordingly whether we exclude it or include it in the definition of FDI does not alter our qualitative results.

[6]This is because, in the absence of a marginal adjustment cost of investment, the marginal Tobin's q is identically equal to one.

[7]Interestingly, a decline in the setup cost affects the average recorded productivity, because the cutoff ε changes. The new spectrum of investing

firms is accordingly adjusted. A similar endogenous-productivity mechanism features in Ghironi and Melitz (2004).

[8]This will be true if the support of G is sufficiently bounded away from -1 or if $K_H^0 = 0$.

Chapter 5

Country-Specific Aggregate Shocks: Representative Firm

5.1 Introduction

So far our setup allowed heterogeneity across firms and our analysis has gone to the details of the firm level. Thus, an empirical investigation based on this analysis would have required firm-level data. Firm-level data are typically available only for a small subset of countries and on a cross-section basis. On the other hand, there is a fairly rich dataset on aggregate bilateral flows. These data enable us to study how cross-country differences in institutions, macro-policies, productivities, etc. affect bilateral FDI flows. But these aggregate data do not allow us to infer whether a reduction in aggregate FDI was caused by each firm reducing its investment or by some low-productivity firms cutting their investment altogether (because of the fixed costs). Therefore, in this chapter, which serves also as abridge to the data, we assume homogeneous firms in each country. In chapter eight we employ a panel data on both OECD and non-OECD countries, across which productivity and setup cost may vary considerably, in order to empirically analyze the determinants of FDI flows.

We concentrate in this chapter on productivity differences across countries as a key factor that drives FDI flows. A high level of productivity in the potential source country versus a low level of productivity in the potential host country would put adverse pressure on FDI flows. We point out that when we take into account threshold barriers, which are typical for FDI as we explain in this book, then this simple prescription needs some substantial modifications. We show that the productivity levels in a pair of source-host countries manifest themselves differently in the two-fold - the selection and flow - FDI decisions. (Recall that with threshold barriers, a firm must decide whether to invest at all and not only how much to invest.) Furthermore, the effects of the productivity shocks depend also on whether FDI is in the form of M&A or in the form of greenfield investment.

5.2 Country-Specific Productivity Shocks

As we focus on the general productivity level of a country, we abstract from the heterogeneity across firms within a country, which was assumed before.

We therefore assume that all firms have the same productivity factor (ε) which is embedded henceforth in the aggregate productivity factor A[1].

Consider again a representative industry in a given host country (H) in a world of free capital mobility, which fixes the world rate of interest, denoted by r. As before, there is a single good which serves both for consumption and investment.

As our focus here is on the country-specific productivity level, we would like to reckon with the possibility that a productivity change affects wages. If the setup cost of a new FDI is in part in domestic (host-country) inputs, we have to take into account the indirect effect of a productivity change on the setup cost.[2] Therefore, we assume that the setup cost is of the form

$$C_H = C_{SH} + w_H L_H^C, \tag{5.1}$$

where C_{SH} is a cost incurred in the source country and L_H^C is a fixed input of domestic labor.

Consider a representative firm which does invest in the first period an amount $I = K - K_H^0$ in order to augment its stock of capital to K. Its present value becomes $V^+(A_H, w_H) - C_H$, where

$$V^+(A_H, w_H) = \max_{(K,L)} \left\{ \frac{A_H F(K,L) - w_H L + K}{1+r} - (K - K_H^0) \right\}. \tag{5.2}$$

(We again assume for notational simplicity that the rate of physical depreciation, δ, is equal to zero.)

The demand of such a firm for K and L are denoted by $K^{+}(A_H, w_H)$ and $L^{+}(A_H, w_H)$, respectively. They are defined by the marginal productivity conditions:

$$A_H F_K(K, L) = r, \tag{5.3}$$

and

$$A_H F_L(K, L) = w_H. \tag{5.4}$$

Note again that the representative firm may choose not to invest at all (that is, to stick to its existing stock of capital K_H^0) and avoid the lumpy setup cost C_H. In this case its present value is

$$V^{-}(A_H, K_H^0, w_H) = \max_L \left\{ \frac{A_H F(K_H^0, L) - w_H L + K_H^0}{1 + r} \right\}, \tag{5.5}$$

and its labor demand, denoted by $L^{-}(A_H, K_H^0, w_H)$, is given by

$$A_H F_L(K_H^0, L) = w_H. \tag{5.6}$$

The firm will make a new investment if, and only if,

$$V^{+}(A_H, w_H) - C_H \geq V^{-}(A_H, K_H^0, w_H). \tag{5.7}$$

That is, the firm makes the amount of investment that is called for by the marginal productivity conditions, (5.3) and (5.4), if and only if, some global selection condition [(5.7)], is met.

As before, we assume that labor is confined within national borders. Denoting the country's endowment of labor by L_H^0, we have the following labor market clearing equation:

$$\left.\begin{array}{ll} L_H^C + L^+(A_H, w_H) = L_H^0 & \text{if } V^+(A_H, w_H) - C_H \geq V^-(A_H, K_H^0, w_H) \\ L^-(A_H, w_H) = L_H^0 & \text{if } V^+(A_H, w_H) - C_H < V^-(A_H, K_H^0, w_H) \end{array}\right\} \tag{5.8}$$

This market clearing equation determines the wage rate in the host country, as a function $w_H(A_H)$ of the host-country productivity factor.

5.3 The Conflicting Effects of the Source- and Host-Country Productivity Shocks

We now turn to discuss determinants FDI flows from source country S to host country H. As before, we treat as FDI the investment of source-country entrepreneurs in the mergers and/or acquisitions of host-country firms. The source-country entrepreneurs are endowed with some "intangible" capital, or know-how, stemming from their specialization or expertise in the industry at hand. This comparative advantage is modelled by assuming that the lumpy setup cost of investment in the host country, when investment is done by the source country entrepreneurs (FDI investors), is below the lumpy setup cost of investment, when carried out by the host country direct investors. This means that the foreign direct investors can bid up the direct investors of the host country in the purchase of the investing firms in the host country. The representative firm is purchased at its value which is $V^{+}[A_H, w_H(A_H)] - C_H$. As before, this essentially assumes that competition among the foreign direct investors pushes the price of the acquired firm to its maximized value. Thus, the FDI investors shift all the gains from their lower setup cost to the host-country original owners of the firm. The new owners also invest an amount $K^{+}[A_H, w_H(A_H)]$ to expand the capital stock of the acquired the firm. On the other hand, if the selection condition (5.7) does not hold, then there will be no FDI flows from country S to country H.

Thus, aggregate foreign direct investment is equal to:

$$FDI = \begin{cases} V^{+}[A_H, w_H(A_H)] - C_H + w_H(A_H)L_H^C & \text{if} \\ +K^{+}[A_H, w_H(A_H)] - K_H^0 & V^{+}[A_H, w_H(A_H)] \geqq V^{-}(A_H, K_H^0, w_H) \\ 0 & \text{if} \\ & V^{+}[A_H, w_H(A_H)] < V^{-}(A_H, K_H^0, w_H) \end{cases} \tag{5.9}$$

Note that the price paid by the FDI investors for the representative firm is $V^{+} - C_H$. Therefore, the specification in equation (5.9) essentially assumes that the domestic component of the fixed setup cost, $w_H(A_H)L_H^C$, constitutes a part of the national accounting definition of FDI. It also assumes that the capital investment $V^+ - C_H + K^+ - K^0$ is financed from abroad. As the source of financing is indeterminant in our simple model, our definition of FDI in the national accounting is irrelevant in our model.

The model thus suggests that if the productivity factor (A_H) is sufficiently high, and/or the wage rate (w_H) is sufficiently low, and/or the setup cost $(C_{SH} + w_H L_H^C)$ is sufficiently low, then FDI flows from country S to country H are positive. Otherwise, the flow of FDI from country S to country H is zero.

As a preamble to our empirical analysis in the next part, we emphasize that the model's special feature is the two-fold mechanism of FDI decisions. First, one decides how much to invest abroad, while ignoring the fixed setup cost. Second, a decision is made whether to invest at all, taking into account this cost. The hallmark of our empirical approach to follow is based on the two equations (conditions) that govern these decisions.

First, the FDI flows from country S to country H (denoted by FDI_N) is governed by a "notional" flow (or gravity) equation:

$$FDI_N = V^+[A_H, w_H(A_H)] - C_H + w_H(A_H)L_H^C + K^+[A_H, w_H(A_H)] - K_H^0. \tag{5.10}$$

That is, the quantity of investment is governed by the marginal productivity conditions (5.3) and (5.4). Note that the representative firm, if forced to invest in circumstances when it does not pay for it to invest, would have invested according to equation (5.10) that is generated by the first-order conditions (5.3) and (5.4).

Second, the question whether FDI flows from country S to country H are at all positive is governed by a "selection" equation (condition):

$$V^{+}[A_H, w_H(A_H)] - C_H - V^{-}[A_H, K_H^0, w_H(A_H)] \geqq 0. \tag{5.11}$$

Consider now the effect of an increase in the host country's productivity factor, A_H, on the flow equation (governing the quantity of notional FDI) and the selection equation (governing the decision whether to invest at all).

As before, suppose initially that the wage rate in the host country (w_H) is fixed [that is, ignore the labor market clearing condition in equation (5.8)]. An increase in A_H raises the quantity of new investment (K^+), if the investment is at all carried out, the acquisition price ($V^+ - C_H$) that FDI investors pay, the amount of notional FDI, and the demand for labor in the host country[3]. However, when wages are not fixed [but are rather determined by the labor-market clearing equation (5.8)], then the increase in the demand for labor raises the wage rate (w_H) in the host country (and the fixed setup cost $w_H L_H^C$), thereby countering the above effects on K^+, $V^+ - C_H$, and the notional FDI. With a unique equilibrium, the initial effects of the increase in A_H are likely to dominate the subsequent counter effects of the rise in w_H, so that the notional FDI still rises[4]. Thus, an increase in the host country's productivity factor (A_H) raises the volume of the notional FDI flows from country S to country H that is governed by the flow equation.

Next, consider the effect of a productivity shock on the selection equation. A rise in A_H increases the value of the domestic component of the setup cost, $w_H(A_H)L_H^C$. This effect by itself weakens the advantage of carrying out positive FDI flows from country S to country H at all. In other words, the gap between $V^+ - C_H$ and V^- in the selection equation may narrow down.

To sum up, a positive productivity shock in the host country raises the observed notional FDI flows in the flow equation and, at the same time, may lower the likelihood of observing positive FDI flows at all.

We demonstrate the possibility of conflicting effects of an increase in the host country productivity level (A_H) on the flow and selection equations with variable wage via some simulations. We employ a production function with constant elasticity of substitution (CES) and diminishing returns to scale:

$$F(K, L) = [\alpha K^\sigma + (1-\alpha)L^\sigma]^{\frac{\gamma}{\sigma}}$$

where $\gamma = 0.95$ (ensuring diminishing returns) and $\sigma = -0.05$ (implying an elasticity of substitution of 0.95).[5] Figure 5.1a depicts the notional FDI flows [FDI_N in equation (5.10)] and the actual FDI flows [FDI in equation (5.9)] as a function of the host country productivity level (A_H). Figure 5.1b depicts the left hand side of the selection condition (5.11) as a function of A_H. As we can see from Figure 5.1a, FDI_N is indeed increasing in A_H. However, Figure 5.1b shows that as A_H increases the left-hand side of the selection condition can drop below zero, thereby eliminating the actual FDI even though the notional FDI still rises.

(Figure 5.1 A&B about here)

Next, consider the effect of the productivity level in the source country (A_S) on the flow and selection equations. Clearly, A_S does not appear in the the flow equation (5.10), so that it does not affect the flow of notional FDI. Neither does A_S affect the selection equation in its current form of equation (5.11). However, A_S comes into play in the selection decision, when we consider again the limited supply of entrepreneurs in the source country. This consideration is particularly relevant for greenfield FDI. A source-country representative entrepreneur then faces a discrete choice of whether to invest either at home or abroad, but not in both. In this case, in order for her to make greenfield FDI, it no longer suffices that $V^+ - C_H$ exceeds V^-; rather $V^+ - C_H$ must also exceed the value of alternative direct investment at home. The latter naturally depends positively on the source-country productivity level, A_S, and we denote it by $B(A_S)$. That is, the selection condition is:

$$V^+[A_H, w_H(A_H)] - C_H > Max\left\{V^-[A_H, K_H^0, w_H(A_H)], B(A_S)\right\}. \quad (5.12)$$

Thus, the source-country productivity level affects negatively the selection decision. But it has no bearing on the flow decision.

5.4 Conclusion

We analyze some determinants of FDI in the presence of setup costs, which give rise to two channels of influence on FDI. In particular, we studied the role of the source and host country productivity levels. A host-country positive productivity shock raises the volume of notional FDI in the flow equation, but may lower the likelihood of selecting positive FDI flows at all. A source-country productivity shock has a negative effect on selecting positive greenfield FDI, but has no bearing on their flows. This chapter provides an analytical underpinning for the econometric investigation to follow.

Notes

[1]For notational simplicity, we also set the number of firm (N) equal one.

[2]This cost may be affected by input prices in other industries. Feliciano and Lipsey (2000) find that a foreign presence in manufacturing in the U.S. raises wages in non-manufacturing establishments.

[3]The proofs of these claims follows in a straightforward manner from the preceding chapter.

[4]However, with fixed setup cost the equilibrium need not to be unique, and an increase in A^H may, somewhat counter-intuitively, reduce FDI, possibly even to zero. For a similar phenomenon, see Razin, Sadka and Coury (2003).

[5]The other parameter values are $\alpha = 0.8$, $C_{SH} = 0$, $L_H^0 = 1$, $K_H^0 = 0.01$, $L_H^C = 0.15$, $r = 0.2$.

Part II

The Econometric Approach

Chapter 6

Overview of the Econometric Equations

6.1 Introduction

In this chapter we present some basic elements of the econometric approach adopted in the empirical investigation of the theoretical implications of our analysis of the bilateral FDI flows in the preceding part. Recall that a crucial feature of the formation of FDI that we emphasized in the first part is the two-fold nature of FDI decisions. There is a decision to make concerning the question whether to invent at all - captured by a "threshold" selection equation; and concerning how much to invest - captured by the flow or gravity equation.

6.2 The Heckman Selection Model

The two-fold nature of FDI decision gives rise to many cases of zero actual FDI flows. With n countries in a sample, there are potentially n(n-1) pairs of source-host (s,h) countries. In fact, the actual number of (s,h) pairs with observed flows is typically much smaller. Therefore, the selection of the actual number of (s,h) pairs, which is naturally *endogenous*, cannot be ignored; that is, this selection cannot be taken as *exogenous*, as has been often a standard practice in gravity models in the literature. This feature of FDI decisions lends itself naturally to the application of the Heckman selection model (1974, 1979).[1] This selection-bias method is adopted to jointly estimate the likelihood of surpassing a certain threshold (the selection equation) and the magnitude of the FDI flow, provided that the threshold is indeed surpassed (the flow equation).

Specify the flow equation [such as equation (5.10)] as

$$Y_{ijt} = \begin{cases} Y^*_{ijt} = X_{ijt}\beta + u_{ijt} & \text{if } \pi^*_{ijt} \geqq 0 \\ 0 & \text{if } \pi^*_{ijt} < 0 \end{cases}, \tag{6.1}$$

where Y^*_{ijt} is a latent variable denoting the flow of notional FDI from source country i to host country j in period t; X_{ijt} is a vector of explanatory variables; β is a coefficient vector; u_{ijt} is an error term; and Y_{ijt} is the actual flow of FDI. Note that Y^*_{ijt} can take both positive and negative values. Note also that the actual flow of FDI, Y_{ijt}, is zero not only when the notional flow, Y^*_{ijt}, is negative; Y_{ijt} may be zero even when Y^*_{ijt} is positive, but does not provide enough profit to surpass the threshold. π^*_{ijt} is the substance of the selection equation [see, for instance, condition (5.11)], and is specified by

$$\pi^*_{ijt} \equiv \pi^{*\prime}_{ijt}/\sigma_{\pi^{*\prime}} = (W_{ijt}\gamma - C_{ijt})/\sigma_{\pi^{*\prime}}, \tag{6.2}$$

where π^*_{ijt} indicates whether an FDI would be made or not (depending whether it is positive or negative); W_{ijt} is a vector of explanatory variables (which may overlap with the explanatory variables of X_{ijt}); C_{ijt} is the fixed cost of setting up new investment; γ is a vector of coefficients; and $\sigma_{\pi^{*\prime}}$ is the standard deviation of $\pi^{*\prime}$. The setup cost C^*_{ijt} is given by

$$C^*_{ijt} = A_{ijt}\delta + v_{ijt}, \tag{6.3}$$

where A_{ijt} is a vector of explanatory variables; δ is a vector of coefficients; and v_{ijt} is an error term. Substituting for C^*_{ijt} in equation (6.2) from equation (6.3), we get:

$$\pi^*_{ijt} = Z_{ijt}\theta + \varepsilon_{ijt}, \tag{6.4}$$

where $Z_{ijt} = (W_{ijt}, A_{ijt})$; $\theta = (\gamma/\sigma_{\pi^{*\prime}}, -\delta/\sigma_{\pi^{*\prime}},)$; and

$$\varepsilon_{ijt} = -v_{ijt}/\sigma_{\pi^{*\prime}}. \tag{6.5}$$

Assuming that u_{ijt} and v_{ijt} are normally distributed with zero means, it follows that $\varepsilon_{ijt} \sim N(0,1)$. The error terms, u_{ijt} and ε_{ijt}, are bivariate normal:

$$\begin{pmatrix} u_{ijt} \\ \varepsilon_{ijt} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{Y^*} & \rho_{Y^*\pi^*}\sigma_{Y^*} \\ \rho_{Y^*\pi^*}\sigma_{Y^*} & 1 \end{pmatrix}\right). \tag{6.6}$$

Define the following indicator function:

$$D_{ijt} = \begin{cases} 1 \text{ if } \pi^*_{ijt} \geqq 0 \\ 0 \text{ otherwise} \end{cases}. \tag{6.7}$$

The latter function indicates whether the threshold is surpassed and an FDI flow is formed or not. Note that π^* itself is a latent variable which is not observed. But we do observe D, that is we do observe whether π^* is positive or not.

The expected value of Y_{ijt}, conditional on the event that there is indeed a positive FDI flow, is given by

$$E\left(Y_{ijt}/D_{ijt}=1\right) = X_{ijt}\beta + E\left(u_{ijt}/D_{ijt}=1\right) \equiv X_{ijt}\beta + \beta_\lambda \lambda_{ijt}, \tag{6.8}$$

where

$$\beta_\lambda = \rho_{Y^*\pi^*}\sigma_{Y^*} \tag{6.9}$$

and

$$\lambda_{ijt} = \frac{\phi(Z_{ijt}\theta)}{\Phi(Z_{ijt}\theta)} \tag{6.10}$$

is the inverse Mills ratio; ϕ and Φ are the density and cumulative unit-normal distribution functions, respectively. Note again that we do not observe π^*_{ijt}, but we do observe D_{ijt}. Because $Prob(D_{ijt} = 1) = Prob(\pi^*_{ijt} \geqq 0) = \Pr ob(\varepsilon_{ijt} \geqq -Z_{ijt}\theta) = \Pr ob(\varepsilon_{ijt} \leqq Z_{ijt}\theta)$, by equation (6.4) and the symmetry of the normal distribution, it follows that

$$Prob(D_{ijt} = 1) = \Phi(Z_{ijt}\theta). \tag{6.11}$$

The maximum likelihood method is then employed to jointly estimate the flow coefficient vector β and the selection coefficient vector θ.

Note that λ_{ijt} depends on X_{ijt}. Therefore, one can see from equation (6.8) that OLS estimates of the coefficient vector β of the flow equation, confined to positive observations of Y_{ijt} (that is, discarding the zero flows), is biased because such estimates include also the effect of X_{ijt} on Y_{ijt} through the term $\beta_\lambda\lambda_{ijt}$. Figure 6.1 explains the intuition for the cause of the bias for the case where $\rho_{Y^*\pi^*} > 0$. Suppose, for instance, that x_{ijt} measures the productivity differential between the ith source country and the potential jth host country, holding all other variables constant. Our theory predicts that the parameter β_x is positive. This is shown by the upward sloping line AB. Note that the slope is an estimate of the "true" marginal effect of x_{ijt} on Y^*_{ijt}. But recall that flows could also be equal to zero, if the setup costs are sufficiently high. A threshold, which is derived from the setup costs, is shown as the curve $TT\prime$ in Figure 6.1. However, if we discard observations with zero actual FDI flows, the remaining sub-sample is no longer random. As equation (6.2) makes clear, the *selection* of country pairs into the sub-sample depends on the vector X_{ijt} (including x_{ijt}). To illustrate, suppose that for high values of x_{ijt} (say, x^H in Figure 6.1), (i, j) pair-wise FDI flows are all positive. That is, for all pairs of countries in the sub-sample the threshold is surpassed and the *observed* average of notional FDI flows for $x_{ijt} = x^H$ is also equal to the conditional population average for FDI flows, point R on line AB. However, this does not hold for low values of x_{ijt} (say, x^L). For these (i, j)-pairs, we observe positive values of $Y_{i,j,t}$ only for a subset of country pairs in the population. Point S is, for instance, excluded from the sub-sample of positive FDI flows. Consequently, for low x_{ijt}'s, we observe only flows between country pairs with low setup costs (namely, with low $v_{i,j,t}$'s). As a result, the observed average of the FDI flows is at point $M\prime$, whereas the "true" average is at point M. As seen in Figure 6.1, the OLD regression line for the sub-sample is therefore the $A\prime B\prime$ line, which underestimates the effect of productivity differentials on bilateral FDI flows.

(Figure 6.1 about here)

If we do not discard the zero FDI flow observations, the OLS estimates of β are still biased, because they are based on observations on Y rather than on Y^*.[2]

6.3 The Tobit Model

As we have seen, our theory suggests that actual FDI flows may be zero even when notional FDI flows are not. In fact, the actual FDI flow variable in a typical sample is zero for a significant fraction of the sample, but is roughly continuously distributed for positive values. In such circumstances, the Tobit model is often employed. The jist of this model is in our case as follows. Suppose that instead of equation (6.1) , we were to have:

$$Y'_{ijt} = \begin{cases} Y^*_{ijt} = X_{ijt}\beta + \mu_{ijt} & \text{if } Y^*_{ijt} \geqq 0 \\ 0 & \text{if } Y^*_{ijt} < 0 \end{cases}, \tag{6.1a}$$

where $\mu_{ijt} \sim N(0, \sigma_Y^2)$. Define now an index function

$$D'_{ijt} = \begin{cases} 1 \text{ if } Y^*_{ijt} \geqq 0 \\ 0 \text{ if } Y^*_{ijt} < 0 \end{cases}. \tag{6.7a}$$

One can show [see, for instance, Woolridge (2003)] that in this case

$$E\left(Y'_{ijt}/D'_{ijt} = 1\right) = X_{ijt}\beta + \beta'_\lambda \lambda'_{ijt}, \tag{6.8a}$$

where

$$\beta'_\lambda = \sigma_{Y^*} \tag{6.9a}$$

and

$$\lambda'_{ijt} = \frac{\phi(X_{ijt}\beta/\sigma_{Y^*})}{\Phi(X_{ijt}\beta/\sigma_{Y^*})} \tag{6.10a}$$

Comparing the Tobit equations (6.8a)-(6.10a) with the Heckman equations (6.8)-(6.10), one can see that the Tobit model is a special case of the Heckman model when Y^*_{ijt} and π^*_{ijt} are fully positively correlated, that is: $\rho_{Y^*\pi^*} = 1$. In this special case, there is a perfect correlation between the event $Y^*_{ijt} \geqq 0$ and the event $\pi^*_{ijt} \geqq 0$ (for any given X_{ijt}), and therefore the two fundamental equations (6.1) and (6.1a) are the same. Thus, if the estimated value of $\rho_{Y^*\pi^*}$ in the Heckman model is significantly below 1, then the Tobit estimates of the coefficient vector β are biased.

6.4 Conclusion

Historically, empirical studies of the flows FDI where cast in a framework of gravity equations. However, fixed setup costs create some thresholds for FDI flows. The latter are therefore positive only when a certain profitability threshold is surpassed. Thus, a typical sample of FDI flows contains a mass (a "big fraction") of zero entries. The empirical literature developed after Tinbergen (1962) has either *omitted* pairs with no FDI flows, or treated reported zero flows as *literally* indicating *zero flows*[3] and employed Tobit estimation methods.[4] In this chapter we overviewed the Heckman Selection and the Tobit methods and illustrated why the former is more appropriate than either the latter or the OLS method in case there are fixed setup costs. Unlike the Tobit and the OLS method, the Heckman method provides unbiased estimates for the determinants of the flows of FDI; it also provides unbiased estimates for the determinants of the endogenous selection of FDI flows, which selection is ignored by both the Tobit and OLS methods.

Notes

[1]See also Keane, Moffit and Runkle (1986) and Kyhazidou (1996) for applications to panel data.

[2]Recently, Silva and Tenreyro (2003) proposed the Poisson pseudo-maximum likelihood estimator to deal with zero values in the bilateral trade models.

[3]Zero flows are sometime also treated as measurement errors. Note that if measurement errors (in the dependent variable) are nor correlated with the explanatory variables, then the estimated parameters are not biased; although they are imprecisely estimated.

[4]Eaton and Tamura (1996) introduced the use of a threshold Tobit estimation to deal with zeros in trade or FDI bilateral flows. The potential for zero FDI activity is also recognized by Brainard (1997) who studies the method multinationals use to serve markets. In robustness checks she provides generalized Tobit estimates that include a probit for the probability of any affiliate sales, combined with the OLS for the import share, where affiliate sales are observed. Carr, Markusen and Maskus (2001) similarly employ Tobit method to study the determinants of goods trade. As we show here, these estimates may be biased, if there are indeed fixed costs that generate thresholds for positive flows of either FDI or goods trade.

Chapter 7

Application to a Base-Line Sample: OLS, Tobit and the Heckman Selection Model

7.1 Introduction

In this chapter we employ data on bilateral FDI flows in a sample of 24 OECD countries over the period from 1981 to 1998 in order to illustrate the application of the Heckman Selection, the Tobit and the OLS methods. This application provides a benchmark empirical study of the determinants of FDI. In the next part we extend this benchmark study to include some major variables which constitute the focus of the theoretical investigation in Part One.

7.2 Data and Variables

Data are drawn from OECD reports (OECD, various years) on a sample of 24 OECD countries, over the period from 1981 to 1998. The FDI data are based on the OECD reports of FDI exports from 17 OECD source countries to 24 OECD countries.

In order to smooth the various variables over business cycles, we employ 3-year averages, so that we have six periods (each consisting of 3 years). The main explanatory variables that we employ are: (1) standard country characteristics such as GDP or GDP per-capita, population size, educational attainment (as measured by average years of schooling), language, financial sound rating (the inverse of financial risk rating), etc.; (2) (s, h) source-host characteristics, such as geographical distance, common language (zero-one variable), (s, h) flows of goods, bilateral telephone traffic per-capita as a proxy for informational distance, etc. Table 7.1 describes the list of the 24 countries in the sample, and indicates for each country whether positive flows are observed in the sample, at least once, as a source or host country (but most source countries do not interact more than with few host countries). Table 7.2 summarizes the data sources.

(Table 7.1 about here)

(Table 7.2 about here)

7.3 Estimation

Table 7.3 and Table 7.4 provide a "first look" at the direction and volume of FDI flows. The first of these two tables describes the frequency of FDI flows between all possible (s,h) pairs. It suggests that source-host differences in GDP per capita look as good predictors of the direction of flows. The frequency of flows is close to one among rich countries, whereas it is very low and often zero among poorer countries. Table 7.4 describes FDI flows as percentages of the host-country GDP. It suggests that source-host differences in GDP per capita are not correlated with the volume of FDI flows for the subset of country pairs with positive flows. For instance, Japan which is the second richest country in the sample received FDI flows from the U.S. amounting to 1.26 percent of Japan's GDP; whereas Spain received FDI from the U.S. amounting to 6.54% of Spain's GDP.[1]

(Table 7.3 about here)

(Table 7.4 about here)

We now turn to the estimation of the determinants of bilateral FDI flows. We consider several potential explanatory variables of the two-fold decisions on FDI flows.[2] We regroup these variables as follows: standard "mass" variables (the source and host population sizes); "distance" variables (physical distance between the source and host countries and whether or not the two countries share a common language); and "economic" variables (source and host GDP per capita, source-host differences in average years of schooling, and source and host financial risk rating). We also control for country and time fixed effects. The dependent variable in all the flow (gravity) equations is the log of the FDI flow, deflated by the unit value of manufactured goods exports.

We employ three alternative econometric procedures. As a benchmark, we ignore the selection equation, and simply estimate the gravity equation twice: (i) by treating all FDI flows in (s, h) pairs with no recorded FDI flows as "zeros" (OLS-zero);[3] (ii) by excluding country pairs with no FDI flows (OLS-D). The rationale for inserting "zeros" in the OLS-zero case is as follows. Generally, when one observes zero FDI flows between a pair of countries, it could be either because the two countries do not wish to have such flows, even in the absence of fixed costs, or because setup costs are prohibitive, or because of measurement errors. But if one assumes that there are no setup costs or measurement errors, (s, h) pairs with zero FDI flows truly indicate zero flows. This is why we assign a negligible value as a common low value for the value of the FDI flows for the zero-flows (s, h) pairs.[4] (All other positive flows have logarithmic value much exceeding zero.) The estimation results for the OLS-zero and OLS-D cases are shown in panel A of Table 7.5.

(Table 7.5 about here)

Next, we continue to assume that there are no fixed costs and that all FDI flows that are below a certain low threshold level ("censor") are due to measurement errors, and employ a Tobit estimator.[5] We present the results in Panel B of Table 7.5, with three censor levels (lowest, 0.0 and 3.00).

Against these two benchmarks, the role played by the unobserved fixed setup costs can be now brought to the limelight, when we employ the Heckman selection method. We jointly estimate the maximum likelihood of the flow (gravity) equation and the selection equation. The Heckman estimation method accommodates both measurement errors and a possible existence of setup costs. Consider a binary variable $D_{i,j,t}$ which is equal to 1 if country i exports positive FDI flows to country j at time t; zero otherwise. Assuming that setup costs are lower if country i already invested in country j in the past, then $D_{i,j,t-k}$ could serve as an instrument in the selection equation (the exclusion restriction). The results are described in Panel C of Table 7.5.

All estimations conform to the notion that the volume of FDI flows is not affected by deviations from long-run averages of GDP per capita in the source and host countries. The coefficient of the GDP per capita variable is not significant in the Heckman selection equation. Turn to the effect of the host country education level, relative to the source country counterpart. Employing Tobit estimation, one may conclude that cross-country educational gaps have a significant effect on the flow of FDI. However, the Heckman method suggests that the cross-country educational gap manifests itself through the selection and has no significant effect on the flow of FDI. To test whether the effect on FDI flows is *non-linear*, we estimate the parameters of interest in the OLS method for different ranges of FDI flows. That is, the OLS regression in the OLS-zero case has different coefficients than the OLS-D regression.

As expected, the common language dummy is positive and significant, and the distance coefficient is negative and significant in all formulations. It is worth noting that only the Heckman model assigns a significant positive role to the host country population (through the selection mechanism). The coefficient of the host-country financial sound rating is significant (and positive) only in Heckman flow equation and the OLS-D case. The source-country financial sound rating has a negative and significant effect on FDI in the Tobit cases and one of the two OLS cases (OLS-zero). However, the Heckman method suggests that this variable works through the selection process rather than having a direct effect on FDI flows. The existence of previous flows of FDI has a significant and positive effect in the selection equation. This may be interpreted as indicating that the existence of FDI flows in the past reduces the fixed cost of setting a new FDI.

Most importantly as a "smoking gun" for the existence of fixed costs in the data, we note that the correlation between the error terms in the flow and the selection equations is negative and significant. This finding, on which we further elaborate in the next section, provides an additional evidence for the relevance of fixed set up costs.

We have a few cases of negative flows in our sample. Negative flows indicate liquidations of previous FDI. In Table 7.6 we use a dummy variable for negative FDI flows(that is, past FDI liquidations) as instruments. It is a reasonable instrument because past FDI liquidations are correlated positively with past FDI flows (liquidations, by definition, are generated from existing stocks), but not apriori correlated with current FDI flows.

(Table 7.6 about here)

The Tobit estimation assigns a positive role (and occasionally significant) to this dummy concerning the flow of FDI. However, the Heckman estimation suggests that the positive effect comes through the selection mechanism. The inclusion of this dummy variable did not significantly affects the results reported in Table 7.5

7.4 Evidence for Fixed Costs

The finding that there is a significant correlation (ρ) between the error terms in the flow and selection equations indicates that the formation of an (s, h) pair of positive FDI, and the size of the FDI flows between this pair of countries are not independent processes. Furthermore, with ρ being negative, this correlation is consistent with the setup costs hypothesis[6]. If some shocks jointly drive marginal productivity of capital and setup costs of FDI, then indeed the error terms in the selection equation may be negatively correlated with the errors terms in the flow equation. For instance, above-average marginal productivity of capital in a host country may yield below-average likelihood of non-zero exports of FDI (because it may yield above-average setup costs); as in Chapter Five.

7.5 Conclusion

In this chapter we illustrate the application of OLS, Tobit and Heckman estimation methods in a sample of bilateral FDI flows among 24 OECD countries over a period ranging from 1981 to 1998. We provide some evidence for the existence of fixed setup costs. In such a case the OLS and Tobit estimates of the determinants of FDI flows are biased. Furthermore, the Heckman method suggests that some of the factors, that are found to be determinants of FDI flows in the OLS and Tobit estimation, in fact influence FDI through the selection mechanism rather than directly through the flows of FDI.

Notes

[1]Note that we aggregate FDI data over 3 years.

[2]As we do not naturally pretend to provide a general theory which explains all potential determinants of FDI flows, we include some other explanatory variables that are found relevant elsewhere in the literature.

[3]More precisely, the log of the FDI flow is set equal to log of the lowest observed flow between any (s, h) country pair in the sample.

[4]We choose this value to be the lowest observed flow between any (s, h) country pair in the sample.

[5]Note that this estimator is appropriate also in the case where the desired FDI flows were actually negative, as in the case where a foreign subsidiary is liquidated, but were recorded as zeros.

[6]For some micro-level evidence for the existence of fixed costs see Caballero and Engel (1999 and 2000).

Part III

Empirical Applications

Chapter 8

Productivity Shocks

8.1 Introduction

Foreign direct investment has become a key channel of international capital flows. In Chapter Five we explain that cross country productivity differences can generate FDI flows. In particular, country-specific productivity shocks within a source-host country pair affect FDI in a variety of possibly conflicting ways. A positive productivity shock in the source country is expected to have a negative effect the selection equation. A positive productivity shock in the host country may also have somewhat surprisingly a negative effect in the selection equation, but a positive effect on the flows of FDI.

The two-fold (selection and flow) FDI decision is generated by the existence of "lumpy" setup costs of new investments that govern the flow of bilateral FDI. The rich and technologically-advanced countries have a comparative advantage in setting up foreign subsidiaries. As this advantage may also be industry-specific, the model is capable of generating two-way rich-rich, and rich-poor FDI flows.

Threshold barriers play also an important role in determining the extent of trade-based foreign direct investment; see, for instance, Zhang and Markusen (1999), Carr, Markusen and Maskus (2001), and Helpman, Melitz and Yeaple (2004). The trade-based literature typically focuses on issues such as the interdependence of FDI and trade in goods and the ensuing industrial structure. For instance, they attempt to explain how a source country can export both FDI and goods to the same host country. The explanation essentially rests on productivity heterogeneity within the source country, and differences in setup costs associated with FDI and export of goods. The trade-based literature on FDI is based on a framework of heterogenous firms, such as in Melitz (2003). Thus, the empirical approach in this literature is geared towards a *firm-level* decisions on exports and FDI in the source country. As explained in Chapter Five, the analytical framework, which leads to the empirical investigation of this chapter, is based on a representative-firm model, as mostly common in macro studies. Our focus is therefore on *aggregate* bilateral FDI. Thus, trade-based empirical applications typically use micro-dataset, whereas we utilize country-wide data set. Note that micro cross-country panel datasets are not available, so that micro-based empirical studies have to be confined typically to a single source or host country and to extremely short time span. In contrast, we employ here data for 62 OECD and non-OECD countries over a large interval of time (1987-2000).

8.2 Data

We consider several potential explanatory variables of the two-fold decisions on FDI flows. As before, these variables include standard "mass" variables (the source and host population sizes); "distance" variables (physical distance between the source and the host countries and whether or not the two countries share a common language); and "economic" variables (source and host real GDP per capita, source-host differences in average years of schooling, and source and host financial risk rating). We also control for country and time fixed effects. The dependent variable in all the flow (gravity) equations is the log of the FDI flows.

As before, the main variables are grouped as follows: (1) standard country characteristics such as real GDP per-capita, population size, educational attainment (as measured by average years of schooling), and financial sound rating (the inverse of financial risk taking); (2) (s,h) source-host characteristics, such as (s,h) FDI flows, geographical distance, common language (zero-one variable). In addition, we focus on a new explanatory variable: (3) productivity, approximated by labor productivity, that is output per worker, as measured by PPP-adjusted real GDP per worker. This variable is at times instrumented by the capital/labor ratio. Table 8.1 summarizes the data sources. Table 8.2 describes the list of the 62 countries in the sample, and indicates for each country, as a source or a host, whether positive bilateral flows are observed in the sample, at least once. Note that most source countries do not interact more than with few host countries. Here we do not smooth the data by taking three-year averages, as in Chapter Seven, but rather employ unfiltered annual data. This enables us to investigate the effects of the explanatory variables over the business cycle.

(Tables 8.1 and 8.2 about here)

Data on FDI flows are drawn from the International Direct Investment dataset (Source OECD), covering the bilateral FDI flows among 62 countries (29 OECD countries and 33 Non-OECD countries) over the period 1987 to 2000. The Source OECD provides data on FDI flows in U.S. dollars, and we deflate them by the U.S. CPI for urban consumers.

8.3 Empirical Evidence

As we explained in Part One, the existence of fixed setup costs generates many observations with zero FDI flows. In our dataset, there are indeed about 62% host-source pairs for which no FDI flows appear. Therefore, as explained in Part Two, the Heckman selection method is adopted to jointly estimate the likelihood of surpassing a threshold generated by the latent fixed setup costs (the selection equation) and the magnitude of the FDI flows, provided that the threshold is indeed surpassed (the flow or gravity equation).

The Source OECD dataset reports FDI flows from OECD countries to OECD and non-OECD countries, as well as FDI flows from non-OECD countries to OECD countries. However, it does not report FDI flows from non-OECD to non-OECD countries. We therefore estimate the model under several alternative assumptions concerning the missing observations on FDI flows from non-OECD to non-OECD countries. Most likely, these missing observations reflect zero FDI flows and this is how we treat them on one alternative. In another alternative, we discard the observations with the missing data on the FDI flows (from non-OECD countries to non-OECD countries). In addition, we present the estimation with and without instrumenting the potentially endogenous output per worker variable, by the capital/labor ratio and years of schooling variables. The estimation results are presented in Tables 8.3-8.4. In each table, Panel A presents the results when the productivity factor is approximated by the output per worker variable. In Panel B, the latter is replaced by instrumental variables: the capital/labor ratio and the years of schooling variables. Table 8.5 presents the results of the estimation of the instrumental equation.

(Table 8.3 about here)

Table 8.3 presents the estimation of the equations for bilateral FDI flow and selection (ignoring missing observations on FDI flows from non-OECD to non-OECD). The effect of the education variable, namely the source-host difference in education levels, on the selection mechanism is significant and negative, but not so on the magnitude of the flows, across different alternative versions of the productivity variable. Host-country financial sound rating is important in most of the specifications in this table, both in the flow and in the selection equations. But we find no evidence for the importance of the source financial sound ratings on bilateral FDI flows, neither in the flow nor in the selection margins. Host GDP per capita is important in the flow equation only. As expected, and consistent with previous "gravity" literature, we find that common language raises, and distance reduces the volume of FDI flows. These two explanatory variables have similar effects in the selection equation. Host population size has a significant coefficient in the flow equation but not in the selection equation. Source population size is insignificant in either equation. The effect of the existence of past FDI relations is positive and significant in the selection equation, as it may help to reduce the setup costs of establishing a new FDI flow, which is in line with our theoretical predictions. We also note that the correlation between the error terms in the flow and the selection equations is negative and significant at the 5%-level. As explained in Chapter Seven, this result may be interpreted as an evidence for the existence of fixed costs, because in their absence the correlation coefficient is one (and the Heckman model reduces to the Tobit model). The past FDI dummy is used as am exclusion restriction variable. The positive coefficient is interpreted as an indication for a lower threshold barrier for pairs of countries that had positive FDI flows in the past.

We turn now to the variables at the focus of the investigation: the productivity factor, as approximated by output per worker. In Panel A the host output per worker has a positive and significant effect in the flow and selection equations. Source-country output per worker has a negative and significant effect only on the selection mechanism, but has no significant effect on the flows of FDI. These results about the effects of the source-country productivity shocks on the flow and selection equations are consistent with the analytical framework developed in Chapter Five. In Panel B, with the productivity variable instrumented by capital per worker and education attainment, the host output per worker affects negatively and significantly the selection mechanism, which is consistent with the model of Chapter Five, but is insignificant in the flow equation (unlike what is predicted by our model). The source instrumented output per worker negatively affects both the flow of FDI and the selection mechanism.

(Table 8.4 about here)

In Table 8.4 we present the estimation of the flow and selection equations, when we treat missing observations on FDI flows from non-OECD to non-OECD countries as "zeros". Results are broadly similar to Table 8.3 and provide evidence consistent with the key hypotheses about the conflicting effects of productivity shocks.

(Table 8.5 about here)

Note that the relationship in the selection equation between the probability (P) of making a new FDI and the explanatory variables (including productivity) is described by the following (non-linear) equation:

$$P(prod_H) = \int_{-\infty}^{\alpha+\theta_P prod_H} (2\pi)^{-1/2} \exp(-y^2/2) dy, \qquad (8.1)$$

where α represents the effect of all the other explanatory variables (held fixed at their sample averages), including country and time fixed effects, and θ_P is the coefficient of $prod_H$ (output per worker in the host country) in the selection equation. Note also that the estimate of θ_P is negative and statistically significant. The marginal effect of $prod_H$ on P is

$$\partial P/\partial prod_H = \theta_P (2\pi)^{-1/2} \exp[-(\alpha + \theta_P prod_H)^2/2] < 0.\text{[1]} \qquad (8.2)$$

Moreover, the expected value of FDI flow is

$$E[FDI] = P(prod_H) \exp(\delta + \beta_P prod_H) + [1 - P(prod_H)] \cdot 0, \qquad (8.3)$$

where δ represents the effect of all the other explanatory variables (held fixed at their sample averages), and β_P is the coefficient of $prod_H$ in the flow equation. Note that we use $\exp(\delta + \beta_P prod_H)$ for the observed FDI flow in that our dependent variable in the flow equation is the log of FDI. Therefore, the marginal effect of $prod_H$ on expected bilateral FDI flows, normalized by $\exp(\delta + \beta_P prod_H)P(prod_H)$, is:

$$\frac{1}{\exp(\delta + \beta_P prod_H)P(prod_H)} \frac{dE\left[FDI\right]}{dprod_H} = \frac{dP(prod_H)}{P(prod_H)dprod_H} + \beta_P. \quad (8.4)$$

The first component, $\frac{dP(prod_H)}{P(prod_H)dprod_H}$, is negative, while the second component, β_P, is positive (see Panel A in Table 8.3). The net effect depends on which component is the dominant force. Figure 8.1 depicts this normalized marginal effect for the U.S. as a source country, with all variables except $prod_H$ fixed at their sample average (based on Panel A in Table 8.3). Figure 8.1 clearly shows that as productivity increases, its marginal impact decreases nonlinearly. Expected FDI flows decline in the level of host country productivity[2]. That is, holding constant US productivity as a source country, the effect of an increase in the host country productivity depends crucially on the initial value of the productivity parameter.

(Figure 8.1 about here)

8.4 Conclusion

In this chapter, we take to the data the prediction of our theory (Chapter Five) about the conflicting effects of productivity changes on FDI flows. A positive productivity shock in the host country typically increases the volume of the desired FDI flows to the host country, through the standard marginal profitability effect (the flow equation). But, at the same time, the same shock may lower the likelihood of making any new FDI flows by the source country (the selection equation), through a total profitability effect, derived from the a general-equilibrium increase in wages and other input prices. Using a sample of 62 OECD and Non-OECD countries, over the period 1987-2000, we provide supporting evidence for the existence of such effects of productivity shocks on bilateral FDI. That is, the empirical findings is that productivity would affect the aggregate flows of FDI in one way and the likelihood of positive FDI flows in another.

Finally, we mention a potential caveat. The predictions from the model with fixed costs are predictions related to investment in capacity, but the FDI flow data captures financial flows associated with such investment. A fraction of FDI investment is often financed in an affiliate's host country, coming from host country sources. To the extent that this fraction is not correlated with the productivity shocks, the empirical predictions though are not biased.

Notes

[1]To complete the picture, note also that $P(prod_H)$ has an inflection point at $prod = -\alpha/\theta_P$.

[2]In our data sample, output per worker in host countries ranges from 2.45 to 86.6.

Chapter 9

Source and Host Corporate Tax Rates

9.1 Introduction

In this chapter we focus the empirical analysis of the determinants of bilateral FDI flows on the effects of taxation. In the era of increased globalization, cross-country differences in taxation may play a major role in the international allocation of investment:

> "European countries have been steadily slashing corporate-tax rates as they vie for foreign investment, potentially adding to pressure on the U.S. for similar cuts as it weighs a tax overhaul. Following the lead of Ireland, which dropped its rates to 12.5% from 24% between 2000 and 2003, one nation after another has

> moved toward flatter, lower corporate rates with fewer loopholes" (Wall Street Journal Europe, January 28-30, 2005).

More recently, Turkey has also jumped on the band of wagon of slushing tax rates in order to attract foreign direct investment:

> "Turkey is to slash corporate and personal tax rates in an effort to attract foreign direct investment. Recep Tayyip Erdogan, prime minister, said yesterday the standard rate of corporate tax would be cut from 30 per cent to 20 per cent to allow Turkey to compete for investment against the new European Union member states. The overall tax burden for all companies would fall to 28 per cent from 37 per cent, while the top rate of personal income tax would fall to 35 per cent. These cuts will attract investment from abroad and greatly increase our competitiveness with neighboring countries and with the European Union." (Financial Times, November 30, 2005).

Indeed, the economic literature has extensively dealt with the effects of taxation on investment, going back to the well-known works of Harbeger (1962) and Hall and Jorgenson (1967). Of particular interest are the effects of international differences in tax rates on foreign direct investment; see, for instance, Auerbach and Hassett (1993), Hines (1999), Desai and Hines (2001), De Mooij and Ederveen (2001), and Devereux and Hubbard

(2003). In Appendix 9A.1 we provide a brief discussion of basic principles of international taxation of capital income.

In this chapter we attempt to provide a new look at the mechanisms through which corporate tax rates influence aggregate FDI flows, in the setup adopted in this book of two-fold investment decisions in the presence of threshold barriers. In this context, the source and host tax rates may have different effects on these two decisions. (the flow and selection equations).

9.2 Source and Host Taxation

Consider for concreteness the case of a parent firm that weighs the development of a new product line. We can think of the fixed setup cost as the outlays of developing the product line. The firm may choose to make the development at home and then carry the production at a subsidiary abroad. This choice may be determined by some "genuine" economic considerations such as source-host differences in labor costs, in infrastructure, in human capital, etc. But it may also be influenced by tax considerations.

In this context of FDI, there arises the issue of double taxation. The income of a foreign affiliate is typically taxed by the host country. If the source country taxes this income too, then the combined (double) tax rate may be very high, and even exceeds 100%[1]. This double taxation is typically relieved at the source country by either exempting foreign-source income altogether or granting tax credits[2]. In the former case, foreign-source income is subject to the tax levied by the host country only. When the source country taxes its resident on their world-wide income and grants full credit for foreign taxes (residence taxation), then in principle the foreign-source income is taxed at the source-country tax rate, so that the host-country tax rate becomes irrelevant for investment decisions in the source country. But, in practice, foreign-source income is far from being taxed at the source country rate. First, there are various reduced tax rates for foreign-source income. Second, foreign-source income is usually taxed only upon repatriation, thereby effectively reducing the present value of the tax. Thus, in practice, the host country tax-rate is much relevant for investment decisions of the parent firm at the source country. The relevance of the host-country tax rate intensifies through transfer pricing.[3]

To highlight the issue of source-host differences in tax rates, suppose that the source country does not tax foreign-source income at all. Denote the fixed cost of development by C. Now, if the host-country tax rate is lower than that of the source country, then the parent firm at the source country attempts to keep this cost at home for tax purposes. The firm may thus charge its subsidiary artificially low royalties for the right to produce the new product. Thus, this cost remains largely deductible in the high-tax source country.[4] As before, denote the (maximized) present value of the cash flows arising from the production and sale of the new product by $V(\tau_H) - C$; it depends (negatively) on the corporate tax rate (τ_H) levied by the host country.

To see this note first that one of the channels through which corporate taxation distorts investment decisions is the depreciation allowed for tax purposes. Denote the true rate of depreciation by δ and the rate allowed for tax purposes by δ‘. In this case $V(\tau_H)$ is given by

$$V(\tau_H) = \max_{(K,L)} \left\{ \frac{[A_H F(K,L) - w_H L](1-\tau_H) + \tau_H \delta\text{‘} K + (1-\delta)K}{1+(1-\tau_H)r} - (K - K_H^0) \right\}, \tag{9.1}$$

where, as before, K demotes the post-investment capital stock, K_H^0 is the existing (post-depreciation) stock of capital, L is the labor input, w_H is the host-country wage rate, and r is the world rate of interest. Note that in the presence of taxation, the discount rate is the after-tax rate $-(1-\tau_H)r$. (This specification assumes that the subsidiary uses debt in the host country to finance the new investment.) Employing the envelope theorem, it follows from equation (9.1) that $\partial V/\partial \tau_H < 0$. That is, the present value of the subsidiary's cash flow falls when the corporate-tax rate in the host country rises, as is indeed expected. Furthermore, the amount of new investment (the flow of FDI) depends negatively on τ_H. The first order condition for the stock of capital becomes now

$$A_H F_K(K,L) = r + \frac{\delta - \delta' \tau_H}{1-\tau_H}. \tag{9.2}$$

This latter equation defines (implicitly) an equation for the flow of FDI. As δ' is typically smaller than δ, it follows that $\partial K/\partial \tau_H < 0$.[5]

The parent firm will indulge into the project if

$$C(1-\tau_S) \leq V(\tau_H), \tag{9.3}$$

where τ_S is the corporate tax rate in the source country. The latter equation defines a selection mechanism generated by a threshold barrier.

To sum up: as is evident from condition (9.3), the tax rate in the source country, τ_S, affects positively the decision by a parent firm in country S whether to carry out a foreign direct investment in country H; whereas the tax rate in the host country, τ_H, has a negative effect on this decision. The tax rate in the source country, τ_S, is irrelevant for the determination of the magnitude of FDI flows; the latter are negatively affected by τ_H.

9.3 Empirical Evidence

As in the preceding chapters, the Heckman selection-bias method is adopted to jointly estimate the likelihood of surpassing a threshold (the selection equation) and the magnitude of the FDI flows, provided that the threshold is indeed surpassed (the flow or gravity equation). The empirical study is confined to 24 OECD countries over the period from 1981 to 1998, as in Chapter Seven. The data source is also described in the latter chapter. We naturally add the host and source corporate-tax rates to the list of explanatory variables.[6],[7]

Table 9.1 presents the effects of several potential explanatory variables of the two-fold decisions on FDI flows. Our focus is on the role of the source and host corporate-tax rates, as all other variables were discussed in the preceding chapters.

(Table 9.1 about here)

As explained in detail in Chapter Six, the OLS estimates of the effects of these variables are biased. This is true for both the OLS-D regression, where the observations with no FDI flows are discarded (leaving only 851 observations out of the 2116 observations in the full sample); and for the OLS-Zero regressions, where the no-flow observations were recorded as having FDI flows of zero[8]. Note that the difference in the coefficients between OLS-D and OLS-Zero indicate that there exist non linear relationships between the dependent variable and the independent variables. The Heckman method is suitable for estimating such non linear relationships, as discussed also in Chapter Seven. The Heckman joint estimation of the flow and selection equations are presented in the last two columns. We exclude certain variables from the flow equation for identification. These are the source financial sound rating variable and the previous FDI dummy variable (one for the existence of previous FDI). The existence of previous FDI may play a role in reducing the fixed costs assiciated with new FDI, but is not expected to affect the current flows of FDI. This is why we put this variable in the selection equation but not in the flow equation. Similarly to what is explained in

Chapter Five with respect to the source-country productivity factor in the case of greenfield FDI, the source-country financial sound rating may play a role in the selection mechanism, but not in the flow equation. The results concerning the non-tax variables are more or less in line with the findings in Chapters Seven and Eight.

For instance, a high gap in education in favor of the source country reduces the probability of having FDI flows to the host country. This is expected because a gap in years of schooling may be a proxy for a productivity gap; see also Chapter Five. The host financial sound rating affects positively the flow of FDI, whereas the analogous variable of the source country (an exclusion restriction variable) is negative and significant in the selection equation. Finally, the existence of past FDI relations (another exclusion restriction variable) is positive and significant in the selection equation, as it may help to reduce the setup costs of establishing a new FDI flow.

We turn now to the main focus of this chapter - the effect of corporate-tax rates. First, the source corporate-tax rate is positive and significant in the selection equation, as indeed predicted by condition (9.3). This rate plays no statistically significant role in the flow equation, again in line with our analysis. The coefficient of the host corporate-tax rate is insignificant in the selection equation. But it is negative and significant in the flow equation, again as predicted by our analysis. Note that it is not merely the source-host tax differential $(\tau_S - \tau_H)$ which is the main determinant of FDI flows.

Interestingly, the role of the source and host corporate-tax rates is not properly revealed by the traditional OLS regressions. In the first regression (OLS-D), only the host corporate-tax rate plays a statistically significant role in reducing FDI flows to the host country; whereas in the other regression (OLS-Zero), it is only the source corporate-tax rate which plays a statistically significant role in promoting FDI flows from the source country[9]. **Thus,OLS analysis does not detect a role for both tax rate stop lay in the determination of FDI.**

As in the preceding chapter, we note that the relationship in the selection equation between the probability (P) of making a new FDI and the explanatory variables (including τ_S) is not linear. It is rather given by

$$P(\tau_S) = \int_{-\infty}^{\alpha+\tau_S\theta_\tau} (2\pi)^{-1/2} \exp(-y^2/2)dy, \tag{9.4}$$

where α represents the effect of all the other explanatory variables (held fixed at their sample averages), including country fixed effects, and θ_τ is the coefficient of τ_S in the selection equation. Note also that the estimate of θ_τ is positive and statistically significant. The marginal effect of τ_S on P is

$$\partial P/\partial\tau_S = \theta_\tau (2\pi)^{-1/2} \exp[-(\alpha + \theta_\tau\tau_S)^2/2] > 0. \tag{9.5}$$

(Figure 9.1 about here)

Figure 9.1 depicts the graph of the function $P(\tau_S)$ for the U.S. as a source country and four EU countries (Denmark, Greece, the Netherlands and the

U.K.) as host countries. The U.S.-U.K. characteristics in the sample are such that the estimated probability of a positive FDI flow from the U.S. to the U.K. is one, unaffected by the source country (namely, U.S.) tax rate. For all other three countries, the U.S. tax rate has a strong positive effect in the relevant range of 0-40%. But the marginal effect of the source-country tax rate is not the same for all three countries, being highest for Greece. Figure 9.2 depicts the flow equation for the U.S., as a source country, and the four EU countries as host countries. The host-country tax rate seems to have a negative effect at all rates, including the very high rates that approach 100%. Notably, the tax rate of the U.K. (as a host country) has a very strong negative marginal effect, whereas the tax rate of Greece has a relatively small marginal effect.

(Figure 9.2 about here)

9.4 Conclusion

We analyze in this chapter the effects of taxes on bilateral FDI flows. Evidently, economists and policymakers reckon with the fact that taxes do affect economic activity. Bilateral FDI flows are no exception. Our aim is to bring out the special mechanisms through which taxes influence FDI, when investment decisions are likely to be two-fold because of the existence of fixed setup costs of new investments. As explained throughout this book, for each pair of source-host countries, there is a set of factors determining whether aggregate FDI flows will occur, and a different set of factors determining the volume of FDI flows, given that they occur at all. We demonstrate in this chapter that the notion that the mere international tax differentials are the main factors behind the direction and magnitude of FDI flows is too simple. We hypothesize that the source-country tax rate works primarily on the selection process, whereas the host-country tax rate affects mainly the magnitude of the FDI, once they occur.

Analyzing an international panel data of 24 OECD countries, we bring empirical evidence, using selection bias methods, in support of the predictions of our theoretical conception. Our findings have some bearing on the issue of race to the bottom in the discussion of international tax competition.[10] We find that the effects on bilateral FDI flows of a tax cut in the source country is quantitatively different than a tax cut in the host country. This means that there could be more intense race-to-the-bottom tax cuts among predominantly host countries compared to source countries.

Appendix 9A: Basic Principles of International Taxation of Capital Income

Two common principles of international taxation that are the foundations of many national tax systems are the *residence principle* and *source principle*. The residence principle employs the place of residency of the taxpayer as the basis for assessment of tax liabilities, whereas the source principle emphasizes the source of income as the basis for assessing tax liabilities. According to the residence principle, residents of a country are taxed uniformly on their worldwide incomes, regardless of the source of that income (domestic or foreign). Similarly, nonresidents are not taxed by the home country on their income originating in that country. According to the source principle, incomes originating in a country is uniformly taxed, regardless of the residency of the income recipient. In addition, residents of a country are not taxed by it on their foreign-source incomes.

In practice, countries adopt mixtures of these two pure (polar) principles of international taxation. For instance, a country may tax its residents on their world-wide income and, at the same time, tax also non-residents on the income that originates in their country.

Consider a standard two-country world (home and foreign) and denote the interest rates in the home and the foreign countries by r and r^*, respectively. In general, the home country may have three different effective tax rates that apply to interest income:

τ_D - tax rate levied on residents on their domestic-source income.

τ_F - effective tax rate levied on residents on their foreign-source income, in addition to the tax already levied in the foreign country.

τ_{ND} - tax rate levied on nonresidents on their interest income originating in the home country.

Correspondingly, the foreign country may also have three tax rates, which we denote by τ_D^*, τ_F^*, τ_{ND}^*. With complete integration of capital markets between the two countries (including the possibility of borrowing in one country in order to invest in the other country), arbitrage possibilities imply that

$$r(1 - \tau_D) = r^*(1 - \tau_{ND}^* - \tau_F), \tag{9A.1}$$

and

$$r^*(1 - \tau_{ND} - \tau_F^*) = r^*(1 - \tau_D^*). \tag{9A.2}$$

Equation (9A.1) applies to the residents of the home country. It implies that in equilibrium these residents are indifferent between investing at home or abroad. If this equity was violated, then the home-country residents could borrow unlimited amounts in the low (net of tax) interest rate country and invest these borrowed funds in the high (net of tax) interest rate country, thereby generating unlimited profits. Similarly, equation (9A.2), which applies to residents of the foreign country, rules out such unlimited profit opportunities to foreign residents.

The two polar international tax principles have different efficiency implications. In a world with international capital mobility the equality between saving and investment need not hold for each country separately but rather for world aggregate saving and investment. This separation raises the issue of the efficiency of the international allocation of the world investments and savings. A detailed analysis is provided in Frenkel, Razin and Sadka (1991).

Here we provide only a very concise treatment.

If the two countries adopt the residence principle, then

$$\tau_D = \tau_{ND}^* + \tau_F, \tag{9A.3}$$

$$\tau_D^* = \tau_{ND} + \tau_F^*, \tag{9A.4}$$

and

$$\tau_{ND} = \tau_{ND}^* = 0. \tag{9A.5}$$

Equation (9A.3) states that a resident of the home country is levied *effectively* the same tax rate whether she invests at home (τ_D) or abroad ($\tau_{ND}^* + \tau_F$). Equation (9A.4) states the same thing for a resident of the foreign country. Equation (9A.5) states that there is indeed no source taxation levied by any country on nonresidents. Substituting equation (9A.3) into equation (9A.1) [or equation (9A.4) into equation (9A.2)] yields the (before-tax) rate of return equalization between countries: $r = r^*$. That is, the international allocation of the world stock of capital is efficient. However, if the tax rates are not the same in all countries, then the net returns accruing to savers in different countries vary [that is, $(1 - \tau_D)r \neq (1 - \tau_D^*)r^*$], and the international allocation of world savings is distorted.

On the other hand, if both countries adopt the source principle, then:

$$\tau_D = \tau_{ND}, \tag{9A.4'}$$

$$\tau_D^* = \tau_{ND}^*, \tag{9A.5'}$$

and

$$\tau_F = \tau_F^* = 0. \tag{9A.6'}$$

Substituting equations (9A.4')-(9A.6') into equations (9A.1)-(9A.2) yields the (net of tax) rate-of-return equalization between residents of different countries: $r(1-\tau_D) = r^*(1-\tau_D^*)$. That is, the international allocation of world savings is efficient. However, if the tax rates are not the same in all countries (that is, $\tau_D \neq \tau_D^*$), then $r \neq r^*$, and the international allocation of the world stock of capital is not efficient.

It is worth mentioning that the public economics literature emphasizes the superiority of the efficiency in the international allocation of the world stock of capital over the efficiency of the international allocation of the world savings when taxes are optimally designed[11].

Notes

[1]For a succinct review of this issue see, for example, Hines (2001).

[2]This is also the recommendation of the OECD model tax treaty (OECD, 1997). A similar recommendation is made also by the United Nations model tax treaty (U.N. 1980).

[3]The recent Jobs Creation Act in the U.S. (2005) allows U.S. companies to pay merely a tax of 5.25% on their foreign-source income.

[4]Of course, there still may be other elements of fixed (setup) cost incurred in the host country, as in Chapter Five. To highlight the differential roles of host and source tax rates we abstract here from these elements.

[5]In an oligopolistic competition environment, the effects of τ_H on investments may be opposite to the effects we find in our perfect competition environment; see, for instance, Devereux and Hubbard (2003).

[6]We simply apply the statutory rates, because they are exogenously given. Average effective tax rates, suggested by Deverux and Griffith (2003) as determinants of the location of investments, are endogenous in the sense that they are determined by the amount of investment. To apply econometrically average effective tax rates, there is a need for a good instrument. The statutory rate is the best available instrument.

[7]The data source for these rates is: World Tax Database (University of Michigan) http://www.bus.umich.edu/otpr/worldtaxdatabase.htm

[8]More accurately, as we measured FDI by logs, we put a large negative

number for these FDI flows.

[9]We performed also several robustness tests. For instance, we excluded the host and source financial risk ratings from the flow equation. We also deleted the variable "previous FDI" from both equations. The results concerning the corporate-tax rates seem quite robust.

[10]For an overview of the traditional analysis of international taxation see, for instance, Frenkel, Razin and Sadka (1991), Wilson (1999), and Haufler (2001).

[11]This is as implication of work of Diamond and Mirrlees (1971). For details see, for instance, Frenkel, Razin and Sadka (1991).

Part IV

Policy in a Globalized Economy

Chapter 10

Tax Competition and Coordination

10.1 Introduction

In the globalized economy the issues of tax competition and coordination are becoming increasingly pressing for policy makers. It is especially relevant for the taxation of income from internationally mobile factors, such as the income generated by FDI. As the Economist (1997, pp. 17-18): succinctly puts it:

> "Globalization is a fax problem... First, firms have more freedom over where to locate... First, firms have more freedom over where to locate... This will make it harder for a country to tax [a business] more heavily than its competitors... Second, globalization

> makes it harder to decide where a company should pay tax, regardless of where it is based... This gives them [the companies] plenty of scope to reduce tax bills by shifting operations around or by crafting transfer-pricing..."

With internationally mobile factors, the tax base can shift from one (high-tax) country to another (low-tax) country, thereby creating a fiscal externality[1]. The latter externality can take also the form of a government trying to impose the burden of financing public goods onto non-residents such as foreign corporations. as with any externality, (tax) competition among countries may lead to inefficient ??? of taxes and public good provision; whereas tax coordination among them may enhance the welfare of all of them.[2]

In this chapter we draw on the findings of the preceding chapter to shed some light on various aspects of international tax competition and coordination concerning the flows of FDI. Specifically, we take another look at the implications of FDI for the effects of taxation and for the tax bases in a source-host country setup. We analyzed in the preceding chapter the asymmetric mechanisms through which source and host taxation affect FDI. Here we analyze this asymmetry to explain the endogenous coexistence in a tax competition environment of high-tax, high public expenditure source countries and low-tax, low-public expenditure host countries[3]. Such differences may be a feature of the enlarged European Union characterizing the asymmetry between the old members countries and the new accession countries. In addition, we attempt to provide some indication about the magnitude of the gains from tax coordination and the implication for the "race to the bottom" hypothesis.

10.2 A Source-Host Country Model of Taxes and Public Goods

10.2.1 Production

Consider a host country with a continuum of firms, each with a productivity level factor of ε, where $\varepsilon > -1$; the density and the cumulative distribution functions are denoted, respectively, by g and G. We normalize the number of firms to one. Unlike in Chapter Three, the productivity factor is not random. It is known to all before any economics decision is made. Firms are thus ex ante and ex post different in their productivity levels.

Assume for simplicity that the initial stock of capital of the firm is zero. A firm with a productivity factor ε (an ε-firm) employs a capital stock of K in the first period and produces an output of $A_H F(K)(1+\varepsilon)$ in the second period, where F exhibits a diminishing marginal productivity of capital ($F' > 0, F'' < 0$). As before, there are setup costs of new investment. Therefore, only firms with a productivity factor above some threshold level (ε_0) will make new investments.

We continue to assume that foreign direct investors (from the source country) have a cutting edge advantage over domestic investors with respect to these setup costs, so that they acquire control over the domestic firms in the host country. Foreign direct investors compete among themselves for these firms. Therefore, the price they pay for an ε-firm (with $\varepsilon \geq \varepsilon_0$) to the original domestic owners is $V_H(\varepsilon, \tau_H) - C^*(1 - \tau_S)$, where $V_H(\varepsilon, \tau_H)$ is defined by:

$$V_H(\varepsilon, \tau_H) = \max_K \left[\frac{A_H F(K)(1+\varepsilon)(1-\tau_H) + \tau_H \delta'_H K + (1-\delta)K}{1 + (1-\tau_H)r} - K \right], \tag{10.1}$$

where C^* is the setup cost borne by the foreign direct investor. As before, this cost is born in the source country and tax-deducted there. The parameters δ and δ'_H denote, as before, the physical and the tax rate of depreciation, respectively, and τ_i denotes corporate tax rate in country $i = H, S$. As explained in the preceding chapter, the foreign direct investors do not pay any further tax at their home country. We assume that the two countries are open to the world credit market. This assumption fixes the rate of interest at the world rate, denoted by r.

The first order condition for the optimal stock of capital of an ε-firm is given by

$$A_H F'(K)(1+\varepsilon) = r + \delta + \frac{\tau_H}{1-\tau_H}(\delta - \delta'_H) \tag{10.2}$$

for firms with $\varepsilon \geq \varepsilon_0$. This condition defines the optimal stock of capital of a firm as a function $K^H(\varepsilon, \tau_H)$ of its productivity factor and the host corporate tax rate.

The cutoff level of the productivity factor is a function $\varepsilon_0(\tau_H, \tau_S)$ of τ_H and τ_S, defined implicitly by:

$$V_H(\varepsilon, \tau_H) - (1 - \tau_S)C^* = 0. \tag{10.3}$$

That is, the ε_0-firm is indifferent between investing and not investing. Note that because of the setup cost advantage of the foreign direct investors, firms that are not purchased by these investors will not invest at all under domestic ownership, and their value is zero. (Recall that the initial stock of capital of the firm is zero.)

As we plausibly assume that the depreciation rate allowed for tax purposes (δ'_H) is below the true physical rate (δ), it follows from equation (10.2) that τ_H depresses the stock of capital of each investing firm. It also follows from condition (10.3) that τ_H reduces the number of investing firms (that is, increases ε_0). Therefore the host corporate tax rate reduces the total stock of capital in the host country. In contrast, it follows from condition (10.3) that τ_S increases the number of investing firms (that is, lowers ε_0). Therefore, an increase in the source corporate tax rate raises the capital stock in the host country.

The source country is modeled similarly. As it is the differences in the production and cost parameters are key to the determination of the direction and magnitude of FDI flows, we simplify by assuming that the fixed costs in the source country are nil. Thus, all firms invest. The value of an $\varepsilon-$firm is

$$V_S(\varepsilon, \tau_S) = \underset{K}{\text{Max}} \left[\frac{A_S F(K)\,(1+\varepsilon)\,(1-\tau_S) + \tau_S \delta'_S K + (1-\delta)K}{1+(1-\tau_S)r} - K \right]. \tag{10.4}$$

The optimal stock of capital of an ε-firms is given by the marginal productivity condition:

$$A_S F'(K)\,(1+\varepsilon) = r + \delta + \frac{\tau_S}{1-\tau_S}(\delta - \delta'_S). \tag{10.5}$$

This equation yields the optimal stock of capital as a function, $K_S(\varepsilon, \tau_S)$ of ε and τ_S.

10.2.2 Private Consumption

A representative consumer in country $i = S, H$ has an initial endowment I_i in the first period and a utility function

$$u\,[v(x_1, x_2), P] \tag{10.6}$$

over first-period consumption (x_1), and second-period consumption (x_2), and public expenditures (P). These expenditures can represent public good provision. Weak separability is assumed between (x_1, x_2) and P, so that public expenditures do not affect private demands for first and second-period consumption. For simplicity, it is assumed that P is incurred in the first period. Note that we consider, purely for simplicity, a representative consumer model; it is straightforward to extend it to a many-consumer model in which the public expenditures can reflect redistributive transfers. The tax rate τ_i applies also to the interest income of the consumers, both at home and abroad.[4] Note that we assume identical preferences in the two countries; that, is the same u and v for the two countries. However, the identical preferences assumption does not mean that the two countries have a demand for the same quantity of the public good (P). This is because they do not have the same income. We assume that I_S is significantly higher than I_H. That is, the source country is rich and the host country is poor. Assuming plausibly that the public good is a normal good, the rich-source country will have a greater demand for the public good (namely, for tax revenues) than the poor-host country. We employ this specification in order to single out the cross-country income gap as the driving force for the ensuing cross-country differences in tax policy in the tax-competition equilibrium. (For this reason, we also specified the same production function F for the two countries.)

Utility maximization yields the individual consumption demands for the first and the second periods:

$$X_j\left[W_i, (1-\tau_i)r\right], \qquad j = 1, 2;\; i = H, S, \tag{10.7}$$

where W_i is the income of a representative consumer in country i. Note that the demand functions are identical for the two countries, as we assumed identical preferences.

The income of a representative consumer in the host country consists of the initial endowment, plus the proceeds from the sales of the domestic firms (with productivity factors above ε_0) to the foreign direct inventors. That is, W_H is given by:

$$W_H(\tau_H, \tau_S) = I_H + \int_{\varepsilon_0(\tau_H,\tau_S)}^{\infty} V_H(\varepsilon, \tau_H) g(\varepsilon) d\varepsilon - (1-\tau_S) C^* \left\{1 - G\left[\varepsilon_0(\tau_H, \tau_S)\right]\right\}. \tag{10.8}$$

(Note that the number of firms within a productivity factor above ε_0 is $1 - G(\varepsilon_0)$.)

The income of the representative consumer in the source country (who retains also all the firms in this country) is similarly given by:

$$W_S(\tau_S) = I_S + \int_{-1}^{\infty} V_S(\varepsilon, \tau_S) g(\varepsilon) d\varepsilon. \tag{10.9}$$

Note that the representative consumer in the source country, who is the foreign direct investor in the host country, pays for the purchased firms prices that exhaust entirely the profits she gets from them. (This follows from the assumed perfect competition among the foreign direct investors.)

10.2.3 Government

Each government balances its budget: tax revenues must suffice to finance public expenditures. This is done over time in present value terms, given the free access to the world credit market. By Walras' law the government's budget constraint can be replaced by an economy-wide resource constraint.

Consider first the host country. The representative consumer sells an ε-firm at a price $V_H(\varepsilon,\tau_H)-(1-\tau_S)C^*$. This price reflects the cash flow of the ε-firm, after taxes paid to the host country government. We emphasize that these taxes are paid to the host country government. Hence, from the point of view of the resources available to the host country, the price paid by the foreign direct investors must include also these taxes (which serve to finance public expenditures). Put differently, the host country extracts from the foreign direct investor the before-tax cash flow of the purchased ε-firm, that is $(1+r)^{-1}\{A_H F[K_H(\varepsilon,\tau_H)](1+\varepsilon)+(1-\delta)K_H(\varepsilon,\tau_H)-K_H(\varepsilon,\tau_H)\}-(1-\tau_S)C^*$. Therefore, the economy-wide resource constraint of the host country is

$$\begin{aligned} P_H = \quad & I_H + (1+r)^{-1} \int_{\varepsilon_0(\tau_H,\tau_S)}^{\infty} \{A_H F[K_H(\varepsilon,\tau_H)](1+\varepsilon) + (1-\delta)K_H(\varepsilon,\tau_H)\} g(\varepsilon) d\varepsilon \\ & - \int_{\varepsilon_0(\tau_H,\tau_S)}^{\infty} K_H(\varepsilon,\tau_H) g(\varepsilon) d\varepsilon - (1-\tau_S)C^* \left\{1 - G\left[\varepsilon_0(\tau_H,\tau_S)\right]\right\} \\ & - X_1\left[W_H(\tau_H,\tau_S),(1-\tau_H)\,r\right] - (1+r)^{-1} X_2\left[W_H(\tau_H,\tau_S),(1-\tau_H)\,r\right]. \end{aligned} \tag{10.10}$$

Note from equation (10.10) that the source country effectively subsidizes the host country through the tax deductibility of the fixed setup costs.[5] The magnitude of this subsidy is $\tau_S C^* \{1 - G[\varepsilon_0(\tau_H, \tau_S)]\}$.

Similarly, the economy-wide resource constraint in the source country is given by

$$\begin{aligned} P_S = \quad & I_S + (1+r)^{-1} \int_{-1}^{\infty} \{A_S F[K_S(\varepsilon, \tau_S)] + (1-\delta) K_S(\varepsilon, \tau_S)\} g(\varepsilon) d\varepsilon \\ & -\tau_S C^* \{1 - G[\varepsilon_0(\tau_H, \tau_S)]\} - \int_{-1}^{\infty} K_S(\varepsilon, \tau_S) g(\varepsilon) d\varepsilon \\ & -X_1 [W_S(\tau_S), (1-\tau_S) r] - (1+r)^{-1} X_2 [W_S(\tau_S), (1-\tau_S) r] . \end{aligned} \tag{10.11}$$

Note again the source country subsidizes the host country by the amount of tax deductions allowed for the fixed setup costs.

10.3 Tax Competition

Each government attempts to maximize the welfare of its representative consumer. In doing so, each government takes the policy of the other government as given. We thus look at a Nash-equilibrium of the two country tax competition game.

Formally, the government of the host country chooses the corporate tax rate (τ_H), so as to maximize the utility of the representative consumer,

$$u\left(v\{X_1\left[W_H\left(\tau_H,\tau_S\right),\left(1-\tau_H\right)r\right],X_2\left[W_H\left(\tau_H,\tau_S\right),\left(1-\tau_H\right)r\right]\},P_H\right), \tag{10.12}$$

where the quantity of the public good (P_H) is given by the economy-wide resource constraint (10.10). The source country tax rate (τ_S) is considered by the host government as exogenously given.

Similarly, the source government chooses τ_S so as to maximize

$$u\left(v\{X_1\left[W_S\left(\tau_S\right),\left(1-\tau_H\right)r\right],X_2\left[W_S\left(\tau_S\right),\left(1-\tau_H\right)r\right]\},P_S\right), \tag{10.13}$$

where P_S is given by the economy-wide resource constraint (10.11), and where τ_H in the latter constraint is taken as exogenously given.

The optimal policy (namely, the corporate tax rate) chosen by the host country depends on the source country tax rate (τ_S). Thus, this policy may be thought of as a best-response function of τ_S; denote it by $\hat{\tau}_H(\tau_S)$. Similarly, denote the best-response function of the source country by $\hat{\tau}_S(\tau_H)$. A Nash-equilibrium is a pair of tax policies, (τ_H^*,τ_S^*), such that $\tau_H^*=\hat{\tau}_H(\tau_S^*)$ and $\tau_S^*=\hat{\tau}_S(\tau_H^*)$.

We resort to numerical simulations in order to characterize the Nash-equilibrium and study the effect of the source-host income gap (I_S/I_H) and the setup cost (C^*) on the divergence or convergence the tax-expenditure policies.

In these simulations, we employ a Cobb-Douglas production function $F(K) = K^{\alpha}$, with $\alpha = 2/3$. The parameter values for the productivity levels are $A_H = A_S = 1$. We employ a Cobb-Douglas utility function, $u = \ln x_1 + \beta \ln x_2 + \gamma \ln P$, with $\beta = 0.99$ and $\gamma = 0.95$. The parameter values for the depreciation rates are $\delta_H = \delta_S = 0.2$ and $\delta'_H = \delta'_S = 0.1$. The world rate of interest is $r = 0.05$. We also set the initial endowment at the host country (I_H) at unity.

(Figure 10.1 about here)

Figure 10.1 illustrates the effect of a rise in the initial endowment in the source country (I_S) on the optimal tax-expenditure policies at the Nash equilibrium. (The parameter value for the setup cost is $C^* = 1$.) The host tax rate (τ_H) and the public expenditures (P_H) are not effected by I_S. But, as the source country becomes richer, its tax rate and expenditures rise, thus yielding an equilibrium with low-tax, low-expenditures in the relatively poor host country, and high-tax, high-expenditures in the relatively rich source country.

(Figure 10.2 about here)

Figure 10.2 depicts the effect of the setup cost on the tax-expenditure policies. (The parameter value for initial endowment in the source country,

I_S, is set to unity.) With $C^* = 0$, the two tax rates are equalized at about 23.5%. As the setup cost rises, both tax rates fall, but the host rate falls more sharply. Thus, we get an equilibrium with a low-tax, low-expenditure host country and a high-tax, high-expenditure source country.

10.4 Tax Coordination

In our two-country setup there are clearly some tax externalities. Perhaps, most pronouncely, the fixed setup costs associated with FDI in the host country is subsidized by the source country through the deductibility of these costs; the amount of the subsidy being $\tau_S C^* \{1 - G[\varepsilon_0(\tau_H, \tau_S)]\}$. Similarly, the amount of this subsidy is affected by the corporate tax in the host country (τ_H) through the number, $1 - G[\varepsilon_0(\tau_H, \tau_S)]$, of FDI firms. Therefore, tax competition yields a Pareto-inefficient outcome from the joint point of view of the source and host countries. Tax coordination between the two countries yields Pareto-improvement.

We use numerical simulations (with the same parameter values as in the preceding section) in order to illustrate the gains from tax coordination, relative to tax competition. The two countries coordinate their tax/expenditure policies but we do not allow direct transfers between them. Thus, the host and source countries abide by the two country-specific resource constraints [equations (10.10) and (10.11), respectively], rather than just by one resource constraint by the summation of these two constraints. But they jointly determine their tax/expenditure policies. For concreteness, we consider the policies that assign all the gains from tax coordination to the host country; the source country attains the same level of utility (for the representative consumer) under either tax competition or coordination. We measure the utility gains from tax coordination by the equivalent percentage increase in first and second period consumption as well as the public good (X_1, X_2 and P, respectively) of the representative consumer in the host country.

Figure 10.3, panel (a) depicts these gains for various values of the setup cost (C^*). The magnitude of the gain is rather small for the parameter values we chose, but what is important is the pattern they follow: as expected, the gains rises with the size of the fixed setup cost. Panel (b) of this figure depicts the gains for various values of the income of the representative consumer in the source (rich) country. The richer the source country the higher is the gain.

(Figure 10.3 about here)

Figure 10.4 illustrates the effect of tax coordination on the tax rates themselves for various values of C^* and I_s. As expected, tax coordination leads to higher tax rates, in both countries, thus checking the "race to the bottom". It is useful also to note the differential effect of the setup cost (C^*) on the competitive and coordinated tax rates in the source country; see panel (a). As was pointed out earlier, the amount by which the source country subsidizes the host country is positively related to C^* and the source country corporate tax rate (τ_S). Under tax competition, as C^* rises the source country cuts it corporate tax rate in order to reduce the amount of the subsidy. This does no longer hold under tax coordination. (Recall also that all the gain accrue to the host country.)

(Figure 10.4 about here)

10.5 Conclusion

The 2004 enlargement of the EU with ten new countries provides a stylized analogue of the predictions of the model. Table 10.1 describes the corporate tax rates in the 25 EU countries in 2003. It reveals a marked gap between the original EU-15 countries and the 10 accession countries. The latter have significantly lower rates. Estonia, for instance, has no corporate tax; the rates in Cyprus and Lithuania are 15%; and in Latvia, Poland, and Slovakia, 19%. In contrast, the rates in Belgium, France, Germany, Greece, Italy, and the Netherlands range from 33% to 40%.

(Table 10.1 about here)

Note, however, that the tax rates mentioned are the *statutory* rates. However, what matters from an economic point of view is the *effective* tax rates, which could be significantly different because if different statutory tax bases, tax loss treatment, etc. Nevertheless, there are some indications that the effective tax rates are also lower in the accession countries than in the EU-15 countries. For instance, Jakubiak and Markiewicz (2005) show that the ratios of corporate tax revenues to GDP in the former countries are on average lower than in the latter countries; see their Figure 2.

Given the fiscal externalities, tax normalization is naturally beneficial to all countries involved. However, depending on the actual specifics of the harmonization, it may benefit some countries at the expense of others. In the EU, some of the original six founding countries may push for tax harmonization, whereas most of the 2004 accession countries are reluctant. Furthermore, given the built-in revenue transfer mechanism within the EU, the former countries feel they are financing the tax cuts of the latter countries. Gerhard Schroder, the former German Chancellor, stated in April 2004 that it was unacceptable "that Germany, as the EU's biggest net payer, finances unfair tax competition against itself".

Notes

[1]Similar phenomenon may occur among jurisdictions within a federation; see, for instance, Gordon (1983) and Mintz and Tulkens (1986).

[2]Again, this is true also within a federation of various local government; see, for instance, Zodrow and Mieszkowski (1986).

[3]When there are also internationally immobile factors of production tax competition may affect also the relative tax burden falling on these factors in a way depending on their complementarity or substitutability with the mobile factors; see, for instance, Wildasin (2004).

[4]Note that we assume that corporate income is taxed only at home - at the corporate level. Each country taxes individuals and corporations at the same rate.

[5]When the source country does manage to tax the (resident) parent company on its income from the FDI subsidiary, then it loses tax revenues to the host country through the foreign tax credit clause that is usually granted to avoid double taxation.

Epilogue

Capital flows of all types have increased over the past few decades, and most recently some of the biggest increases have occurred in foreign direct investment. This is especially true of rich countries, but is also increasingly true of developing nations too.

These flows in the form of FDI are also important because it is believed that FDI has special benefits over other forms of capital flows. First, these flows are thought to be more stable, and do not leave the host country exposed to financial crises. Second, FDI if often supposed to be associated with technology transfer which may have spillover benefits for the host country. Third, FDI is often attributed a special role in disciplining host country governments: the threat of moving business offshore limits the ability of host countries to extract taxes and introduce inefficient regulations. Offsetting this, FDI is also often associated with special domestic costs in cases where foreign ownership has caused domestic political unrest in the host country (this is often especially true in natural resource industries).

This book studies determinants of flows of foreign direct investment. A key feature throughout the analysis in this book is the existence of fixed costs associated with FDI decisions. It presents a suite of new models that are useful in thinking about FDI, presents some new data on FDI and provides some empirical techniques that help understand some of the patterns that are present in FDI flows among developed countries and between developed and developing countries.

It is commonly heard in policy circles that FDI is illiquid and that foreign direct investors trade off any potential cost or other advantage over domestic investors they may bring against the fact that their investments will be illiquid. This intuition is formally captured in the book. In particular, ownership is modeled as conveying earlier access to information about the productivity of the firm. This generates benefits in terms of planning the investment strategy. However, because this information is private to the foreign direct investor, it also leads to a lemon problem. That is, if the investors need to sell the project, they face the problem that potential buyers fear that the sale is motivated by private information about low productivity, instead of a genuine need for liquidity. This means that firms that are eventually sold attract a lower price than otherwise: they are illiquid.

Foreign direct investors may possess some "intangible capital" that grant them some advantage over domestic investors in skimming the best projects. The benefits from this unique advantage may be fully or partially shifted to the domestic country, depending on the intensity of competition among foreign direct investors. This shift of benefits occurs through the acquisition price the foreign direct investors pay at the host country for the best projects. This adds a new element to the gains-from-trade argument, over the standard trade-based and rate of return equalization gains.

The fixed costs makes FDI decisions twofold. Standard equalization of marginal productivities with markets costs (of capital and labor) determine the volume of FDI flows, if they are worthwhile at all. But the likelihood that these flows actually occur depends on the magnitude of the fixed costs. The Heckman selection bias method is applied to jointly study the determinants of both the likelihood and the volume of FDI. Most importantly, an empirical evidence is provided for the existence of fixed costs and for their role in determining the likelihood of FDI flows.

The fixed costs contain an important component denominated in domestic labor. Therefore, a positive productivity shock in the host country, which raises labor costs, may be associated with a tendency for a lower likelihood of attracting new FDI. However, if new FDI are attracted, their volume tend to be higher.

The fixed costs also adds a twist to the analysis of tax competition with respect to the flow of FDI. Interestingly, tax competition does not necessarily lead to a "race of the bottom" in the taxation of capital income. Also, higher fixed costs in the host country (relative to the source country) increases the tax rate gap between them.

A future research agenda may analyze some related topics.

One striking feature of FDI flows to developing countries is that their share in total inflows is often higher in riskier countries, with risk measured either by countries' credit ratings for sovereign (government) debt or by other indicators of country risk; see Loungani and Razin (2001). There is also some evidence that we share of FDI in total inflows is often higher in countries where the quality of institutions is lower. One possible explanation is that FDI is more likely than other forms of capital flows to take place in countries with missing, or inefficient, markets. In such settings, foreign investors will prefer to operate directly instead of relying on local financial markets, suppliers, or legal arrangements. The policy implications of this view, according to Albuquerque (2003), are "that countries trying to expand their access to international capital markets should concentrate on developing credible enforcement mechanisms instead of trying to get more FDI." This topic would benefit from further theoretical and empirical analysis.

The econometrics and the data could be further usually elaborated on in order to provide empirical analysis of some structural features of the models presented in the book. For instance, an important implication of the theory hinges on the share of domestic labor in the fixed costs. As the fixed costs are unobserved, it would pose a challenge to elicit from the data the share of labor in these costs. Another research issue relates to the role of illiquidity in determining the allocation of foreign investors between FDI and FPI. Again, it would be useful to confront the predictions of the model concerning liquidity and foreign investment with data.

Bibliography

[1] Admati, Anat R. and P. Pfleiderer (1988), "A Theory of Intraday Patterns: Volume and Price Variability", *Review of Financial Studies*, 1, 3-40.

[2] Admati, Anat R. and P. Pfleiderer (1991), "Sunshine Trading and Financial Market Equilibrium", *Review of Financial Studies*, 4, 443-481.

[3] Akerlof, George. A. (1970), "The Market for Lemons: Quality Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, 84, 488-500.

[4] Albuquerque, Rui (2003), "The Composition of International Capital Flows: Risk-Sharing through Foreign Direct Investment", *Journal of International Economics*, 61(2), 353-383.

[5] Alfaro, Laura, Areendam Chanda, Sebnem Kalemli-Ozcan, and Selin Sayek (2004), "FDI and Economic Growth: The Role of Local Financial Markets", *Journal of International Economics*, 64(1), 89-112.

[6] Amiti, Mary (2005), "Location of Vertically Linked Industries: Agglomeration versus Comparative Advantage", *European Economic Review*, 49(4), 809-832.

[7] Antras, Pol (2004), "Advanced Topics in International Trade: Firms and International Trade", Lecture Notes (Harvard University).

[8] Auerbach, Alan J. and Kevin Hassett (1993), "Taxation and Foreign direct Investment in the United States: A Reconsideration of the Evidence", in: Alberto Giovannini, R. Glenn Hubbard, and Joel Slemrod (eds.) *Studies in International Taxation,* University of Chicago Press, 119-144.

[9] Blonigen, Bruce A. (1997), "Firm-Specific Assets and the Link between Exchange Rates and Foreign Direct Investment", *American Economic Review*, 87(3), 447-465.

[10] Bolton, P. and E. L. von Thadden (1998), "Blocks, Liquidity, and Corporate Control", *Journal of Finance*, 53, 1-25.

[11] Borenzstein, Eduardo, Jose De Gregorio, and J. W. Lee, (1998), "How Does Foreign Direct Investment Affect Economic Growth?", *Journal of International Economics*, 45, 115-35.

[12] Bosworth, Barry, and Susan Collins (1999), "Capital Flows to Developing Economies: Implications for Saving and Investment", *Brookings Papers on Economic Activity*, 1, 143-69.

[13] Brainard, S. Lael (1997), "An Empirical Assessment of the Proximity-Concentration Trade-off Between Multinational Sales and Trade", *American Economic Review,* 87(4), 520-544.

[14] Caballero, Ricardo and Eduardo Engel (1999), "Explaining Investment Dynamics in US Manufacturing: A Generalized (S;s) Approach", *Econometrica*, July, 741-82.

[15] ________________ (2000), "Lumpy Adjustment and Aggregate Investment Equations: A "Simple" Approach Relying on Cash Flow "Information", mimeo.

[16] Carr, David L., James R. Markusen and Keith E. Maskus (2001), "Estimating the Knowledge-Capital Model of the Multinational Enterprise," *American Economic Review*, 91(3), 693-708.

[17] Caves, Richard E. (1971), "International Corporations: The Industrial Economics of Foreign Investment", *Economica*, 38, 1-27.

[18] Chan, L. and J. Lakonishok (1995), "The Behavior of Stock Prices around Institutional Trades", *Journal of Finance,* 50, 1147-1174.

[19] Chari, A., P. P. Ouimet and L. Tesar (2005), "Acquiring Control in Emerging Markets: Evidence from the Stock Market", Working Paper, University of Michigan.

[20] Chuhan, P., G. Perez-Quiros and H. Popper (1996), "Do Short-term Investment and Direct Investment Differ?", World Bank Policy Research Working Paper No. 1669.

[21] Dadush Uri, Dipak Dasgupta, and Dilip Ratha (2000), "The Role of Short-Term Debt in Recent Crises", *Finance and Development*, 37, 54-57.

[22] De Mooij, R. A., and S. Ederveen (2001), "Taxation and Foreign Direct Investment: A Synthesis of Empirical Research", CESifo Working Paper No. 588.

[23] Desai, Mihir A. and James R. Hines (2001), "Foreign Direct Investment in a World of Multiple Taxes", University of Michigan.

[24] Devereux, Michael P. and R. Glenn Hubbard (2003), "Taxing Multinationals", *International Tax and Public Finance*, 10(4), 469-488.

[25] Devereux, Michael P. and Rachel Griffith (2003), "Evaluating Tax Policy for Location Decisions", *International Tax and Public Finance*, 10(2), 107-126.

[26] Diamond, Peter A., and James Mirrlees (1971), "Optimal Taxation and Public Production: I and II," *American Economic Review* (March and June), 8-17 and 261-278.

[27] Diamond, D. W. and P. H. Dybvig (1983), "Bank Runs, Deposit Insurance, and Liquidity", *Journal of Political Economy*, 91, 401-419.

[28] Djankov, Simeon, and Bernard Hoekman (2000), "Foreign Investment and Productivity Growth in Czech Enterprises", *The World Bank Economic Review*, 14 (1), 49-64.

[29] Eaton, Jonathan, and Akiko Tamura (1994), "Bi-lateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment," *The Journal of The Japanese and International Economies*, 8, 478-510.

[30] Fedelstein, Martin (2000), "Aspects of Global Economic Integration: Outlook for the Future", NBER Working Paper no. 7899 (Cambridge, Massachusetts: National Bureau of economic Research).

[31] Feenstra, Robert C. (1999), "Facts and Fallacies About Foreign Direct Investment", in Martin Feldstein, ed.: *International Capital Flows*, (University of Chicago Press, Chicago, III).

[32] Feliciano Z. and Lipsey, Robert E. (1999), "Foreign Ownership and Wages in the United States, 1987-92", NBER Working Paper No. 6923.

[33] Foster, F. D. and S. Viswanathan (1990), "A Theory of the Interday Variations in Volume, Variance, and Trading Costs in Securities Markets", *Review of Financial Studies*, 3, 593-624.

[34] Frankel, J. A. and A. K. Rose (1996), "Currency Crashes in Emerging Markets: An Empirical Treatment", *Journal of International Economics*, 41, 351-366.

[35] Frenkel, Jacob A., Assaf Razin and Efraim Sadka (1991), *International Taxation in an Integrated World Economy*, MIT press.

[36] Froot, K. (1991), "Japanese Foreign Direct Investment", in Feldstein, Martin S. and Y. Kosai (eds.), *US-Japan Economic Forum.*

[37] Froot, K. A. and J. C. Stein (1992), Exchange Rates and Foreign Direct Investment: An Imperfect Capital Markets Approach", *Quarterly Journal of Economics,* 106(5), 1191-1217.

[38] Ghironi, Fabio and Marc J. Melitz (2004), "International Trade and Macroeconomic Dynamics with Heterogenous Firms," NBER Working Paper 10540.

[39] Goldstein, Itay and Assaf Razin (2006), "An information - based Trade-off between Foreign Direct Investment and Foreign Portfolio investment", *Journal of International Economics.*

[40] Gopinath, Gita (2004), "Lending Booms, Sharp Reversals and the Real Exchange Dynamics", *Journal of International Economics*, 62 (1), 1-23.

[41] Gordon, Ruger H. (1983), "An Optimal Taxation Approach to Fiscal Federalism", *Quarterly Journal of Economics*, 98, 567-586.

[42] Griffith, Rachel and Helen Simpson (2003), "Characteristics of Foreign-Owned Firms in British Manufacturing", NBER Working Paper No. 9573 (March).

[43] Gordon, R. H. and A. Lans Bovenberg (1996), "Why Is Capital So Immobile Internationally? Possible Explanations and Implications for Capital Income Taxation", *American Economic Review*, 86, 1057-1075.

[44] Hall, Robert E. and Dale W. Jorgenson (1967), "Pax Policy and Investment Behavior", *American Economic Review*, 57, 391-414.

[45] Hanson G. Mataloni R. and M. Slaughter (2001), "Expansion Strategies of U.S. Multinational Firms", in Dani Rodrick and Susan Collins (eds.), *Brooking Trade Forum* 2001, 245-282.

[46] Harberger, Arnold C. (1962), "The Incidence of The Corporation Income Tax", *Journal of Political Economy*, 70, 215-40.

[47] Hart, O. (2001), "Financial Contracting", *Journal of Economic Literature*, 34, 1079-1100.

[48] Haufler, Andreas (2001), *Taxation in a Global Economy*, Cambridge University Press.

[49] Heckman, James. J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 42, 153-168.

[50] Heckman, James. J. (1974), "Shadow Prices, Market Wages and Labor Supply", *Economtrica*, 47, 153-161.

[51] Helpman, Elhanan (2006), "Trade, FDI, and the Organization of Firms", NBER Working Paper No. 12091.

[52] Helpman, Elhanan, Mark Melitz, and Steve Yeaple (2004), ”Exports vs FDI”, *American Economic Review*.

[53] Helpman, Elhanan (1984), "A Simple Theory of International Trade with Multinational Corporations", *Journal of Political Economy*, 92, 451-471.

[54] Hennart, J., D. Kim and M. Zeng (1998), "The Impact of Joint Venture Status on the Longevity of Japanese Stakes in U.S. Manufacturing Affiliates", *Organization Science*, 9, 382-395.

[55] Hertzel, M. and R. Smith (1993), "Market Discounts and Shareholder Gains for Placing Equity Privately", *Journal of Finance*, 48, 459-485.

[56] Hines, James R. (1999), "Lessons from Behavioral Responses to International Taxation", *National Tax Journal*, 53, 305-22.

[57] Hines, James R. (2001), "Corporate Taxation", in Neil J. Smelser and Paul B. Balts (eds.), *International Encyclopaedia of the Social and Behavioral Sciences*.

[58] Holthausen, R. W., R. W. Leftwich and D. Mayers (1990), "Large-Block Transactions, the Speed of Response, and Temporary and Permanent Stock-Price Effects", *Journal of Financial Economics*, 26, 71-95.

[59] Jakubiak, Malgorzata and Malgorzata Markiewicz (2005), "Capital Mobility and Tax Competition in the EU after Enlargement", CASE (mimeo).

[60] Jensen, M. (1986), "The Agency Costs of Free Cash Flow: Corporate Finance and Takeovers", *American Economic Review*, 76, 323-330.

[61] Kahn, C. and A. Winton (1998), "Ownership Structure, Speculation, and Shareholder Intervention", *Journal of Finance*, 53, 99-129.

[62] Keane, Michael, Robert Moffit and David Runkle (1988), "Real Wages over the Business Cycle: Estimating the Impact of Heterogeneity with Micro Data", *Journal of Political Economy*, 96(6), 1232-66.

[63] Klein, M. W. and E. S. Rosengren (1994), "The Real Exchange Rate and Foreign Direct Investment in the United States", *Journal of International Economics*, 36, 373-389.

[64] Klein, M. W., J. Peek and E. S. Rosengren (2002), "Troubled Banks, Impaired Foreign Direct Investment: The Role of Relative Access to Credit", *American Economic Review*, 92, 664-682.

[65] Korajczyk, R. A., D. J. Lucas and R. L. McDonald (1991), "The Effect of Information Releases on the Pricing and Timing of Equity Issues", *Review of Financial Studies*, 4, 685-708.

[66] Krugman, Paul (2000), "Fise-Sale FDI", in Edwards, Sebastian (ed.), *Capital Flows and the Emerging Economies: Theory, Evidence, and Controversies*, University of Chicago Press.

[67] Kyriazidou, Ekaterini (1996), "Estimation of a Panel Data Sample Selection Model," *Econometrica,* 65, 1335-1364.

[68] La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. W. Vishny (1998), "Law and Finance", *Journal of Political economy*, 106, 1113-1155.

[69] Lipsey, Robert E. (2000), "The Role of Foreign Direct Investment in International Capital Flows", NBER Working Paper # 7094.

[70] Lipsey, Robert E. (2000/2001), "Foreign Direct Investment and the Operations of Multinational Firms", *NBER Reporter*, winter, NBER, Cambridge, USA.

[71] Lipsey, Robert E. (2001), "Foreign Direct Investors in Three Financial Crises", NBER Working Paper # 8084.

[72] Lipsey, Robert E. (2002), "Home and Host Country Effects of FDI", NBER Working Paper # 9293.

[73] Loungani, Prakash, and Assaf Razin (2001), "How Beneficial Is Foreign Direct Investment for Developing Countries", *Finance and Development*, 38(2).

[74] Lucas, Robert E. (1990), "Why Doesn't Capital Flow from Rich to Poor Countries?", *American Economic Review: Papers and Proceedings*, 80(2), 92-96.

[75] Markusen, Lames R. (2002), *Multinational Firms and the Theory of International Trade*, Cambridge, MIT Press.

[76] Maug, E. (1998), "Large Shareholders as Monitors: Is There a Trade off Between Liquidity and Control?", *Journal of Finance*, 53, 65-98.

[77] Mikkelson, W. H. and M. Partch (1985), "Stock Price Effects and Costs of Secondary Distributions", *Journal of Financial Economics*, 14, 165-194.

[78] Melitz, Marc, J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695-1725.

[79] Mintz, Jack M. and Henry Tulkens (1986), "Commodity Tax Competition between Member States of a Federation: Equilibrium and Efficiency", *Journal of Public Economics*, 29, 133-172.

[80] Obstfeld, M. and A. M. Taylor (2002), "Globalization and the Capital Markets", NBER Working Paper No. 8846.

[81] OECD (1997), *Model Tax Convention on Income and on Capital*, Paris, OECD Committee on fiscal Affairs.

[82] Perez-Gonzalez, F. (2005), "The Impact of Acquiring Control on Productivity", Working Paper, Columbia University.

[83] Razin, Assaf (2004), "The Contribution of FDI Flows to Domestic Investment in Capacity, and Vice Versa", in Ito, Takatoshi, and Andrew K. Rose (eds.), *Growth and Productivity in East Asia*, NBER-East Asia Seminar on Economics, Vol. 13, 149-167.

[84] Razin, Assaf, and Efraim Sadka (2003), "Gains from FDI Inflows with Incomplete Information", *Economics Letters,* 7, 71-77.

[85] Razin, Assaf, and Efraim Sadka (2003), *Labor, Capital and Finance: International Flows* (New York and Cambridge, England: Cambridge University Press).

[86] Razin, Assaf and Efraim Sadka (forthcoming), "Corporate Transparency, Cream-Skimming and FDI", *European Economic Review.*

[87] Razin, Assaf, Yona Rubinstein and Efraim Sadka (2004), "Fixed Costs and FDI: The Conflicting Effects of Productivity Shocks", NBER Working Paper No. 10864.

[88] Razin, Assaf, Yona Rubinstein and Efraim Sadka (2005), "Corporate Taxation and Bilateral FDI with Threshold Barriers", NBER Working Paper 11196.

[89] Razin, Assaf, Efraim Sadka, and Tarek Coury (2003), "Trade Openness, Investment Instability, and Terms-of-Trade Volatility," *Journal of International Economics*, 59 (2).

[90] Razin, Assaf, Efraim Sadka and Hui Tong (2005), "Bilateral FDI flows: Threshold Barriers and Productivity Shocks", NBER Working Paper 11639.

[91] Razin, Assaf, Efraim Sadka, and Chi-Wa Yuen (1998), “A Pecking Order Theory of Capital Flows and International Tax Principles”, *Journal of International Economics*, 44, 45-68.

[92] Santos, Joao Silva, and Silvana Tenreyro (2003), "Gravity-Defying Trade," Working Papers 03-1, Federal Reserve Bank of Boston.

[93] Sarno, L. and M. P. Taylor (1999), "Hot Money, Accounting Labels and the Permanence of Capital Flows to Developing Countries: An Empirical Investigation", *Journal of Development Economics*, 59, 337-364.

[94] Shleifer, A. and R. W. Vishny (1986), Large Shareholders and Corporate Control", *Journal of Political Economy*, 94, 461-488.

[95] Tinbergen, Jan (1962), "*Sharing the World Economy*", New York, NY: Twentieth Century Fund.

[96] Zhang, Kevin H. and James R. Markusen (1999), "Vertical Multinationals and Host-Country Characteristics," *Journal of Development Economics*, 59 (2).

[97] U.N. (1980), "U.N. Model Double Taxation Convention between Developed and Developing Countries", U.N. Document # ST/ESA/102.

[98] U.N. (2005), *World Investment Report*, U.N.: New York and Geneva.

[99] Wei, Shang-Jin (2000), "How Taxing is Corruption on International Investors", *Review of Economics and Statistics*, 1, 1-11.

[100] Wildasin, David E. (2004), "Competitive Fiscal Structures", University of Kentucky (mimeo).

[101] Wilson, John D. (1999), "Theories of Tax Competition", *National Tax Journal*, 52, 269-304.

[102] Wruck, K. (1989), Equity Ownership Concentration and firm Value: Evidence from Private Equity Financing", *Journal of Financial economics*, 23, 3-28.

[103] Zodrow, George R. and Peter Mieszkowski (1986), "Pigon, Tiebout, Property Taxation, and the Underprovision of Local Public Goods", Journal of Urban Economics, 19, 356-370.