# Firms, Contracts, and Trade Structure

by

## Pol Antràs

B.A., Universitat Pompeu Fabra (1998)
M.Sc., Universitat Pompeu Fabra (1999)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2003

© Pol Antràs, MMIII. All rights reserved.

The author hereby grants to Massachusetts Institute of Technology permission to reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Department of Economics
May 2, 2003

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Daron K. Acemoglu
Professor of Economics
Thesis Supervisor

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Jaume Ventura
Pentti J. K. Kouri Associate Professor of Economics
Thesis Supervisor

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Peter Temin
Elisha Gray II Professor of Economics
Chairman, Departmental Committee on Graduate Studies

# Firms, Contracts, and Trade Structure

by

## Pol Antràs

Submitted to the Department of Economics
on May 2, 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

## Abstract

This dissertation consists of three essays in the intersection of the theory of international trade and the theory of the firm.

The first essay starts by unveiling two systematic patterns in the volume of intrafirm trade. I then show that these patterns can be rationalized in a theoretical framework that combines a Grossman-Hart-Moore view of the firm with a Helpman-Krugman view of international trade. In particular, I develop an incomplete-contracting, property-rights model of the boundaries of the firm, which I then incorporate into a standard trade model with imperfect competition and product differentiation. The model pins down the boundaries of multinational firms as well as the international location of production. Econometric evidence reveals that the model is consistent with other qualitative and quantitative features of the data.

In the second essay, I develop a dynamic, general-equilibrium Ricardian model of North-South trade, in which the incomplete nature of contracts governing international transactions leads to the emergence of product cycles. Following the property-rights approach to the theory of the firm, the same force that creates product cycles, i.e. incomplete contracts, opens the door to a parallel analysis of the determinants of the mode of organization. The model is shown to deliver endogenous organizational cycles. I discuss several macroeconomic and microeconomic implications of the model and relate them to the previous literature on the product cycle.

The third essay, co-authored with Professor Helpman from Harvard University, provides a theoretical framework for studying global sourcing strategies. In our model, heterogeneous final-good producers choose organizational forms. That is, they choose ownership structures and locations for the production of intermediate inputs. We describe an equilibrium in which firms with different productivity levels choose different ownership structures and supplier locations. We then study the effects of within-sectoral heterogeneity and variations in industry characteristics on the relative prevalence of these organizational forms. The analysis sheds light on the structure of foreign trade within and across industries.

Thesis Supervisor: Daron K. Acemoglu
Title: Professor of Economics

Thesis Supervisor: Jaume Ventura
Title: Pentti J. K. Kouri Associate Professor of Economics

# Acknowledgments

This thesis would have never been written without the love and continuous support of my family. Their intellectual breadth and curiosity motivated me to pursue a Ph.D. Being away from them was the toughest part of my experience at MIT.

I am also grateful to my professors at Universitat Pompeu Fabra in Barcelona for teaching me to think as an economist and, more importantly, for encouraging me to keep learning about economics. In particular, I thank Jan Eeckhout, Andreu Mas-Colell, Jordi Galí, Gilles Saint-Paul, and Xavier Sala-i-Martin for their help in finding my way to MIT.

At MIT, I have been incredibly fortunate to work under the supervision of Daron Acemoglu and Jaume Ventura. Daron was encouraging and supportive since my first day at MIT. His comments and suggestions were always instrumental in transforming my random thoughts into research projects and, later, into actual research papers. He amazes me day after day and I look forward to keep learning from him in the future. Jaume has been a wonderful mentor and a friend. His constant faith in me and insightful advice literally changed the course of my research and life. He will always be my role model for an advisor. Also at MIT, Marios Angeletos and Bengt Holmstrom provided excellent suggestions that enormously improved this dissertation. I am also grateful to them for their support throughout the job-market process. Special thanks also go to Gary King and Katherine Swan for helping me navigate the sometimes intricate administrative requirements at MIT. During my stay at MIT, I have also benefited from financial support from the Bank of Spain.

In the last few months, I have had the privilege of interacting with the two leading figures in the field of international trade. Gene Grossman kindly invited me to spend

part of the fall term of 2002 in Princeton University. He provided me with an apartment, an office, and an incredibly stimulating intellectual environment. The second essay in this dissertation was written during my stay at Princeton. Gene also tiredlessly read the bulk of this dissertation and provided most helpful comments. I feel deeply indebted to him. Elhanan Helpman from Harvard University also offered excellent suggestions on the first two essays in this dissertation and co-authored the third one with me. Working with him was an extraordinary experience and I look forward to learning more from him over the next few years.

My classmates and friends in Cambridge were constant sources of laughter, joy, and support. Special thanks go to Alexis León Dimas, Gerard Padró Miquel, and Pau Casas Arce who made life in West Somerville infinitely more pleasant.

And finally, I want to thank my dearest one, Lucia Breierova, who has given me strength and love throughout the years of this project.

*To my family, who made this possible,*
*and to Lucia, who made it happen.*

# Contents

---

[1]**This chapter is joint work with Professor Helpman from Harvard University.**

# Chapter 1

# Introduction

In recent years, we have witnessed a spectacular increase in the way firms organize production on a global scale. Feenstra (1998), citing Tempest (1996), describes Mattel's global sourcing strategies in the manufacturing of its star product, the Barbie doll:

> The raw materials for the doll (plastic and hair) are obtained from Taiwan and Japan. Assembly used to be done in those countries, as well as the Philippines, but it has now migrated to lower-cost locations in Indonesia, Malaysia, and China. The molds themselves come from the United States, as do additional paints used in decorating the dolls. Other than labor, China supplies only the cotton cloth used for dresses. Of the $2 export value for the dolls when they leave Hong Kong for the United States, about 35 cents covers Chinese labor, 65 cents covers the cost of materials, and the remainder covers transportation and overheads, including profits earned in Hong Kong. (Feenstra, 1998, p. 35-36).

A variety of terms have been used to refer to this phenomenon: the "slicing of the value chain", "international outsourcing", "fragmentation of the production process", "vertical specialization", "global production sharing", and many more.

In developing their global sourcing strategies, firms not only decide on where to *locate* the different stages of the value chain, but also on the extent of *control* they want to exert over these processes. This is the classical "make-or-buy" decision in industrial organization. Firms may decide to keep the production of intermediate inputs within

13

firm boundaries, thus engaging in foreign direct investment when the integrated supplier is in a foreign country, or they may choose to contract with arm's length suppliers for the procurement of these components. An example of the former is Intel Corporation, which assembles most of its microchips in wholly-owned subsidiaries in China, Costa Rica, Malaysia, and Philippines. Conversely, Nike subcontracts most of the manufacturing of its products to independent producers in Thailand, Indonesia, Cambodia, and Vietnam, while keeping within firm boundaries the design and marketing stages of production.

This dissertation consists of three essays that attempt to shed some light on how firms organize production in the global economy. On the empirical side, the essays discuss some systematic patterns in the way production is organized in the world economy. On the theoretical side, I build on workhorse models in international trade and the theory of the firm to develop simple models that can account for these important facts.

The first essay starts by unveiling two systematic patterns in the volume of intrafirm trade, which constitutes roughly one-third of world trade. In a panel of industries, the share of intrafirm imports in total U.S. imports is significantly higher, the higher the capital intensity in production of the exporting industry. In a cross-section of countries, the share of intrafirm imports in total U.S. imports is significantly higher, the higher the aggregate capital-labor ratio of the exporting country. I then show that these patterns can be rationalized in a theoretical framework that combines a Grossman-Hart-Moore view of the firm with a Helpman-Krugman view of international trade. In particular, I develop an incomplete-contracting, property-rights model of the boundaries of the firm, which I then incorporate into a standard trade model with imperfect competition and product differentiation. The model pins down the boundaries of multinational firms as well as the international location of production, and it is shown to predict the patterns of intrafirm trade identified above. Econometric evidence reveals that the model is consistent with other qualitative and quantitative features of the data.

The second essay attempts to provide a new theory of the product cycle based on the notion that the incomplete nature of contracts governing international transactions limits the extent to which the production process can be fragmented across borders. In a dynamic, general-equilibrium Ricardian model of North-South trade, contract incompleteness is shown to lead to the emergence of product cycles. Goods are initially

manufactured in the North, where product development takes place. As the good matures and becomes more standardized, manufacturing is shifted to the South to benefit from lower wages. Following the property-rights approach to the theory of the firm, the same force that creates product cycles, i.e. incomplete contracts, opens the door to a parallel analysis of the determinants of the mode of organization. As a result, a new version of the product cycle emerges, in which manufacturing is shifted to the South first within firm boundaries, and only at a later stage to independent firms in the South. Relative to a world with only arm's length transacting, allowing for intrafirm production transfer by multinational firms is shown to accelerate the shift of production towards the South, while having an ambiguous effect on relative wages. The model delivers macroeconomic implications that complement the work of Krugman (1979) and Helpman (1993). I also discuss several microeconomic implications of the model and show how they are consistent with the empirical literature on the product cycle.

The third essay, co-authored with Professor Helpman from Harvard University, provides a theoretical framework for studying global sourcing strategies. In our model, heterogeneous final-good producers develop differentiated products in the North and choose organizational forms. That is, they choose ownership structures and locations for the production of intermediate inputs. A firm that chooses to keep the production of an intermediate input within its boundaries can produce it at home or in a foreign country. When it keeps it at home, it engages in standard vertical integration. And when it makes it abroad, it engages in foreign direct investment (FDI) and intra-firm trade. Alternatively, a firm may choose to outsource an input in the home country or in a foreign country. When it buys the input at home, it engages in domestic outsourcing. And when it buys it abroad, it engages in arm's-length trade. We describe an equilibrium in which firms with different productivity levels choose different ownership structures and supplier locations. We then study the effects of within-sectoral heterogeneity and variations in industry characteristics on the relative prevalence of these organizational forms. The analysis sheds light on the structure of foreign trade within and across industries. The model can be used to investigate the implications of falling trade costs for the relative prevalence of domestic vertical integration over domestic outsourcing, as well as of foreign direct investment over foreign outsourcing. We provide suggestive evidence of trends in

the data that are consistent with the predictions of our model.

While the three essays are closely related, they have been written so that each may be read independently of the others.

# Chapter 2

# Firms, Contracts, and Trade Structure

## 2.1  Introduction

Roughly one-third of world trade is intrafirm trade. In 1994, 42.7 percent of the total volume of U.S. imports of goods took place within the boundaries of multinational firms, with the share being 36.3 percent for U.S. exports of goods (Zeile, 1997). In spite of the clear significance of these international flows of goods between affiliated units of multi-national firms, the available empirical studies on intrafirm trade provide little guidance to international trade theorists. In this paper I unveil some novel patterns exhibited by the volume of U.S. intrafirm imports and I argue that these patterns can be rationalized combining a Grossman-Hart-Moore view of the firm, together with a Helpman-Krugman view of international trade.

In a hypothetical world in which firm boundaries had no bearing on the pattern of international trade, one would expect only random differences between the behavior of the volume of intrafirm trade and that of the total volume of trade. In particular, the share of intrafirm trade in total trade would not be expected to correlate significantly with any of the classical determinants of international trade.

Figure 2-1 provides a first illustration of how different the real world is from this

17

Figure 2-1: Share of Intrafirm U.S. Imports and Relative Factor Intensities
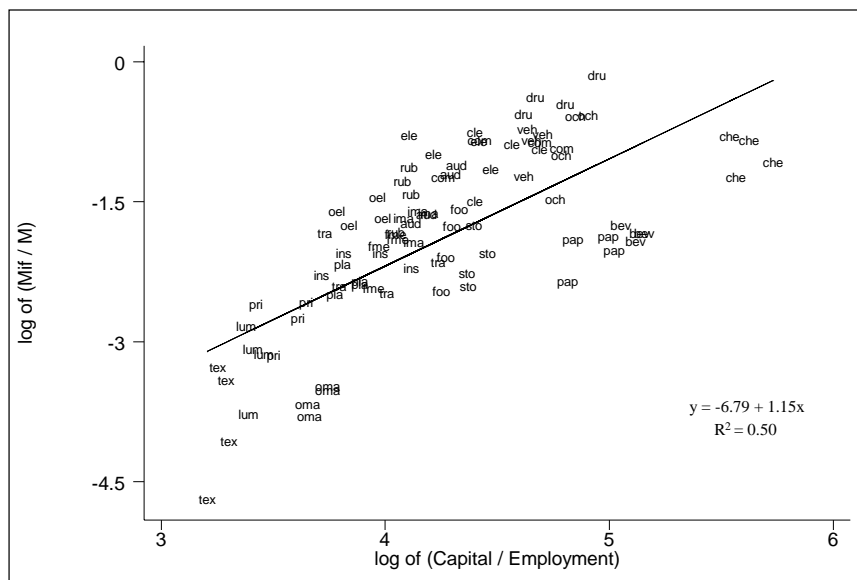


**Notes:** The Y-axis corresponds to the logarithm of the share of intrafirm imports in total U.S. imports for 23 manufacturing industries and 4 years: 1987, 1989, 1992, 1994. The X-axis measures the log of that industry's ratio of capital stock to total employment in the corresponding year, using U.S. data. See Table A.1. for industry codes and Appendix A.4. for data sources.

Figure 2-2: Share of Intrafirm U.S. Imports and Relative Factor Endowments
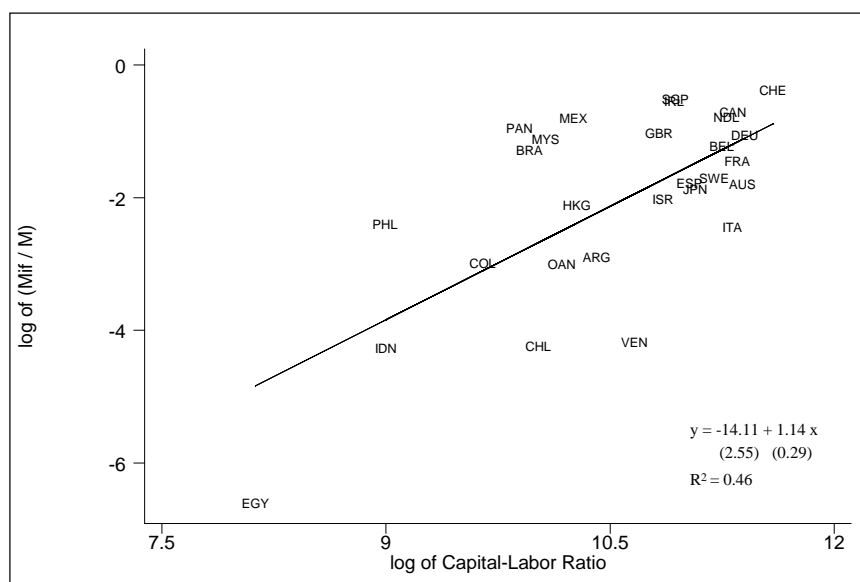


**Notes:** The Y-axis corresponds to the logarithm of the share of intrafirm imports in total U.S. imports for 28 exporting countries in 1992. The X-axis measures the log of the exporting country's physical capital stock divided by its total number of workers. See Table A.2. for country codes and Appendix A.4. for details on data sources.

18

hypothetical world. In a panel consisting of 23 manufacturing industries and four years of data (1987, 1989, 1992, and 1994), the share of intrafirm imports in total U.S. imports is significantly higher, the higher the capital intensity in production of the exporting industry. Figure 2-1 indicates that firms in the U.S. tend to import capital-intensive goods, such as chemical products, within the boundaries of their firms, while they tend to import labor-intensive goods, such as textile products, from unaffiliated parties.[1]

Figure 2-2 unveils a second strong pattern in the share of intrafirm imports. In a cross-section of 28 countries, the share of intrafirm imports in total U.S. imports is significantly higher, the higher the capital-labor ratio of the exporting country. U.S. imports from capital-abundant countries, such as Switzerland, tend to take place between affiliated units of multinational firms. Conversely, U.S. imports from capital-scarce countries, such as Egypt, occur mostly at arm's length. This second fact suggests that the well-known predominance of North-North trade in total world trade is even more pronounced within the intrafirm component of trade.[2]

Why are capital-intensive goods transacted within the boundaries of multinational firms, while labor-intensive goods are traded at arm's length?[3] To answer this question, I build on the theory of the firm initially exposited in Coase (1937) and later developed by Williamson (1985) and Grossman and Hart (1986), by which activities take place wherever transaction costs are minimized. In particular, I develop a property-rights model of the boundaries of the firm in which, in equilibrium, transaction costs of using the market are increasing in the capital intensity of the imported good. To explain the cross-country pattern in Figure 2-2, I embed this partial-equilibrium framework in a general-

---

[1] The pattern in Figure 1 is consistent with Gereffi's (1999) distinction between 'producer-driven' and 'buyer-driven' international economic networks. The first, he writes, is "characteristic of *capital*- and technology-*intensive* industries [...] in which large, usually *transnational*, manufacturers play the central roles in coordinating production networks" (p. 41). Conversely, 'buyer-driven' networks are common in "*labor-intensive*, consumer goods industries" and are characterized by "highly competitive, *locally owned*, and globally dispersed production systems" (pp. 42-43). The italics are my own.

[2] This is consistent with comparisons based on foreign direct investment (FDI) data. In the year 2000, more than 85% of FDI flows occured between developed countries (UNCTAD, 2001), while the share of North-North trade in total world trade was only roughly 70% (World Trade Organization, 2001).

[3] At this point, a natural question is whether capital intensity and capital abundance are truly the crucial factors driving the correlations in Figures 2-1 and 2-2. In particular, these patterns could in principle be driven by other omitted factors. Section 2.4 will present formal econometric evidence in favor of the emphasis placed on capital intensity and capital abundance in this paper.

equilibrium, factor-proportions model of international trade, with imperfect competition and product differentiation, along the lines of Helpman and Krugman (1985). The model pins down the boundaries of multinational firms as well as the international location of production. Bilateral trade flows between any two countries are uniquely determined and the implied relationship between intrafirm trade and relative factor endowments is shown to correspond to that in Figure 2-2. The result naturally follows from the interaction of comparative advantage and transaction-cost minimization.

In drawing firm boundaries, I build on the seminal work of Grossman and Hart (1986). I consider a world of incomplete contracts in which ownership corresponds to the entitlement of some residual rights of control. When parties undertake noncontractible, relationship-specific investments, the allocation of residual rights has a critical effect on each party's *ex-post* outside option, which in turn determines each party's *ex-ante* incentives to invest. Ex-ante efficiency (i.e., transaction-cost minimization) then dictates that residual rights should be controlled by the party whose investment contributes most to the value of the relationship.

To explain the higher propensity to integrate in capital-intensive industries, I extend the framework of Grossman and Hart (1986) by allowing the transferability of certain investment decisions. In situations in which the default option for one of the parties (a supplier in the model) is too unfavorable, the allocation of residual rights may not suffice to induce adequate levels of investment. In such situations, I show that the hold-up problem faced by the party with the weaker bargaining position may be alleviated by having another party (a final-good producer in the model) contribute to the former's relationship-specific investments. Investment-sharing alleviates the hold-up problem faced by suppliers, but naturally increases the exposure of final-good producers to opportunistic behavior, with the exposure being an increasing function of the contribution to investment costs. If cost sharing is large enough, ex-ante efficiency is shown to command that residual rights of control, and thus ownership, be assigned to the final-good producer, thus giving rise to vertical integration. Conversely, when the contribution of the final-good producer is relatively minor, the model predicts outsourcing.

What determines then the extent of cost sharing? Business practices suggest that, in many situations, investments in physical capital are easier to share than investments in

labor input. Dunning (1993, p. 455-456) describes several cost-sharing practices of multi-national firms in their relations with independent subcontractors. Among others, these include provision of used machinery and specialized tools and equipment, prefinancing of machinery and tools, and procurement assistance in obtaining capital equipment and raw materials. There is no reference to cost sharing in labor costs, other than in labor training. Milgrom and Roberts (1993) discuss the particular example of General Motors, which pays for firm- and product-specific capital equipment needed by their suppliers, even when this equipment is located in the suppliers' facilities. Similarly, in his review article on Japanese firms, Aoki (1990, p. 25) describes the close connections between final-good manufacturers and their suppliers but writes that "suppliers have considerable autonomy in other respects, for example in personnel administration". Even within firm boundaries, cost sharing seems to mostly take place when capital investments are involved. In particular, Table 2.1 indicates that British affiliates of U.S.-based multinationals tend to have much more independence in their employment decisions (e.g., in hiring of workers) than in their financial decisions (e.g., in their choice of capital investment projects).

Table 2.1. Decision-Making in U.S. based multinationals

| % of British affiliates in which parent influence on decision is strong or decisive | | | |
|---|---|---|---|
| Financial decisions | | Employment/personnel decisions | |
| Setting of financial targets | 51 | Union recognition | 4 |
| Preparation of yearly budget | 20 | Collective bargaining | 1 |
| Acquisition of funds for working capital | 44 | Wage increases | 8 |
| Choice of capital investment projects | 33 | Numbers employed | 13 |
| Financing of investment projects | 46 | Lay-offs/redundancies | 10 |
| Target rate of return on investment | 68 | Hiring of workers | 10 |
| Sale of fixed assets | 30 | Recruitment of executives | 16 |
| Dividend policy | 82 | Recruitment of senior managers | 13 |
| Royalty payments to parent company | 82 | | |

Source: Dunning (1993, p. 227). Originally from Young, Hood and Hamill (1985).

In this paper, I do not intend to explain why cost sharing is more significant in physical capital investments than in labor input investments. This may be the result of suppliers having superior local knowledge in hiring workers, or it may be explained by the fact that managing workers requires a physical presence in the production plant. Regardless of the source of this asymmetry, the model developed in section 2.2 shows that if cost sharing is indeed more significant in capital-intensive industries, the propensity to integrate will also be higher in these industries. In order to explain the trade patterns shown in Figures 2-1 and 2-2, I then embed the partial-equilibrium relationship between final-good producers and suppliers into a general-equilibrium framework with a continuum of goods in each of two industries. In section 2.3, I open this economy to international trade, allowing final-good producers to obtain intermediate inputs from foreign suppliers. In doing so, I embrace a Helpman-Krugman view of international trade with imperfect competition and product differentiation, by which countries specialize in producing certain varieties of intermediate inputs and export them worldwide. Trade in capital-intensive intermediate inputs will be transacted within firm boundaries. Trade in labor-intensive goods will instead take place at arm's length. The model solves for bilateral trade flows between any two countries, and predicts the share of intrafirm imports in total imports to be increasing in the capital-labor ratio of the exporting country.[4] This is the correlation implied by Figure 2-2. Moreover, some of the quantitative implications of the model are successfully tested in section 2.4.

This paper is related to several branches in the literature. On the one hand, it is related to previous theoretical studies that have rationalized the existence of multinational firms in general-equilibrium models of international trade.[5] Helpman's (1984) model introduced a distinction between firm-level and plant-level economies of scale that has proven crucial in later work. In his model, multinationals arise only outside the factor price equalization set, when a firm has an incentive to geographically separate

---

[4]This second part of the argument is based on the premise that capital-abundant countries tend to produce mostly capital-intensive commodities. In an important contribution, Romalis (2002) has recently shown that the empirical evidence is indeed consistent with factor proportions being a key determinant of the structure of international trade.

[5]The literature builds on the seminal work of Helpman (1984) and Markusen (1984). For extensive reviews see Caves (1996) and Markusen and Maskus (2001).

the capital-intensive production of an intangible asset (headquarter services) from the more labor-intensive production of goods. Following the work of Markusen (1984) and Brainard (1997), an alternative branch of the literature has developed models rationalizing the emergence of multinational firms in the absence of factor endowment differences. In these models, multinationals will exist in equilibrium whenever transport costs are high and whenever firm-specific economies of scale are high relative to plant-specific economies of scale.[6,7]

These two approaches to the multinational firm share a common failure to properly model the crucial issue of internalization. These models can explain why a domestic firm might have an incentive to undertake part of its production process abroad, but they fail to explain why this foreign production will occur within firm boundaries (i.e., within multinationals), rather than through arm's length subcontracting or licensing. In the same way that a theory of the firm based purely on technological considerations does not constitute a satisfactory theory of the firm (c.f., Tirole, 1988, Hart, 1995), a theory of the multinational firm based solely on economies of scale and transport costs cannot be satisfactory either. As described above, I will instead set forth a purely organizational, property-rights model of the multinational firm. My model will make no distinction between firm-specific and plant-specific economies of scale. Furthermore, trade will be costless and factor prices will not differ across countries. Yet multinationals will emerge in equilibrium, and their implied intrafirm trade flows will match the strong patterns identified above.

This paper is also related to previous attempts to model the internalization decision of multinationals firms. Following the insights from the seminal work of Casson (1979), Rugman (1981) and others, this literature has constructed models studying the role of informational asymmetries and knowledge non-excludability in determining the choice

---

[6]The intuition for this result is straightforward: when firm-specific economies of scale are important, costs are minimized by undertaking all production within a single firm. If transport costs are high and plant-specific economies of scale are small, then it will be profitable to set up multiple production plants to service the different local markets. Multinationals are thus of the "horizontal type".

[7]Recently, the literature seems to have converged to a "unified" view of the multinational firm, merging the factor-proportions (or "vertical") approach of Helpman (1984), together with the "proximity-concentration" trade-off implicit in Brainard (1997) and others. Markusen and Maskus (2001) refer to this approach as the "Knowledge-Capital Model" and claim that its predictions are widely supported by the evidence.

between direct investment and licensing (e.g., Ethier, 1986, Ethier and Markusen, 1996). Among other things, this paper differs from this literature in stressing the importance of capital intensity and the allocation of residual rights in the internalization decision, and perhaps more importantly, in describing and testing the implications of such a decision for the pattern of intrafirm trade.

Finally, this paper is also related to an emerging literature on general-equilibrium models of industry structure (e.g., McLaren, 2000, Grossman and Helpman, 2002a). My theoretical framework shares some features with the recent contribution by Grossman and Helpman. In their model, however, the costs of transacting inside the firm are introduced by having integrated suppliers incur exogenously higher variable costs (as in Williamson, 1985). More importantly, theirs is a closed-economy model and therefore does not consider international trade in goods, which of course is central in my contribution.[8]

The rest of the paper is organized as follows. Section 2 describes the closed-economy version of the model and studies the role of factor intensity in determining the equilibrium mode of organization in a given industry. Section 3 describes the multi-country version of the model and discusses the international location of production as well as the implied patterns of intrafirm trade. Section 4 presents econometric evidence supporting the view that both capital intensity and capital abundance are significant factors in explaining the pattern of intrafirm U.S. imports. Section 5 concludes. The proofs of the main results are relegated to the Appendix.

## 2.2 The Closed-Economy Model: Ownership and Capital Intensity

This section describes the closed-economy version of the model. In section 2.3 below, I will reinterpret the equilibrium of this closed economy as that of an integrated world

---

[8]Although in this paper I show that a Grossman-Hart-Moore view of the firm is consistent with the facts in Figures 2-1 and 2-2, neither my theoretical model nor the available empirical evidence is rich enough to test this view of the firm against alternative ones. This would be a major undertaking on its own. See Baker and Hubbard (2002) and Whinston (2002) for more formal treatments of these issues.

economy. The features of this equilibrium will then be used to analyze the patterns of specialization and trade in a world in which the endowments of the integrated economy are divided up among countries.

## 2.2.1 Set-up

**Environment**   Consider a closed economy that employs two factors of production, capital and labor, to produce a continuum of varieties in two sectors, $Y$ and $Z$. Capital and labor are inelastically supplied and freely mobile across sectors. The economy is inhabited by a unit measure of identical consumers that view the varieties in each industry as differentiated. In particular, letting $y(i)$ and $z(i)$ be consumption of variety $i$ in sectors $Y$ and $Z$, preferences of the representative consumer are of the form

$$U = \left( \int_0^{n_Y} y(i)^\alpha di \right)^{\frac{\mu}{\alpha}} \left( \int_0^{n_Z} z(i)^\alpha di \right)^{\frac{1-\mu}{\alpha}} , \tag{2.1}$$

where $n_Y$ $(n_Z)$ is the endogenously determined measure of varieties in industry $Y$ $(Z)$. Consumers allocate a constant share $\mu \in (0,1)$ of their spending in sector $Y$ and a share $1 - \mu$ in sector $Z$. The elasticity of substitution between any two varieties in a given sector, $1/(1-\alpha)$, is assumed to be greater than one.

**Technology**   Goods are also differentiated in the eyes of producers. In particular, each variety $y(i)$ requires a special and distinct intermediate input which I denote by $x_Y(i)$. Similarly, in sector $Z$, each variety $z(i)$ requires a distinct component $x_Z(i)$. The specialized intermediate input must be of high quality, otherwise the output of the final good is zero. If the input is of high quality, production of the final good requires no further costs and $y(i) = x_Y(i)$ (or $z(i) = x_Z(i)$ in sector $Z$).

Production of a high-quality intermediate input requires capital and labor. For simplicity, technology is assumed to be Cobb-Douglas:

$$x_k(i) = \left( \frac{K_{x,k}(i)}{\beta_k} \right)^{\beta_k} \left( \frac{L_{x,k}(i)}{1-\beta_k} \right)^{1-\beta_k} , \quad k \in \{Y, Z\} \tag{2.2}$$

where $K_{x,k}(i)$ and $L_{x,k}(i)$ denote the amount of capital and labor employed in production of variety $i$ in industry $k \in \{Y, Z\}$. I assume that industry $Y$ is more capital-intensive than industry $Z$, i.e. $1 \geq \beta_Y > \beta_Z \geq 0$.

Low-quality intermediate inputs can be produced at a negligible cost in both sectors.

There are also fixed costs associated with the production of an intermediate input. For simplicity, it is assumed that fixed costs in each industry have the same factor intensity as variable costs, so that the total cost functions are homothetic. In particular, fixed costs for each variety in industry $k \in \{Y, Z\}$ are $f r^{\beta_k} w^{1-\beta_k}$, where $r$ is the rental rate of capital and $w$ the wage rate.

**Firm structure**   There are two types of producers: final-good producers and suppliers of intermediate inputs. Before any investment is made, a final-good producer decides whether it wants to enter a given market, and if so, whether to obtain the component from a vertically integrated supplier or from a stand-alone supplier. An integrated supplier is just a division of the final-good producer and thus has no control rights over the amount of input produced. Figuratively, at any point in time the parent firm could selectively fire the manager of the supplying division and seize production. Conversely, a stand-alone supplier does indeed have these residual rights of control. In Hart and Moore's (1990) words, in such a case the final-good producer could only "fire" the entire supplying firm, including its production. Integrated and non-integrated suppliers differ only in the residual rights they are entitled to, and in particular both have access to the same technology as specified in (2.2).[9]

As discussed in the introduction, a premise of this paper is that investments in physical capital are easier to share than investments in labor input. To capture this idea, I assume that while the labor input is necessarily provided by the supplier, capital expenditures $r K_{x,k}(i)$ are instead transferable, in the sense that the final-good producer can decide whether to let the supplier incur this factor cost too, or rather rent the capital itself and hand it to the supplier at no charge.[10] Irrespective of who bears their cost, the investments

---

[9]This is in contrast with the transaction-cost literature that usually assumes that integration leads to an exogenous increase in variable costs (e.g. Williamson, 1985, Grossman and Helpman, 2002a).

[10]Alternatively, one could assume that labor costs are also transferrable, but that their transfer leads to a significant fall in productivity. This fall in productivity could be explained, in an international

in capital and labor are chosen simultaneously and non-cooperatively.[11] Once a final-good producer and its supplier enter the market, they are locked into the relationship: the investments $rK_{x,k}(i)$ and $wL_{x,k}(i)$ are incurred upon entry and are useless outside the relationship. In Williamson's (1985) words, the initially *competitive* environment is *fundamentally transformed* into one of bilateral monopoly. Regardless of firm structure and the choice of cost sharing, fixed costs associated with production of the component are divided as follows: $f_F r^{\beta_k} w^{1-\beta_k}$ for the final-good producer and $f_S r^{\beta_k} w^{1-\beta_k}$ for the supplier, with $f_F + f_S = f$.[12]

Free entry into each sector ensures zero expected profits for a potential entrant. To simplify the description of the industry equilibrium, I assume that upon entry the supplier makes a lump-sum transfer $T_k(i)$ to the final-good producer, which can vary by industry and variety. Ex-ante, there is a large number of identical, potential suppliers for each variety in each industry, so that competition among these suppliers will make $T_k(i)$ adjust so as to make them break even. The final-good producer chooses the mode of organization so as to maximize its ex-ante profits, which include the transfer.

**Contract Incompleteness**   The setting is one of incomplete contracts. In particular, it is assumed that an outside party cannot distinguish between a high-quality and a low-quality intermediate input. Hence, input suppliers and final-good producers cannot sign enforceable contracts specifying the purchase of a certain type of intermediate input

---

context, by the inability of multinational firms to cope with idiosyncratic labor markets (c.f., Caves, 1996, p. 123).

[11] The assumption that the final-good producer decides between bearing *all* or *none* of the capital expenditures can be relaxed to a case of partial transferability. For instance, imagine that $x_k(i)$ was produced according to:

$$x_k(i) = \left( \frac{K_{x,k}^F(i)}{\beta_k} \right)^{\beta_k} \left( \frac{K_{x,k}^S(i)}{\eta(\beta_k)(1-\beta_k)} \right)^{\eta(\beta_k)(1-\beta_k)} \left( \frac{L_{x,k}(i)}{(1-\eta(\beta_k))(1-\beta_k)} \right)^{(1-\eta(\beta_k))(1-\beta_k)}$$

where $K_{x,k}^F(i)$ represents the part of the capital input that is transferable, and where $K_{x,k}^S(i)$ is inalienable to the supplier. As long as the elasticity of output with respect to transferable capital is higher, the higher the capital intensity in production, the same qualitative results would go through. In particular, as long as $\beta_k + \eta(\beta_k)(1-\beta_k)$ increases with $\beta_k$, the model would still predict more integration in capital-intensive industries (see footnote 24). I follow the simpler specification in (2.2) because it greatly simplifies the algebra of the general equilibrium.

[12] Henceforth, I associate a subscript $F$ with the final-good producer and a subscript $S$ with the supplier.

for a certain price. If they did, input suppliers would have an incentive to produce a low-quality input at the lower cost and still cash the same revenues. I take the existence of contract incompleteness as a fact of life, and will not complicate the model to relax the informational assumptions needed for this incompleteness to exist.[13] It is equally assumed that no outside party can verify the amount of ex-ante investments $rK_{x,k}(i)$ and $wL_{x,k}(i)$. If these were verifiable, then final-good producers and suppliers could contract on them, and the cost-reducing benefit of producing a low-quality input would disappear. For the same reason, it is assumed that the parties cannot write contracts contingent on the volume of sale revenues obtained when the final good is sold. Following Grossman and Hart (1986), the only contractibles ex-ante are the allocation of residual rights and the transfer $T_k(i)$ between the parties.[14]

If the supplier incurs all variable costs, the contract incompleteness gives rise to a standard hold-up problem. The final-good producer will want to renegotiate the price after $x_k(i)$ has been produced, since at this point the intermediate input is useless outside the relationship. Foreseeing this renegotiation, the input supplier will undertake suboptimal investments. The severity of the underinvestment problem is directly related to how weak the supplier's bargaining power is ex-post.

If the final-good producer shares capital expenditures with the supplier, the hold-up problem becomes two-sided. Because the investment in capital is also specific to the pair, the final-good producer is equally locked in the relationship, and thus its investment in capital will also tend to be suboptimal, with the extent of the underinvestment being inversely related to its bargaining power in the negotiation.

Because no enforceable contract will be signed ex-ante, the two firms will bargain over

---

[13]From the work of Aghion, Dewatripont and Rey (1994), Nöldeke and Schmidt (1995) and others, it is well-known that allowing for specific-performance contracts can lead, under certain circumstances, to efficient ex-ante relationship-specific investments. Che and Hausch (1997) have shown, however, that when ex-ante investments are cooperative (in the sense, that one party's investment benefits the other party), specific-performance contracts may not lead to first-best investment levels and may actually have no value.

[14]The assumption of non-contractibility of ex-ante investments could be relaxed to a case of partial contractibility. I have investigated an extension of the model in which production requires both contractible and non-contractible investments. If the marginal cost of non-contractible investments is increasing in the amount of contractible investments, the ability to set the contractible investments in the ex-ante contract is not sufficient to solve the underinvestment problem discussed below, and the model delivers results analogous to the ones discussed in the main text.

the surplus of the relationship after production takes place. At this point, the ex-ante investments as well as the quality of the input are observable to both parties and thus the costless bargaining will yield an ex-post efficient outcome. I assume that Generalized Nash Bargaining leaves the final-good producer with a fraction $\phi \in (0, 1)$ of the ex-post gains from trade. For reasons that will become clear below, I make the following assumption:

**Assumption 1:** $\phi > 1/2$.

Following the work of Grossman and Hart (1986) and Hart and Moore (1990), and contrary to the older transaction-cost literature, I assume that integration of the supplier does not eliminate the opportunistic behavior at the heart of the hold-up problem. Bargaining will therefore occur even when the final-good producer and the supplier are integrated. Ownership, however, crucially affects the distribution of ex-post surplus through its effect on each party's outside option. More specifically, the outside option for a final-good producer will be different when it owns the supplier and when it does not. In the latter case, the amount $x_k(i)$ is owned by the supplier and thus if the two parties fail to agree on a division of the surplus, the final-good producer is left with nothing. Conversely, under integration, the manager of the final-good producer can always fire the manager of the supplying division and seize the amount of input already produced.

If the final-good producer could fully appropriate $x_k(i)$ under integration, there would be no surplus to bargain over after production, and the supplier would optimally set $L_{x,k}(i) = 0$ (which of course would imply $x_k(i) = 0$). In such case, integration would never be chosen. To make things more interesting, I assume that by integrating the supplier, the final-good producer obtains the residual rights over only a fraction $\delta \in (0, 1)$ of the amount of $x_k(i)$ produced, so that the surplus of the relationship remains positive even under integration. I take the fact that $\delta$ is strictly less than one as given, but this assumption could be rationalized in a richer framework.[15]

---

[15]For instance, consider the following alternative set-up. Production of intermediates procedes in two stages. When firms enter the bargaining, only a fraction $\delta \in (0, 1)$ of $x_k(i)$ has been produced. After the bargaining and *immediately* before the delivery of the input, the supplier (and only him) can costlessly refine the component, increasing the amount produced from $\delta x_k(i)$ to $x_k(i)$ (one could think of this

Figure 2-3: Timing of Events
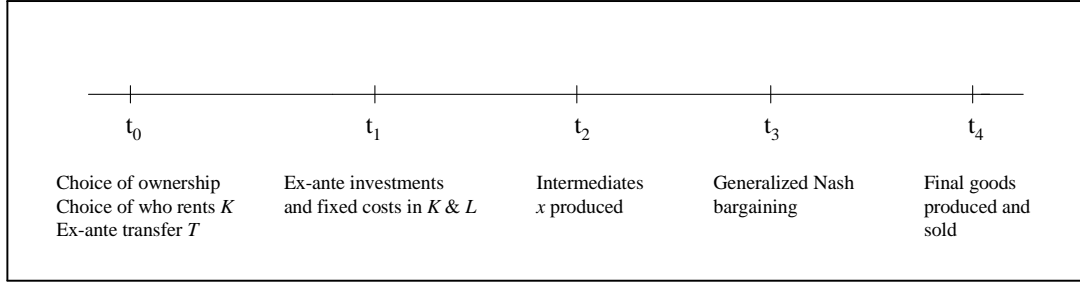


On the other hand, and because the component is completely specific to the final-good producer, the outside option for the intermediate input producer is zero regardless of ownership structure.

In choosing whether to enter the market with an integrated or a stand-alone supplier, the final-good producer considers the benefits and costs of integration. By owning the supplier, the final-good producer tilts the bargaining power in its favor but reduces the incentives for the supplier to make an efficient ex-ante investment in labor (and perhaps capital).

I now summarize the timing of events (see also Figure 2-3). At $t_0$, the final-good producer decides whether it wants to enter a given market. At this point, residual rights are assigned, the extent of cost sharing is decided, and the supplier makes a lump-sum transfer to the final-good producer. At $t_1$, firms choose their investments in capital and labor and also incur their fixed costs. At $t_2$, the final-good producer hands the specifications of the component (and perhaps the capital stock $K_{x,k}$) to its partner, and this latter produces the intermediate input (which can be of high or low quality). At $t_3$, the quality of the component becomes observable and the two parties bargain over the division of the surplus. Finally at $t_4$, the final good is produced and sold. For simplicity, I assume that agents do not discount the future between $t_0$ and $t_4$.

---

second stage as the branding of the product). Suppose, further, that the supplier does not perform this product refinement unless the two firms agree in the bargaining (this strategy is, in fact, subgame perfect). In such case, the surplus of the relationship would also be strictly positive.

## 2.2.2 Firm Behavior for a Given Demand

The model is solved by starting at $t_4$ and moving backwards. I will initially assume that final-good producers always choose to incur the variable costs $rK_{x,k}(i)$ themselves. Below, I will show that Assumption 1 is in fact sufficient to ensure that this is the case in equilibrium.

The assumption of a unit elasticity of substitution between varieties in industry $Y$ and $Z$ implies that we can analyze firm behavior in each industry independently. Consider industry $Y$, and suppose that at $t_4$, $n_{Y,V}$ pairs of integrated firms and $n_{Y,O}$ pairs of stand-alone firms are producing.[16] Let $p_{Y,V}(i)$ be the price charged by an integrating final-good producer for variety $i$ in industry $Y$. Let $p_{Y,O}(i)$ be the corresponding price for a non-integrating final-good producer.

From equation (2.1), demand for variety $i$ in industry $Y$ is given by

$$y(i) = A_Y p_Y(i)^{-1/(1-\alpha)}, \tag{2.3}$$

where

$$A_Y = \frac{\mu E}{\int_0^{n_{Y,V}} p_{Y,V}(j)^{-\alpha/(1-\alpha)}dj + \int_0^{n_{Y,O}} p_{Y,O}(j)^{-\alpha/(1-\alpha)}dj}, \tag{2.4}$$

and $E$ denotes total spending in the economy. I treat the number of firms as a continuum, implying that firms take $A_Y$ as given.

**Integrated pairs** Consider first the problem faced by a final-good producer and its integrated supplier. If the latter produces a high-quality intermediate input and the firms agree in the bargaining, the potential revenues from the sale of the final good are $R_Y(i) = p_Y(i)y(i)$, which using (2.2) and (2.3) can be written as

$$R_Y(i) = A_Y^{1-\alpha} \left(\frac{K_{x,Y}(i)}{\beta_Y}\right)^{\alpha\beta_Y} \left(\frac{L_{x,Y}(i)}{1-\beta_Y}\right)^{\alpha(1-\beta_Y)}.$$

---

[16]Henceforth, a subscript $V$ will be used to denote equilibrium values for final-good producers that vertically integrate their suppliers. A subscript $O$ will be used for those that outsource the production of the input.

On the other hand, if the parties fail to agree in the bargaining, the final-good producer will only be able to sell an amount $\delta y(i)$, which again using (2.3) will translate into sale revenues of $\delta^\alpha R_Y(i)$. The ex-post opportunity cost for the supplier is zero. The quasi-rents of the relationship are therefore $(1 - \delta^\alpha) R_Y(i)$.

The contract incompleteness will give rise to renegotiation at $t_3$. In the bargaining, Generalized Nash bargaining leaves the final-good division with its default option plus a fraction $\phi$ of the quasi-rents. On the other, the integrated supplier receives the remaining fraction $1 - \phi$ of the quasi-rents. Since both $\phi$ and $\delta$ are assumed to be strictly less than one, the supplier's ex-post revenues from producing a high-quality input are always strictly positive. Low-quality inputs will therefore never be produced at $t_2$.

Rolling back to $t_1$, the final-good producer will therefore set its investment in capital $K_{x,Y}(i)$ to maximize $\overline{\phi} R_Y(i) - r K_{x,Y}(i)$ where

$$\overline{\phi} = \delta^\alpha + \phi (1 - \delta^\alpha) > \phi.$$

The program yields a best-response investment $K_{x,Y}(i)$ in terms of factor prices, the level of demand as captured by $A_Y$, and the investment in labor $L_{x,Y}(i)$. On the other hand, the integrated supplier simultaneously sets $L_{x,Y}(i)$ to maximize $(1 - \overline{\phi}) R_Y(i) - w L_{x,Y}(i)$, from which an analogous reaction function $L_{x,Y}(i)$ is obtained.[17] Solving for the intersection of the two best-response functions yields the equilibrium ex-ante investments.[18] Plugging these investments into (2.2) and (2.3) and rearranging, yields the following

---

[17]The supplier could in principle find it optimal to complement the capital investment of the final-good division with some extra investment of its own, call it $K_{x,Y}^S$. Nevertheless, if the two investments in capital are perfect substitutes in production, Assumption 1 is sufficient to ensure that the optimal capital investment of the supplier is 0. To see this, notice that $\overline{\phi} (\partial R_Y(i)/\partial K_{x,Y}) > (1 - \overline{\phi}) (\partial R_Y(i)/\partial K_{x,Y}^S)$ for $\overline{\phi} > \phi > 1/2$. The complementary slackness condition thus implies that $K_{x,Y}^S = 0$.

[18]In particular, these are $K_{x,Y,V}(i) = \frac{\alpha \beta_Y \overline{\phi}}{r} A_Y \left( \frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha \overline{\phi}^{\beta_Y} (1-\overline{\phi})^{1-\beta_Y}} \right)^{-\alpha/(1-\alpha)}$ and $L_{x,Y,V}(i) = \frac{\alpha(1-\beta_Y)(1-\overline{\phi})}{w} A_Y \left( \frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha \overline{\phi}^{\beta_Y} (1-\overline{\phi})^{1-\beta_Y}} \right)^{-\alpha/(1-\alpha)}$.

identical optimal output and price for all varieties in industry $Y$:

$$y_V = x_{Y,V} = A_Y \left( \frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha \overline{\phi}^{\beta_Y} \left(1 - \overline{\phi}\right)^{1-\beta_Y}} \right)^{-1/(1-\alpha)} \tag{2.5}$$

$$p_{Y,V} = \frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha \overline{\phi}^{\beta_Y} \left(1 - \overline{\phi}\right)^{1-\beta_Y}}. \tag{2.6}$$

Facing a constant elasticity of demand, the final-good producer charges a constant mark-up over marginal cost. The mark-up is $1/(\overline{\phi}^{\beta_Y} \left(1 - \overline{\phi}\right)^{1-\beta_Y})$ times higher than the mark-up that would be charged if contracts were complete. From equation (2.6), if $\beta_Y$ is high, the mark-up is relatively higher, the lower is $\overline{\phi}$. Conversely, if production of $x_Y$ requires mostly labor ($\beta_Y$ low), the mark-up is relatively higher, the higher is $\overline{\phi}$.

Using the expressions for $y_V$ and $p_{Y,V}$, it is easy to check that the equilibrium investment levels are also identical for all varieties and satisfy $rK_{x,Y,V} = \alpha \beta_Y \overline{\phi} p_{Y,V} y_V$ and $wL_{x,Y,V} = \alpha(1 - \beta_Y)(1 - \overline{\phi}) p_{Y,V} y_V$.

At $t_1$, the two parties also choose how much capital and labor to rent in incurring the fixed costs. Applying Shepard's lemma, factor demands in the fixed costs sector are

$$
\begin{aligned}
K_{f,h,Y} &= \beta_Y f_h \left(\frac{w}{r}\right)^{1-\beta_Y} \\
L_{f,h,Y} &= (1 - \beta_Y) f_h \left(\frac{w}{r}\right)^{-\beta_Y},
\end{aligned}
\tag{2.7}
$$

for $h \in \{F, S\}$.

Finally, at $t_0$ the supplier makes a lump-sum transfer $T_{Y,V}$ to the final-good producer. As discussed above, at $t_0$, there is a large number of potential suppliers, so that ex-ante competition among them ensures that this transfer exactly equals the supplier's ex-ante profits.[19] Using the value of this transfer, ex-ante profits for an integrating final-good producer can finally be expressed as

$$\pi_{F,V,Y} = \left(1 - \alpha(1 - \beta_Y) + \alpha \overline{\phi}(1 - 2\beta_Y)\right) A_Y p_{Y,V}^{-\alpha/(1-\alpha)} - f r^{\beta_Y} w^{1-\beta_Y}, \tag{2.8}$$

---

[19] In particular, this transfer is $T_{Y,V} = (1 - \overline{\phi})(1 - \alpha(1 - \beta_Y)) A_Y p_{Y,V}^{-\alpha/(1-\alpha)} - f_S r^{\beta_Y} w^{1-\beta_Y}$.

where $p_{Y,V}$ is given in (2.6).

**Pairs of stand-alone firms**  If the firms enter the market as stand-alone firms, the supplier is entitled to the residual rights of control over the amount of input produced at $t_2$. The ex-post opportunity cost for the final-good producer is therefore zero in this case. As for the supplier, since the component is specific to the final-producer, the value of $x_Y(i)$ outside the relationship is also again zero. It follows that if the intermediate input producer hands a component with the correct specification, the potential sale revenues $R_Y(i)$ will entirely be quasi-rents. The final-good producer will obtain a fraction $\phi$ of this surplus in the bargaining, so at $t_1$ it will choose $K_{x,Y}(i)$ to maximize $\phi R_Y(i) - rK_{x,Y}(i)$. On the other hand, the supplier will set $L_{x,Y}(i)$ so as to maximize $(1 - \phi) R_Y(i) - wL_{x,Y}(i)$.

From here on, it is clear that the solution to the problem is completely analogous to that for pairs of integrated firms, with $\phi$ replacing $\overline{\phi}$ in equations (2.5) through (2.8). In particular, profits for a final-good producer that chooses to outsource the production of the intermediate input will be

$$\pi_{F,Y,O} = (1 - \alpha(1 - \beta_Y) + \alpha\phi(1 - 2\beta_Y)) A_Y p_{Y,O}^{-\alpha/(1-\alpha)} - fr^{\beta_Y} w^{1-\beta_Y}, \qquad (2.9)$$

where $p_{Y,O} = r^{\beta_Y} w^{1-\beta_Y}/(\alpha\phi^{\beta_Y}(1-\phi)^{1-\beta_Y})$.

**Comparison with an environment with complete contracts**  We can compare the previous two situations to one in which the quantity and quality of the component (as well the ex-ante investments) were verifiable. In such a case, the two parties would bargain over the division of the surplus upon entry and the contract would not be renegotiated ex-post. Upon entry, the threat point for both parties would be zero . The surplus of the relationship would thus be given by $S_Y(i) = p_Y(i)y(i) - rK_{x,Y}(i) - wL_{x,Y}(i) - fr^{\beta_Y} w^{1-\beta_Y}$. At $t_1$, the final-good producer would choose $K_{x,Y}(i)$ to maximize $\phi S_Y(i)$, while the supplier would set $L_{x,Y}(i)$ to maximize $(1 - \phi) S_Y(i)$. It is straightforward to check that the impossibility of writing enforceable contracts leads to underinvestment in both $K_{x,Y}$ and $L_{x,Y}$. In particular, letting $K_{x,Y}^*$ and $L_{x,Y}^*$ denote the optimal contractible investments,

34

Figure 2-4: Complete vs. Incomplete Contracts
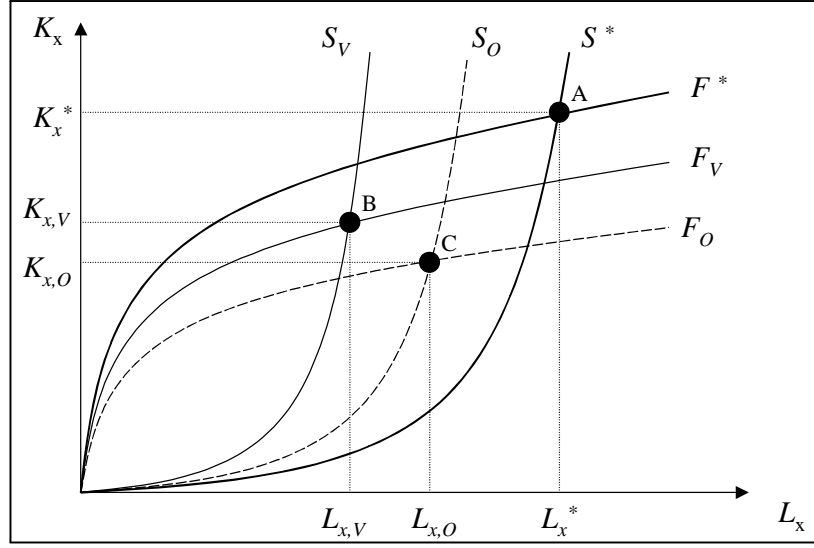


it is easy to show that $K^*_{x,Y} > \max\{K_{x,Y,V}, K_{x,Y,O}\}$ and $L^*_{x,Y} > \max\{L_{x,Y,V}, L_{x,Y,O}\}$.[20]

Underinvestment stems from the fact that, with incomplete contracts, each firm receives only a fraction of the marginal return to its ex-ante investment. The inefficiency is depicted in Figure 2-4. The curves $F^*$ and $S^*$ represent the reaction functions $K^*_{x,Y}(L_{x,Y})$ and $L^*_{x,Y}(K_{x,Y})$ under complete contracts, with the corresponding equilibrium in point A. Similarly, B and C depict the incomplete-contract equilibria corresponding to integration and outsourcing, respectively. An important point to notice from Figure 2-4 is that the underinvestment in labor relative to that in capital tends to be greater under integration that under outsourcing.[21] This follows from the fact that under integration, the supplier has a relatively weaker bargaining power and thus receives a smaller fraction of the marginal return to its ex-ante investment. A similar argument explains why the

---

[20] In the case of capital this follows from

$$\frac{\alpha\beta_Y A_Y}{r}\left(\frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha}\right)^{\frac{-\alpha}{1-\alpha}} > \max\left\{\frac{\alpha\beta_Y \overline{\phi} A_Y}{r}\left(\frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha\overline{\phi}^{\beta_Y}\left(1-\overline{\phi}\right)^{1-\beta_Y}}\right)^{\frac{-\alpha}{1-\alpha}}, \frac{\alpha\beta_Y \phi A_Y}{r}\left(\frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha\phi^{\beta_Y}\left(1-\phi\right)^{1-\beta_Y}}\right)^{\frac{-\alpha}{1-\alpha}}\right\}.$$

[21] By this I mean that $\left(L^*_{x,Y}/L_{x,Y,V}\right)/\left(K^*_{x,Y}/K_{x,Y,V}\right) > \left(L^*_{x,Y}/L_{x,Y,O}\right)/\left(K^*_{x,Y}/K_{x,Y,O}\right)$. Note that this also implies that controlling for industry characteristics, integrated suppliers should be using a higher capital-labor ratio in production than nonintegrated ones. This is consistent with the results of some empirical studies, discussed in Caves (1996, pp. 230-231) and Dunning (1993, p. 296), that compare capital intensity in overseas subsidiaries of multinational firms with that of independent domestic firms in the host country.

investment in capital tends to be relatively more inefficient under outsourcing.

**The rationale for cost sharing**  Consider now the problem faced by an independent supplier when the final-good producer decides not to contribute to variable costs. In such a case, the supplier chooses $K_{x,Y}(i)$ and $L_{x,Y}(i)$ to maximize $(1 - \phi) R_Y(i) - r K_{x,Y}(i) - w L_{x,Y}(i)$, and the final-good producer simply receives $\phi R_Y(i)$ ex-post. Following similar steps as before, it is easy to show that ex-ante profits for a final-good producer can now be expressed as

$$\widetilde{\pi}_{F,Y,O} = (\phi + (1 - \alpha)(1 - \phi)) A_Y \left( \frac{r^{\beta_Y} w^{1-\beta_Y}}{\alpha(1 - \phi)} \right)^{-\alpha/(1-\alpha)} - f r^{\beta_Y} w^{1-\beta_Y}. \qquad (2.10)$$

The case of an integrated supplier is completely analogous. In particular, the same expression (2.10) applies with $\overline{\phi}$ replacing $\phi$.

The following result follows from comparing equation (2.10) with (2.8) and (2.9):

**Lemma 1** *Under Assumption 1 (i.e., if $\phi > 1/2$), final-good producers will always decide to bear the cost of renting the capital required to produce the intermediate input.*

**Proof.** See Appendix 2.6.1. ■

The intuition for this result is that the higher is $\phi$, the smaller is the fraction of the marginal return to its ex-ante investments that the supplier receives, and thus the less it will invest in $K_{x,Y}$. This underinvestment will have a negative effect on the value of the relationship, which is what the final-good producer maximizes ex-ante. For a large enough $\phi$ (in this case $1/2$), the detrimental effect of the underinvestment in capital is large enough so as to make it worthwhile for the final-good producer to bear the cost of renting $K_{x,Y}$ itself, even if by doing so it now exposes itself to a hold-up by the supplier. In other words, for $\phi > 1/2$, a supplier incurring all variable costs faces a too severe hold-up problem, which the final-good producer finds optimal to alleviate by sharing part of the required ex-ante investments.

### 2.2.3 Factor Intensity and Ownership Structure

At $t_0$, a final-good producer in industry $k = \{Y, Z\}$ chooses the ownership structure that maximizes its ex-ante profits. Comparing equations (2.8) and (2.9), it is straightforward to check that a final-good producer will choose to integrate its supplier whenever

$$\Theta = \frac{\pi_{F,V,k} + fr^{\beta_Y}w^{1-\beta_Y}}{\pi_{F,O,k} + fr^{\beta_Y}w^{1-\beta_Y}} = \frac{1 - \alpha(1-\beta_Y) + \alpha\overline{\phi}(1-2\beta_Y)}{1 - \alpha(1-\beta_Y) + \alpha\phi(1-2\beta_Y)} \left(\frac{p_{Y,V}}{p_{Y,O}}\right)^{-\alpha/(1-\alpha)} > 1.$$

This inequality is more likely to hold, the lower is $p_{Y,V}$ relative to $p_{Y,O}$, that is, the less distorted is the mark-up under integration relative to that under outsourcing. Plugging the equilibrium prices and using $\overline{\phi} = \delta^\alpha + \phi(1 - \delta^\alpha)$, it is possible to express $\Theta$ in terms of the fundamental parameters of the model:

$$\Theta(\beta_k, \alpha, \phi, \delta) = \left(1 + \frac{\alpha(1-\phi)\delta^\alpha(1-2\beta_k)}{1 - \alpha(1-\beta_k) + \alpha\phi(1-2\beta_k)}\right)\left(1 + \frac{\delta^\alpha}{\phi(1-\delta^\alpha)}\right)^{\frac{\alpha\beta_k}{1-\alpha}} (1-\delta^\alpha)^{\frac{\alpha}{1-\alpha}}.$$

$$(2.11)$$

An important point to notice here is that $\Theta(\cdot)$ is not a function of factor prices. This follows directly from the assumption of Cobb-Douglas technology and isolates the partial equilibrium decision to integrate or outsource from any potential general-equilibrium feedbacks. This implied block-recursiveness is a useful property for solving the model sequentially, but the main results should be robust to more general specifications for technology.[22]

In order to explain the pattern in Figure 2-1, it is central to understand how the relative attractiveness of integration (as captured by $\Theta$) is affected by the capital intensity in production. The following lemma states that $\Theta(\beta_k, \alpha, \phi, \delta)$ is an increasing function

---

[22] In particular, with a more general CES technology of the type

$$x_k = \left(\beta_k \left(\frac{K_{x,k}}{\beta_k}\right)^{\frac{\sigma-1}{\sigma}} + (1-\beta_k)\left(\frac{L_{x,k}}{1-\beta_k}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}},$$

$\Theta(\cdot)$ also becomes a function of the wage-rental ratio in the economy. Interestingly, for $\sigma < 1$, the model predicts that, *ceteris paribus*, the propensity to outsource will be higher in countries with a higher wage-rental ratio. The drawback of this generalization is that the model turns out to be beyond the reach of paper-and-pencil analysis.

of $\beta_k$.

**Lemma 2** *The attractiveness of integration, as measured by $\Theta(\cdot)$, increases with the capital intensity of intermediate input production $\beta_k$: $\partial \Theta(\cdot)/\partial \beta_k > 0$ for all $\beta_k \in [0,1]$.*

    **Proof.** *See Appendix 2.6.2.* ∎

The intuition for why $\Theta(\beta_k, \alpha, \phi, \delta)$ is increasing in $\beta_k$ is straightforward. The higher the capital intensity of an industry, the more value-reducing will the underinvestment in capital be. Furthermore, as discussed above, the underinvestment in capital tends to be more severe under integration than under outsourcing. It thus follows that profits under integration relative to those under outsourcing will tend to be higher, the higher the capital intensity in production.[23]

Lemma 2 paves the way for the following central result:

**Proposition 1** *Given a triplet $\alpha, \phi, \delta \in (0,1)$, there exists a unique threshold capital intensity $\widehat{\beta}(\alpha, \phi, \delta) \in (0,1)$ such that all firms with $\beta_k < \widehat{\beta}(\alpha, \phi, \delta)$ choose to outsource production of the intermediate input (i.e., $\Theta(\beta_k, \cdot) < 1$), while all firms with $\beta_k > \widehat{\beta}(\alpha, \phi, \delta)$ choose to integrate their suppliers (i.e., $\Theta(\beta_k, \cdot) > 1$). Only firms with capital intensity $\widehat{\beta}(\alpha, \phi, \delta)$ are indifferent between these two options.*

    **Proof.** *See Appendix 2.6.3.* ∎

The logic of this result lies at the heart of Grossman and Hart's (1986) seminal contribution. Ex-ante efficiency dictates that residual rights should be controlled by the party undertaking a relatively more important investment. If production of the intermediate input requires mostly labor, then the investment made by the final-good producer will be relatively small, and thus it will be optimal to assign the residual rights of control to the supplier. Conversely, when the capital investment is important, the

---

[23]Despite this clear intuition, proving that $\partial \Theta(\cdot)/\partial \beta_k$ is positive is somewhat cumbersome (see Appendix 2.6.2). This is due to a counterbalancing effect. Integration enhances efficiency in capital-intensive industries by reducing the underinvestment problem. But this, of course, comes at the expense of higher capital expenditures which, *ceteris paribus,* tend to reduce profits. Lemma 2 shows, however, that this latter effect is always outweighed by the former.

final-good producer will optimally choose to tilt the bargaining power in its favor by obtaining these residual rights, thus giving rise to vertical integration.[24]

Proposition 1 advances a rationale for the first fact identified in the introduction. To the extent that vertical integration of suppliers occurs mostly in capital-intensive industries, one would expect the share of intrafirm trade in those industries to be relatively higher than that in labor-intensive industries. Nevertheless, Proposition 1 cannot by itself justify the trade pattern in Figure 2-1. An explanation of this fact requires a proper modelling of international trade flows, which I carry out in section 2.3. Furthermore, the open-economy version of the model will naturally give rise to the cross-country pattern unveiled in Figure 2-2. Before moving on, however, a characterization of the general equilibrium of the closed economy is necessary.

**Other comparative statics**   Equation (2.11) lends itself to other comparative static exercises. For instance, it is possible to show that $\Theta(\cdot)$ is a decreasing function of $\phi$, which by the implicit function theorem implies that the cut-off $\widehat{\beta}(\alpha, \phi, \delta)$ is an increasing function of $\phi$. To understand this result, notice that an increase in $\phi$ shifts bargaining power from the supplier to the final-good producer regardless of ownership structure (since $\overline{\phi}$ increases with $\phi$). It thus follows that increasing $\phi$ necessarily worsens the incentives for the supplier. To compensate for this, the final-good producer will now find it profitable to outsource in a larger measure of capital intensities.

The effect of $\alpha$ is in general ambiguous as it appears in several terms in equation (2.11). Numerical analysis indicates, however, that an increase in competition (a higher $\alpha$) tends to enhance outsourcing in sufficiently labor-intensive firms, while promoting integration in the most capital-intensive ones. The intuition for this result is that the higher the elasticity of substitution in demand, the more sensitive will profits be to the price charged by the final-good producer. A natural response to an increase in $\alpha$ is thus a shift towards higher efficiency, which translates into giving more bargaining power to suppliers in labor-intensive pairs, and better incentives to final-good producers in capital-

---

[24]The result goes through if the input is produced according to the technology in footnote 11 and $\beta_k + \eta(\beta_k)(1 - \beta_k)$ increases with $\beta_k$. In particular, the function $\Theta(\beta_k, \alpha, \phi, \delta)$ is identical in this more general case, so that Proposition 1 still holds for the same $\widehat{\beta}$. Having the final-good producer incur all capital expenditures is therefore not an important asumption.

intensive pairs.[25]

Finally, an increase in $\delta$ corresponds to an increase in $\overline{\phi}$ holding constant $\phi$, i.e., a fall in the bargaining power of the supplier under integration. The effect of such an increase depends again on the capital intensity of the production process. In labor-intensive firms the incentives of the supplier are very important and thus efficiency considerations will dictate a shift towards more outsourcing in response to an increase in $\delta$. On the other hand, in capital-intensive firms, an increase in $\delta$ may make integration more attractive, as it now secures the more significant investor a larger fraction of the marginal return to its investment. Numerical analysis tends to broadly support these intuitions.

### 2.2.4   Industry Equilibrium

In this section, I describe the partial equilibrium in a particular industry taking factor prices as given. Again, without loss of generality, I focus on industry $Y$. In equilibrium, free entry implies that no firm makes positive expected profits. In principle, three equilibrium modes of organization are possible: (i) a mixed equilibrium with some varieties being produced by integrated pairs and others by non-integrated pairs; (ii) an equilibrium with pervasive integration in which no final-good producer finds it profitable to outsource the production of the intermediate input; and (iii) an equilibrium with pervasive outsourcing in which no final-good producer chooses to vertically integrate its supplier.

The assumption that all firms in a given industry share the same capital intensity greatly simplifies the analysis. In particular, the following is a straightforward corollary of Proposition 1:

**Lemma 3** *A mixed equilibrium in industry $Y$ only exists in a knife-edge case, namely when $\beta_Y = \widehat{\beta}(\alpha, \phi, \delta)$ (i.e., $\Theta(\beta_Y, \cdot) = 1$). An equilibrium with pervasive integration in industry $Y$ exists only if $\beta_Y > \widehat{\beta}(\alpha, \phi, \delta)$ (i.e., $\Theta(\beta_Y, \cdot) > 1$). An equilibrium with*

---

[25]To see where the result is coming from, ignore the first term in (2.11) as well as the effect of $\alpha$ through the terms $\delta^\alpha$. Then the effect of $\alpha$ is positive as long as $(1 - \delta^\alpha)(1 + \delta^\alpha/(\phi(1 - \delta^\alpha)))^\beta > 1$, that is if $\beta > \overline{\beta}$ for some $\overline{\beta}(\phi, \delta, \alpha) \in (0, 1)$. Naturally, the sign of the derivative also depends on the values of $\phi$ and $\delta$. I stress the role of factor intensity here since the channel is absent in other papers that have studied the relationship between market competition and the attractiveness of outsourcing (e.g. Grossman and Helpman, 2002a, and Marin and Verdier, 2001).

*pervasive outsourcing in industry $Y$ exists only if $\beta_Y < \widehat{\beta}(\alpha, \phi, \delta)$ (i.e., $\Theta(\beta_Y, \cdot) < 1$).*

Since a mixed equilibrium does not generically exist, I focus below on a characterization of the two remaining types of equilibria.

**Equilibrium with Pervasive Integration**   Consider an equilibrium in which only integrating final-good producers enter the market. As discussed above, the ex-ante transfer $T_{Y,V}$ ensures that suppliers always break even. If no final-good producer outsources the production of $x_Y$, all firms will charge a price for $y(i)$ given by equation (2.6). Since $n_{Y,O} = 0$, equation (2.4) simplifies to $A_{Y,V} = \mu E p_{Y,V}^{\alpha/(1-\alpha)}/n_{Y,V}$. On the other hand, from equation (2.8), for integrating final-good producers to make zero profits, demand must also satisfy:

$$A_{Y,V} = \frac{fr^{\beta_Y}w^{1-\beta_Y}}{1 - \alpha(1 - \beta_Y) + \alpha\overline{\phi}(1 - 2\beta_Y)}p_{Y,V}^{\alpha/(1-\alpha)}. \qquad (2.12)$$

It thus follows that the equilibrium number of varieties in an equilibrium with pervasive integration must be given by:

$$n_{Y,V} = \frac{1 - \alpha(1 - \beta_Y) + \alpha\overline{\phi}(1 - 2\beta_Y)}{fr^{\beta_Y}w^{1-\beta_Y}}\mu E. \qquad (2.13)$$

Naturally, the equilibrium number of varieties in industry $Y$ depends positively on total spending in the industry and negatively on fixed costs. The equilibrium level of output of each variety can be obtained by plugging the equilibrium demand (2.12) into equation (2.5):

$$y_V = \frac{\alpha\overline{\phi}^{\beta_Y}\left(1 - \overline{\phi}\right)^{1-\beta_Y} f}{1 - \alpha(1 - \beta_Y) + \alpha\overline{\phi}(1 - 2\beta_Y)}. \qquad (2.14)$$

Equilibrium factor demands can similarly be obtained by plugging (2.12) into the expressions in footnote 18.

**Equilibrium with Pervasive Outsourcing**   Consider next an equilibrium in which no final-good producer vertically integrates its supplier. In such an equilibrium every firm charges a price given by $p_{Y,O}$ which makes equation (2.4) simplify to $A_{Y,O} =$

$\mu E p_{Y,O}^{\alpha/(1-\alpha)}/n_{Y,O}$. Combining this expression with the free-entry condition

$$A_{Y,O} = \frac{fr^{\beta_Y}w^{1-\beta_Y}}{1 - \alpha(1 - \beta_Y) + \alpha\phi(1 - 2\beta_Y)}p_{Y,O}^{\alpha/(1-\alpha)}, \tag{2.15}$$

yields the equilibrium number of pairs undertaking outsourcing,

$$n_{Y,O} = \frac{1 - \alpha(1 - \beta_Y) + \alpha\phi(1 - 2\beta_Y)}{fr^{\beta_Y}w^{1-\beta_Y}}\mu E, \tag{2.16}$$

which is identical to (2.13) with $\phi$ replacing $\overline{\phi}$. The equilibrium values for output and factor demands are also analogous to those for the equilibrium with pervasive integration.

## 2.2.5 General Equilibrium

Having described the equilibrium in a particular industry, we can now move to the general equilibrium of the closed economy, in which income equals spending

$$E = rK + wL, \tag{2.17}$$

and the product, capital and labor markets clear.

By Walras' law, we can focus on the equilibrium in the labor market.[26] Letting $L_Y$ and $L_Z$ denote total labor demand by each pair in industries $Y$ and $Z$, labor market clearing requires $n_Y L_Y + n_Z L_Z = L$. We can decompose $L_Y$ into three components, depending on the equilibrium mode of organization. In an equilibrium with pervasive integration,

$$L_Y = L_{x,Y,V} + L_{f,F,Y} + L_{f,S,Y}. \tag{2.18}$$

The first term is the total amount of labor hired by integrated suppliers for the manufacturing of intermediate inputs. The remaining terms are the amounts of labor hired to cover fixed costs: $L_{f,F,Y}$ is the amount of labor employed in total fixed costs by final-good producers and $L_{f,S,Y}$ is the analogous demand by suppliers. From equation (2.7), notice that neither $L_{f,F,Y}$ nor $L_{f,S,Y}$ are affected by the equilibrium organization mode.

---

[26]The product market has already been assumed to clear in the previous sections.

Plugging (2.7) and (2.17) into equation (2.18), and substituting $n_{Y,V}$ and $L_{x,Y,V}$ for their equilibrium values, it is possible to simplify to:

$$n_{Y,V} L_Y = (1 - \beta_Y)\left(1 - \alpha\beta_Y(2\overline{\phi} - 1)\right)\frac{\mu\,(rK + wL)}{w}. \tag{2.19}$$

Following the same steps, it is straightforward to show that in an equilibrium with pervasive outsourcing,

$$n_{Y,O} L_Y = (1 - \beta_Y)\left(1 - \alpha\beta_Y(2\phi - 1)\right)\frac{\mu\,(rK + wL)}{w}. \tag{2.20}$$

Equations (2.19) and (2.20) imply that the share of income that labor receives is sensitive to the equilibrium mode of organization. Given the assumption of Cobb-Douglas technology, in a world of complete contracts, the share of income accruing to labor in industry $Y$ would be $\mu(1 - \beta_Y)$. With incomplete contracts, the share received by labor will be larger or smaller than $\mu(1 - \beta_Y)$ depending on whether $\phi$ or $\overline{\phi}$ are greater or smaller than $1/2$. Under Assumption 1, incomplete contracts tend to bias the distribution of income towards owners of capital. Intuitively, with $\phi > 1/2$, the underinvestment in labor is relatively more severe. For a given supply of factors, the relatively higher demand for capital tends to push up its price and thus its share in total income. As is clear from equations (2.19) and (2.20), this bias is greater under integration than under outsourcing.

To set the stage for an analysis of the share of intrafirm trade in total trade, I make the following assumption:

**Assumption 2:** $\beta_Y > \widehat{\beta} > \beta_Z$.

In words, I assume that the equilibrium in industry $Y$ is one with pervasive integration. Conversely, firms in the more labor-intensive industry $Z$ are assumed to outsource pervasively. It is useful to define the shares of income that accrues to capital in each sector, which using equations (2.19) and (2.20) are given by

$$\widetilde{\beta_Y} = \beta_Y(1 + \alpha\,(1 - \beta_Y)\,(2\overline{\phi} - 1))$$

and

$$\widetilde{\beta_Z} = \beta_Z \left(1 + \alpha \left(1 - \beta_Z\right) \left(2\phi - 1\right)\right).$$

Notice that $\beta_Y > \beta_Z$ implies $\widetilde{\beta_Y} > \widetilde{\beta_Z}$ and the presence of incomplete contracts does not create factor intensity reversals. Denoting the average labor share in the economy by $\sigma_L \equiv \mu(1 - \widetilde{\beta_Y}) + (1 - \mu)(1 - \widetilde{\beta_Z})$ and imposing the condition $n_{Y,V} L_Y + n_{Z,O} L_Z = L$, the equilibrium wage-rental ratio in the economy can be expressed as:

$$\frac{w}{r} = \frac{\sigma_L}{1 - \sigma_L} \frac{K}{L}, \tag{2.21}$$

The equilibrium wage-rental ratio is a linear function of the aggregate capital-labor ratio. This is a direct implication of the assumption of a Cobb-Douglas technology in both industries. The factor of proportionality is equal to the average labor share in the economy divided by the average capital share. As discussed above, Assumption 1 implies that labor shares are depressed relative to their values in a world with complete contracts. It follows that incomplete contracts also tend to depress the equilibrium wage-rental ratio in the economy.

## 2.3   The Multi-Country Model: Capital Abundance and Intrafirm Trade

Suppose now that the closed-economy described above is split into $J \geq 2$ countries, with each country receiving an endowment $K^j$ of capital and an endowment $L^j$ of labor. Factors of production are internationally immobile. Countries differ only in their factor endowments. In particular, individuals in all $J$ countries have identical preferences as specified in (2.1) and share access to the same technology in (2.2). The parameters $\phi$ and $\delta$ are also assumed to be identical everywhere. Countries are allowed to trade intermediate inputs at zero cost. Final goods are instead assumed to be nontradable, so that final-good producers produce their varieties in all $J$ countries. To be more specific,

each final-good producer has a (costless) plant in each of the $J$ countries.[27] Conversely, varieties of intermediates inputs will be produced in only one location in order to exploit economies of scale. I assume that for all $j \in J$, the capital-labor ratio $K^j/L^j$ is not *too different* from $K/L$, so that factor price equalization (FPE) holds, and the equilibrium prices and aggregate allocations are those of the integrated economy described above. Below, I derive both necessary and sufficient conditions for FPE to be achieved.

This section is in three parts. I first study the international location of production of intermediate inputs and show how the cross-country differences in factor endowments naturally give rise to cross-country differences in the relative number of varieties produced in each industry. I then analyze the implied patterns of international trade and discuss the determinants of its intrafirm component. Finally, I study the robustness of the results to alternative assumptions on the tradability of final goods.

### 2.3.1 Pattern of Production

Because countries differ only in their factor endowments, the cut-off capital intensity $\widehat{\beta}$ will be identical in all countries, and by Assumption 2, suppliers in industry $Y$ will be vertically integrated while those in industry $Z$ will remain non-integrated.

The factor market clearing conditions in country $j \in J$ can be written as:

$$n_Y^j \left( K_{x,Y}^j + K_{f,F,Y}^j + K_{f,S,Y}^j \right) + n_Z^j \left( K_{x,Z}^j + K_{f,F,Z}^j + K_{f,S,Z}^j \right) = K^j$$

and

$$n_Y^j \left( L_{x,Y}^j + L_{f,F,Y}^j + L_{f,S,Y}^j \right) + n_Z^j \left( L_{x,Z}^j + L_{f,F,Z}^j + L_{f,S,Z}^j \right) = L^j,$$

where $n_k^j$ refers now to the number of industry $k$ varieties of intermediate inputs produced in country $j$.[28] It is straightforward to check that factor demands for each variety depend

---

[27] Because final goods are costlessly produced, the model cannot endogenously pin down where their production is located. Assuming that they are not traded resolves this indeterminacy. In section 2.3.3, I show that the main result goes through under alternative set-ups that equally resolve the indeterminacy.

[28] To simplify notation, I drop all subscripts associated with the equilibrium mode of organization. For instance, I will denote the equilibrium number of industry $Y$ ($X$) varieties of intermediate inputs produced in country $j$ as $n_Y^j$ ($n_Z^j$) instead of $n_{Y,V}^j$ ($n_{Z,O}^j$).

only on worldwide identical parameters and on aggregate prices, which by assumption are also common in all countries. This implies that differences in the production patterns between countries will be channelled through the number of industry varieties produced in each country. In particular, using the integrated economy equilibrium values for $n_Y$ and $n_Z$, the factor demand conditions can be simplified to:

$$
(rK + wL) \left( \mu \widetilde{\beta_Y} \frac{n_Y^j}{n_Y} + (1 - \mu) \widetilde{\beta_Z} \frac{n_Z^j}{n_Z} \right) = rK^j
$$

$$
(rK + wL) \left( \mu \left( 1 - \widetilde{\beta_Y} \right) \frac{n_Y^j}{n_Y} + (1 - \mu) \left( 1 - \widetilde{\beta_Z} \right) \frac{n_Z^j}{n_Z} \right) = wL^j
$$

Combining these two expressions and plugging the equilibrium wage-rental ratio of the integrated economy, $w/r = (\sigma_L / 1 - \sigma_L) K/L$, yields the number of varieties of intermediate inputs of each industry produced in each country:

$$
n_Y^j = \left( \left( 1 - \widetilde{\beta_Z} \right) (1 - \sigma_L) \frac{K^j}{K} - \widetilde{\beta_Z} \sigma_L \frac{L^j}{L} \right) \frac{n_Y}{\left( \widetilde{\beta_Y} - \widetilde{\beta_Z} \right) \mu} \tag{2.22}
$$

and

$$
n_Z^j = \left( \widetilde{\beta_Y} \sigma_L \frac{L^j}{L} - \left( 1 - \widetilde{\beta_Y} \right) (1 - \sigma_L) \frac{K^j}{K} \right) \frac{n_Z}{\left( \widetilde{\beta_Y} - \widetilde{\beta_Z} \right) (1 - \mu)}, \tag{2.23}
$$

where $n_Y$ is given by (2.13) and $n_Z$ by (2.16) with $\beta_Z$ replacing $\beta_Y$. Equation (2.22) states that a given country $j \in J$ will produce a larger measure of varieties of intermediates in industry $Y$ the larger its capital-labor ratio. Conversely, from equation (2.22), the measure of industry-$Z$ varieties it produces is a decreasing function of its capital-labor ratio.

Note also that for a given $K^j/L^j$ both $n_Y^j$ and $n_Z^j$ are increasing in the size of the country, as measured by its share in world GDP, $s_j \equiv (rK^j + wL^j) / (rK + wL)$. In fact, it is straightforward to check that $n_Y^j > s^j n_Y$ if and only if $K^j/L^j > K/L$, and $n_Z^j > s^j n_Z$ if and only if $K^j/L^j < K/L$. In words, capital (labor)-abundant countries tend to produce a share of input varieties in the capital (labor)-intensive industries that exceeds their share in world income.

Figure 2-5: Pattern of Production for $J = 2$

For the above allocation to be consistent with FPE, it is necessary and sufficient that $n_Y^j > 0$ and $n_Z^j > 0$ for all $j \in J$, i.e., that no country fully specializes in any one sector.[29] Using equations (2.22) and (2.23), this condition can be written as:

**Assumption 3:** $\overline{\kappa} = \dfrac{\widetilde{\beta_Y}\sigma_L}{\left(1-\widetilde{\beta_Y}\right)(1-\sigma_L)} > \dfrac{K^j/L^j}{K/L} > \dfrac{\widetilde{\beta_Z}\sigma_L}{\left(1-\widetilde{\beta_Z}\right)(1-\sigma_L)} = \underline{\kappa}$  for all $j \in J$.

It can be checked that the upper bound $\overline{\kappa}$ is greater than one, while the lower bound $\underline{\kappa}$ is smaller than one. Assumption 3 thus requires the capital-labor ratio $K^j/L^j$ to be sufficiently similar to $K/L$.

Figure 2-5 provides a graphical representation of the production pattern for the case of two countries, the North ($N$) and the South ($S$). The graph should be familiar to readers of Helpman and Krugman (1985). $O^N$ and $O^S$ represent the origins for the North and the

---

[29]To understand necessity, notice that when factor prices depend only on *world* factor endowments, the capital-labor ratio in production is fixed and identical for all countries. Therefore, a given country cannot employ all its factors by producing in only one industry except in the knife-edge case in which its endowment of $K^j$ and $L^j$ exactly match that industry's factor intensity. For a discussion of sufficiency, see Helpman and Krugman (1985, p. 13-14).

South, respectively. The vectors $O^N Y$ and $O^N Z$ represent world employment of capital and labor in industries $Y$ and $Z$ in the equilibrium of the integrated economy. The set of factor endowments that satisfy Assumption 3 (i.e., FPE) corresponds to the set of points inside the parallelogram $O^N Y O^S Z$. Point $E$ defines the distribution of factor endowments. In the graph, the North is capital-abundant relative to the South. Line $BB'$ goes through point E and has a slope of $w/r$. The relative income of each country is thus held fixed for all points in line $BB'$ and inside the FPE set.

To map this figure to the pattern of production described above, I follow Helpman and Krugman (1985) in choosing units of measurement so that $\left\| O^N Y \right\| = n_Y y$, $\left\| O^N Z \right\| = n_Z z$, and $\left\| O^N O^S \right\| = E = rK + wL$. With the first two normalizations, we can graphically determine the number of varieties of intermediate inputs produced in each country. Moreover, with the last normalization, we can write $s^N = \left\| O^N C \right\| / \left\| O^N O^S \right\|$. Basic geometry then implies that $n_Y^N > s^N n_Y$ and $n_Z^N < s^N n_Z$, which is what we expected given that the North is capital abundant in the graph.

So far I have assumed that factors of production are internationally immobile. I therefore have not allowed final-good producers to rent the capital stock in their home country and export it to the country where intermediates are produced. Allowing for such international factor movements would not invalidate the equilibrium described above. In fact, by equalizing factor prices everywhere, international trade in intermediate inputs eliminates the incentives for capital to flow across countries.[30]

### 2.3.2 Pattern of Trade

Having described the international location of production of intermediate inputs, we can now move to a study of trade patterns, with a special emphasis on the share of intrafirm imports. Since the final good is nontradable, the entire volume of world trade will be in intermediate inputs.

A given country $N \in J$, will host $n_Y + n_Z$ plants producing an identical measure of varieties of final goods. Of the $n_Y$ plants in industry $Y$, a measure $n_Y^j$ will be importing

---

[30]More generally, I only require that costs of capital mobility are higher than costs of trading goods, so that international differences in rates of return are arbitraged away through trade flows rather than capital flows (c.f., Mundell, 1957).

the intermediate input from their integrated suppliers in country $j \neq N$. This volume of trade will thus be *intrafirm* trade. On the other hand, of the $n_Z$ plants in industry $Z$, a measure $n_Z^j$ will be importing the input from independent suppliers in country $j \neq N$. These transactions will thus occur at *arm's length*.

Before describing these flows in more detail, we must first confront the problem of how to value them. As I discussed above, the fact that contracts are incomplete precludes the purchase of a certain type of intermediate input for a certain price. In fact, there is no explicit price for these varieties. Because all variable costs are incurred in the country where the input is produced, a plausible assumption is to value these intermediates at average cost. But since the final good is produced at no cost, this implies that the *implicit* price of an intermediate input is simply $p_{Y,V}$ in industry $Y$ and $p_{Z,O}$ in industry $Z$.[31]

Without loss of generality, consider now a given country $N \in J$. On the production side, suppliers in country $N$ will be producing $n_Y^N$ and $n_Z^N$ varieties of intermediate inputs. On the consumption side, and since preferences are identical everywhere, consumers in country $N$ will incur a fraction $s^N$ of world spending on each variety. Since the final good is nontradable, this implies that country $N$ will be exporting a fraction $1 - s^N$ of the output each variety of intermediate input it produces, and will be importing a fraction $s^N$ of the output of each variety it does not produce.

Consider now a second country $S \in J$. From the above discussion, it is clear that the volume of $N$ imports from $S$ will be $s^N \left( n_Y^S p_Y y + n_Z^S p_Z z \right)$, or simply

$$M^{N,S} = s^N s^S \left( rK + wL \right). \tag{2.24}$$

Similarly, the volume of country $N$ exports to country $S$ is $s^S s^N \left( rK + wL \right)$. It thus follows that trade is balanced. Since both industries produce differentiated goods, for a given $s^N + s^S$, the volume of bilateral trade is maximized when both countries are of equal size (c.f., Helpman and Krugman, 1985).

Now let us look more closely at the composition of imports. Since only in industry $Y$

---

[31] Alternatively, intermediates could be valued according to the supplier's average revenues. In such case, the implicit prices would be $\left( 1 - \overline{\phi} \right) p_{Y,V}$ and $\left( 1 - \phi \right) p_{Z,O}$. Since, $\overline{\phi} > \phi$, the value of trade flows in industry $Y$ would be relatively more depressed. This would tend to attenuate the link between factor endowments and the volume of trade established in Proposition 2 below.

do final-good producers vertically integrate their suppliers, only imports in this industry will occur within firm boundaries. Denoting the volume of country $N$ intrafirm imports from $S$ by $M_{i-f}^{N,S}$, it follows that $M_{i-f}^{N,S} = s^N n_Y^S p_Y y$. Plugging the equilibrium value for $n_Y^S$ and rearranging, it is possible to express intrafirm imports as

$$M_{i-f}^{N,S} = s^N s^S \left(rK + wL\right) \frac{\left(1 - \widetilde{\beta_Z}\right)(1 - \sigma_L)\frac{K^S}{L^S} - \widetilde{\beta_Z}\sigma_L\frac{K}{L}}{\left(\widetilde{\beta_Y} - \widetilde{\beta_Z}\right)\left((1 - \sigma_L)\frac{K^S}{L^S} + \sigma_L\frac{K}{L}\right)}. \tag{2.25}$$

Intrafirm imports are thus increasing in the size of both the importing and exporting countries, and, from simple differentiation of (2.25), are also increasing in the capital-labor ratio of the exporting country.

Figure 2-6 depicts combinations of factor endowments that yield the same volume of intrafirm imports $M_{i-f}^{N,S}$, for the case in which there are only two countries, $N$ and $S$. The arrows in the graph point in the direction of increasing intrafirm imports. Point $C$ is such that $\|O^N C\| = \|C O^S\|$, implying that the line $BB'$ contains all points for which $s^N = s^S$. The graph shows how for a given capital-labor ratio of the exporting country (i.e., the South), $M_{i-f}^{N,S}$ is maximized when the two countries are of equal size. On the other hand, for a given relative size of the two countries, $M_{i-f}^{N,S}$ is increasing in the capital-labor ratio of the exporting country. In sum,

**Lemma 4** *For any pair of countries $N, S \in J$ with $S \neq N$, the volume of $N$'s intrafirm imports from $S$, $M_{i-f}^{N,S}$, is, for a given size $s^N$ of the importing country, an increasing function of the capital-labor ratio $K^S/L^S$ and the size $s^S$ of the exporting country. Furthermore, for a given $K^S/L^S$ and $s^S$, $S_{i-f}^{N,S}$ is also increasing in the size $s^N$ of the importing country.*

Now let $S_{i-f}^{N,S}$ denote the share of intrafirm imports in total imports, i.e. $S_{i-f}^{N,S} \equiv M_{i-f}^{N,S}/M^{N,S}$. Dividing equation (2.25) by (2.24) yields

$$S_{i-f}^{N,S} = \frac{\left(1 - \widetilde{\beta_Z}\right)(1 - \sigma_L)\frac{K^S}{L^S} - \widetilde{\beta_Z}\sigma_L\frac{K}{L}}{\left(\widetilde{\beta_Y} - \widetilde{\beta_Z}\right)\left((1 - \sigma_L)\frac{K^S}{L^S} + \sigma_L\frac{K}{L}\right)}. \tag{2.26}$$

Notice that by Assumption 3, $S_{i-f}^{N,S} \in (0,1)$. Furthermore, when $K^S/L^S$ goes to $\underline{\kappa} \cdot K/L$,

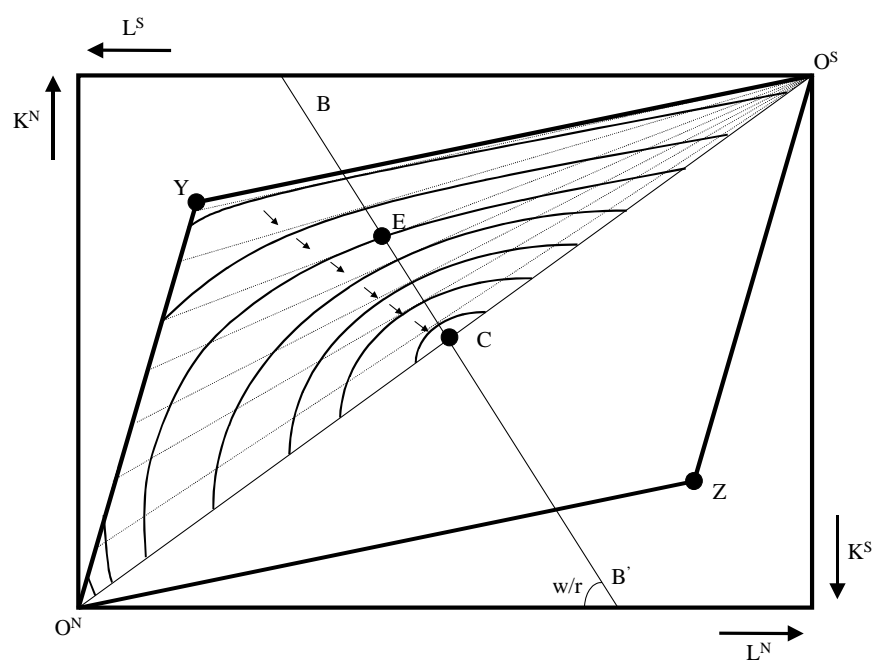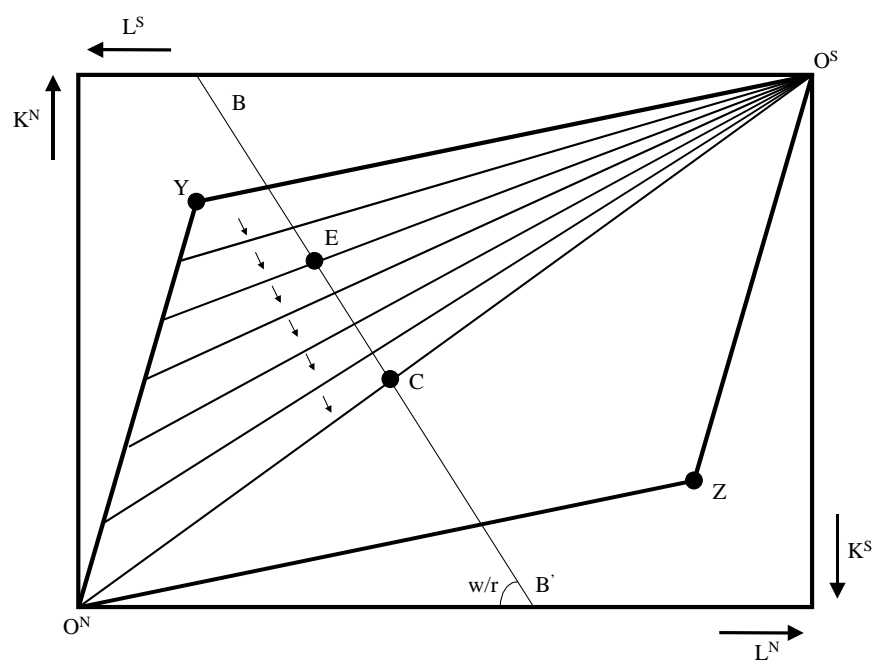Figure 2-6: Volume of Intrafirm Imports



Figure 2-7: Share of Intrafirm Imports

the South stops producing varieties of intermediates in industry $Y$, and thus $S_{i-f}^{N,S}$ goes to 0. Similarly, when $K^S/L^S$ goes to $\overline{\kappa} \cdot K/L$, the South fully specializes in industry $Y$, and thus $S_{i-f}^{N,S}$ goes to 1. More importantly, simple differentiation of (2.26) reveals that:

**Proposition 2** *For any pair of countries $N, S \in J$ with $S \neq N$, the share $S_{i-f}^{N,S}$ of intrafirm imports in total $N$'s imports from $S$ is an increasing function of the capital-labor ratio $K^S/L^S$ of the exporting country. Furthermore, for a given $K^S/L^S$, $S_{i-f}^{N,S}$ is unaffected by the relative size of each country.*

The first statement is one of the key results of the paper. In particular, it shows how in a world with international trade, the pattern of Figure 2-2 in the introduction is a direct implication of the pattern in Figure 2-1.

Figure 2-7 provides a graphical illustration of Proposition 2 for the case of two countries. Since $S_{i-f}^{N,S}$ is uniquely determined by $K^S/L^S$, the sets of points for which $S_{i-f}^{N,S}$ is constant are simple straight lines from the origin of the South. The arrows indicate that for any relative size of each country, $S_{i-f}^{N,S}$ is increasing in $K^S/L^S$.

In the next section, I will test the theoretical predictions on $S_{i-f}^{N,S}$ and $M_{i-f}^{N,S}$ using data on U.S. imports. Before doing so, I briefly argue that the assumption of nontradability of the final-good varieties is not crucial for the results derived above.

### 2.3.3 Alternative Assumptions on the Tradability of Final Goods

**A. Probabilistic Location of Final-Good Production** Consider the case in which final-good varieties can be traded, but assume that each variety is produced in only one location.[32] Let final goods be traded only at arm's length. Assume also that production of these varieties is distributed across countries in a probabilistic manner. Denote by $P_k^j(i)$ the probability that a given final-good variety $i$ of industry $k$ is assigned to country $j$. Let $P_k^j(i|j')$ be the same probability conditional on the intermediate input corresponding to that variety being produced in country $j'$. The international location of intermediate-input production is as described in section 2.3.1. Consider the imports of a

---

[32]Imagine, for instance, that there is a negligible but positive fixed cost of final-good production. The discussion characterizes the limit in which this cost goes to zero.

given country $N \in J$ from another country $S \in J$. These consist of final-good varieties assigned to country $S$, as well as of intermediate inputs required for the production of final-good varieties assigned to $N$. Expected total imports can be expressed as:

$$M^{N,S} = s^N p_Y y \int_0^{n_Y} P_Y^S(i)di + s^N p_Z z \int_0^{n_Z} P_Z^S(i)di + p_Y y \int_0^{n_Y^S} P_Y^N(i|S)di + p_Z z \int_0^{n_Z^S} P_Z^N(i|S)di,$$

where the first two terms correspond to imports of final-good varieties and the remaining ones are imports of intermediate inputs. Of these imports, only the third component will be traded within firm boundaries. As a benchmark, consider the case in which the location of final-good varieties is completely random and $P_k^j(i) = P_k^j(i|j') = s^j$. In words, the probability that a given final-good variety is produced in a certain country is proportional to the size of the country, but is independent of the variety, industry and location of intermediate-input production. It is straightforward to check that, in this probabilistic framework, expected imports are twice as large as in the previous section, i.e., $M^{N,S} = 2s^N s^S (rK + wL)$, while expected intrafirm imports are identical to those in the case of nontradable final goods, i.e., $M_{i-f}^{N,S} = s^N n_Y^S p_Y y$. It thus follows that the share of intrafirm imports in total imports will be one-half of that in equation (2.26). Therefore, all the claims in Lemma 4 and Proposition 2 remain valid in this alternative set-up. This example illustrates that in order to eliminate the prediction relating the share of intrafirm imports and relative capital abundance, one needs to introduce *ad hoc* correlations in the cross-country distribution of final-good varieties.[33]

**B. Technical Know-How and North-South Trade**   Consider yet an alternative set-up. Countries not only differ in factor endowments but also in their stock of knowledge. In particular, there is a set of countries, the North, that possess the know-how to produce varieties of final-goods. The remaining countries, the South, do not have this know-how and thus export only intermediate inputs and import only final goods. Denote the former set by $\mathcal{N} \subseteq J$ and the latter by $\mathcal{S} \subseteq J$. The final good is traded at arm's

---

[33]For instance, if the locations of intermediate and final-good production were perfectly correlated, then $P_k^j(i|j') = 0$ for all $j \neq j'$, and the share of intrafirm imports would be zero for all countries.

length at a negligible but positive cost. Since production of final varieties is costless, this implies that countries belonging to $\mathcal{N}$ will only trade intermediates inputs between themselves (it follows that if $\mathcal{S} = \emptyset$, the pattern of trade would be identical to the one discussed above). In order to service the Southern markets for final goods, countries in the North will import varieties of intermediate input in excess of the amount needed to satisfy their domestic demand. Assume that these *excess* imports are not biased towards any particular industry, in the sense that these excess imports constitute the same fraction of the exporter's production in each industry.

Since countries in $\mathcal{S}$ import only final-good varieties, the bilateral share of intrafirm imports in total imports is trivially 0 for all $s \in \mathcal{S}$. Conversely, a given country $n \in \mathcal{N}$ will only import intermediate inputs. In particular, given an exporting country $j \in J$, imports from $j$ will be comprised of those intermediates required for domestic consumption of final goods and those required to service Southern countries. Denote by $\eta_n^j$ the share of $j$'s production that is imported to service Southern countries. Then total imports from $j$ will be $(s^n + \eta_n^j)s^j E$, with their intrafirm component being $(s^n + \eta_n^j)n_Y^S p_Y y$. When taking the ratio, the terms in $\eta_n$ cancel, and the share of intrafirm imports in total imports is again given by an expression identical to equation (2.26). Conditional on the importing country belonging to $\mathcal{N}$ (which is surely the case for the United States), Proposition 2 thus still remains valid. Notice also that the total volume of intrafirm imports $M_{i-f}^{n,j}$ will again increase in the capital-labor ratio and the size of the exporting country. Furthermore, unless $\eta_n^j$ is negatively correlated with the size of country $n$, the second statement in Lemma 4 will also still apply.

## 2.4 Econometric Evidence

In this section, I use data on intrafirm and total U.S. imports to test more formally the empirical validity of the main results of the paper. I start by studying more closely the relationship between the factor intensity of the exporting industry and the share of intrafirm imports in total imports. In particular, I show that the clear correlation in Figure 2-1 does not seem to be driven by the omission of other relevant variables. Next, I move on to the relationship between relative factor endowments and the share of

intrafirm imports. The link predicted by Proposition 2 is confirmed even after controlling for additional factors that could reasonably be expected to affect this share. Furthermore, as predicted by the theory, the size of the exporting country is shown not to have an independent effect on the share of intrafirm imports. Finally, I analyze the determinants of the total volume of intrafirm imports and show that, consistently with Lemma 4, total intrafirm imports are indeed significantly affected by the size of the exporting country.

Before discussing the econometric results, however, the next two sections discuss the specifications and the data I use to test the hypotheses.

## 2.4.1 Specification

**Cross-Industry Tests** The first hypothesis to test is that the share of intrafirm imports is higher, the higher the capital intensity of the exporting industry. The model presented above actually predicts that the share should be 0 for industries with capital intensity $\beta_k$ below a certain threshold $\widehat{\beta}$ and 1 for industries with $\beta_k > \widehat{\beta}$, a prediction that does not seem to be borne by the data. Imagine, however, that the model provides a valid description of the world, but the statistician disaggregates the industry data under a criterion different from the one dictated by preferences or technology. In particular, instead of the sectors $Y$ and $Z$ in the model, the statistician disaggregates the data into $M$ industries. Denote by $n_Y^j(m)$ the measure of firms in country $j$ that produce varieties of intermediate inputs of sector $Y$ and that are included in industry $m \in M$ by the statistician. Let $n_Z^j(m)$ be the analogous measure for sector $Z$. The statistician will report an average capital intensity in industry $m$ and country $j$ equal to:[34]

$$k^j(m) = \frac{n_Y^j(m)K_Y + n_Z^j(m)K_Z}{n_Y^j(m)L_Y + n_Z^j(m)L_Z}. \tag{2.27}$$

On the other hand, letting $j$ be the rest of the world, the statistician will record U.S. imports in industry $m$ amounting to $s^{US}(n_Y^j(m)p_Y y + n_Z^j(m)p_Z z)$, of which $s^{US}n_Y^j(m)p_Y y$ will be reported as intrafirm imports and $s^{US}n_Z^j(m)p_Z z$ as imports at arm's length. The

---

[34]I drop the superscript $j$ when the theory dicates that a certain variable will be identical in all countries. This is the case, for instance, of $L_Y$, which in the model is given by equation (2.18).

share of intrafirm imports in industry $m$ will therefore be:

$$S_{i-f}^{US,j}(m) = \frac{n_Y^j(m)p_Y y}{n_Y^j(m)p_Y y + n_Z^j(m)p_Z z}. \qquad (2.28)$$

Combining equations (2.27) and (2.28) yields

$$S_{i-f}^{US,j}(m) = \frac{k^j(m) - K_Z/L_Z}{(1 - qL_Y/L_Z)k^j(m) + qK_Y/L_Z - K_Z/L_Z}, \qquad (2.29)$$

where $q = p_Z z/p_Y y$. It is straightforward to check that the recorded share of intrafirm imports in industry $m \in M$ is an increasing function of the recorded average capital intensity in that industry.[35] The model can thus deliver the smooth pattern in Figure 2-1.[36]

In the econometric results below, I report estimates from regressions of the form:

$$\ln \left( S_{i-f}^{US,ROW} \right)_m = \theta_1 + \theta_2 \ln (K/L)_m + W_m' \theta_3 + \epsilon_m, \qquad (2.30)$$

where $(S_{i-f}^{US,ROW})_m$ is industry $m$'s share of intrafirm imports in total U.S. imports from the rest of the world, $(K/L)_m$ is the average capital-labor ratio in the exporting industry, $W_m$ is a vector of controls, and $\epsilon_m$ is an error term, which is assumed to be orthogonal to the regressors. The vector $W_m$ is included to control for other possible industry-specific determinants of vertical integration. Since I observe the share $(S_{i-f})_m$ in four different years, I also include industry effects in some of the regressions. In light of Proposition 1, I hypothesize that $\theta_2 > 0$.

**Cross-Country Tests**  The second hypothesis that I test is that, in the cross-section of countries, the share of intrafirm imports in total imports is higher, the higher the capital-labor ratio of the exporting country. Equation (2.26) actually provides a closed-

---

[35]In particular, $S_{i-f}^{US,j}(m)$ increases with $k^j(m)$ as long as $K_Y/L_Y > K_Z/L_Z$, which is true in the model since $\widetilde{\beta_Y} > \widetilde{\beta_Z}$.

[36]As pointed out by a referee, the smooth pattern in Figure 2-1 could also be the result of idiosyncratic preferences by firms concerning the outsourcing decision. An explicit modelling of such firm-level heterogeneity would, however, greatly complicate the general-equilibrium analysis.

form solution for this relationship. Denoting the importing country by $US$ and the exporting country by $j$, and applying a log-linear approximation to (2.26) leads to the following specification:[37]

$$\ln \left( S_{i-f}^{US,j} \right) = \gamma_1 + \gamma_2 \ln \left( K^j / L^j \right) + \gamma_3 \ln \left( L^j \right) + W_j' \gamma_4 + \varepsilon_j, \tag{2.31}$$

where $S_{i-f}^{US,j}$ is the share of intrafirm imports in total U.S. imports from country $j$, $K^j/L^j$ is capital-labor ratio of country $j$, $W_j$ is a vector of controls, and $\epsilon_j$ is an orthogonal error term. The theory predicts that $\gamma_2$ should be positive. In fact, from the log-linearization, we can derive a much more precise prediction, i.e. $\gamma_2 = (1 - \sigma_L) \sigma_L / \left( 1 - \sigma_L - \widetilde{\beta_Z} \right)$. This implies that the elasticity of the share of intrafirm imports to the capital-labor ratio should not be lower than the labor share in the economy. Furthermore, from the last statement in Proposition 2, we should not expect $\gamma_3$ to be significantly different from zero.

The third test I conduct consists in running a regression analogous to (2.31) but with the log of *total* intrafirm imports (instead of its share in total imports) in the left-hand side. In particular, I consider the specification:

$$\ln \left( M_{i-f}^{US,j} \right) = \omega_1 + \omega_2 \ln \left( K^j / L^j \right) + \omega_3 \ln \left( L^j \right) + W_j' \omega_4 + \varepsilon_j. \tag{2.32}$$

In view of Lemma 4, both $\omega_2$ *and* $\omega_3$ should be positive. Furthermore, it is easy to show that the model imposes the restrictions $\omega_2 > \gamma_2$ and $\omega_3 = 1$.[38] In words, the total volume of intrafirm imports should be more responsive to the capital-labor ratio of the exporting country than its share in total imports, while its elasticity with respect to the size of the exporting country should be one.

---

[37]In particular, I log-linearize (2.26) around $K^j/L^j = K/L$, and obtain $\ln S_{i-f}^{US,j} \Big|_{K^j/L^j = K/L} \simeq \ln \left( \mu \right) + \frac{(1-\sigma_L)\sigma_L}{1-\sigma_L-\beta_Z} \left( \ln K^j / L^j - \ln K/L \right)$.

[38]A log-linear approximation of equation (2.25) around $K^j/L^j = K/L$ yields $\omega_2 = (1 - \sigma_L) \left( 1 - \widetilde{\beta_Z} \right) / \left( 1 - \sigma_L - \widetilde{\beta_Z} \right) > \gamma_2$ and $\omega_3 = 1$

### 2.4.2 Data

**Dependent Variables**  The left-hand side variables are constructed combining data on intrafirm U.S. imports and on overall U.S. imports. Intrafirm U.S. imports include (i) imports shipped by overseas affiliates to their U.S. parents; and (ii) imports shipped to U.S. affiliates by their foreign parent group. The series were obtained from the direct investment dataset available from the Bureau of Economic Activity (BEA) website. The BEA suppresses data cells in order to avoid disclosure of individual firm data. This limits the scope for testing the hypotheses of the paper in a fully satisfactory manner. For reasons discussed in Appendix 2.6.4, I end up running equation (2.30) for a panel consisting of 23 manufacturing industries and four years: 1987, 1989, 1992, and 1994. As for equations (2.31) and (2.32), data availability limits the analysis to a cross-section of 28 countries in 1992 (see Appendix 2.6.5 for a complete list of the industries and countries included in the regressions).

In the panel of industries, the share of intrafirm imports in total U.S. imports ranges from a value slightly below 1% for textiles in 1987 to around 82% for drugs in 1994, for an overall average of 21.2%. In the cross-section of countries, the share ranges from an almost negligible 0.1% for Egypt up to 64.1% for Switzerland, for an overall average of 22.4% (see Table 2.2).

Table 2.2 Share of Intrafirm Imports in total U.S. Imports (%)

| by Industry (avg. 1987-94) | | | | by Country (1992) | | | |
|------|------|------|------|------|------|------|------|
| DRU | 65.5 | FOO | 13.9 | CHE | 64.1 | ESP | 15.5 |
| OCH | 40.9 | PAP | 12.7 | SGP | 55.4 | AUS | 15.5 |
| VEH | 39.8 | FME | 12.6 | IRL | 53.7 | JPN | 14.2 |
| ELE | 37.3 | STO | 11.8 | CAN | 45.1 | ISR | 12.4 |
| COM | 36.7 | INS | 11.1 | NDL | 42.2 | HKG | 11.2 |
| CHE | 35.9 | TRA | 10.7 | MEX | 41.7 | PHL | 8.4 |
| CLE | 35.7 | PLA | 9.1 | PAN | 35.8 | ITA | 8.1 |
| RUB | 23.9 | PRI | 6.1 | GBR | 33.2 | ARG | 5.1 |
| AUD | 23.8 | LUM | 4.1 | DEU | 31.9 | COL | 4.6 |
| OEL | 18.9 | OMA | 2.6 | MYS | 30.1 | OAN | 4.6 |
| IMA | 17.3 | TEX | 2.3 | BEL | 27.3 | VEN | 1.4 |
| BEV | 15.1 | | | BRA | 25.9 | CHL | 1.3 |
| | | | | FRA | 21.6 | IDN | 1.3 |
| | | | | SWE | 16.8 | EGY | 0.1 |

Note: see Appendix 2.6.5 for a list of industries and countries.

**Industry Variables**   Most right-hand side variables in the cross-industry regressions are taken from the NBER Manufacturing Industry Productivity Database.[39] Capital intensity is measured as the ratio of the total capital stock to total employment in the corresponding exporting industry.[40] To control for other potential determinants of internalization, I run equation (2.30) including other industry characteristics one at a time.

---

[39]The SIC classification used in the NBER Manufacturing Database was converted to BEA industry categories using a concordance table available from the BEA and reproduced in Table A.1.

[40]This presupposes that U.S. industry capital intensities are similar to those in the rest of the world. In a world with factor price equalization this would naturally be the case. In a more general set-up, the much weaker assumption of no factor intensity reversals is sufficient to ensure that the same qualitative results would be obtained by using the source country factor intensity data.

First, I allow for the possibility that the integration decision might be determined by the human-capital intensity of the production process. To the extent that final-good producers also contribute to their suppliers' costs related to the acquisition of human capital (e.g., by financing training programs), a model identical to the one developed above with human capital $H$ replacing $K$ would indeed predict an effect of human-capital intensity. Furthermore, if physical capital and human-capital intensity are positively correlated, the patterns in Figure 2-1 would then be overstating the effect of capital intensity. I measure human-capital intensity as the ratio of nonproduction workers to production workers in a given industry, as reported in the NBER Manufacturing dataset. A completely analogous argument could be used to defend the inclusion of some measure of the importance of R&D and advertising in the production process. R&D intensity and advertising intensity are defined, respectively, as the ratio of R&D expenditures to sales and advertising expenditures to sales, and are obtained from a 1977 FTC survey.[41] I also control for the possibility that the integration decision may be driven by the size of scale economies at the plant level, as measured by average capital stock per establishment.[42] Finally, the decision to integrate could also be related to the importance of suppliers' production in the overall value chain. A rough way of proxying for this is to control for the share of value added in total industry sales, as reported in the NBER manufacturing dataset.

**Country Variables**   The main right-hand side variables in equations (2.31) and (2.32), including the capital-labor ratio of the exporting country and its total population, are taken from the cross-section of country variables for the year 1988 constructed by Hall and Jones (1999). In the present paper, I have adopted the view that capital abundance is a crucial determinant of the amount of multinational activity in a given country. Zhang and Markusen (2001) develop a model in which the volume of foreign direct investment in a given country is instead crucially affected by its skilled-labor abundance. To control for these possible effects, I include the measure of human capital abundance reported in Hall and Jones (1999). Other authors have stressed the importance of fiscal and

---

[41] This measure has been widely used in the literature (e.g., Cohen and Klepper, 1992, Brainard, 1997).

[42] This variable was constructed combining the capital stock figures from the NBER dataset with data on the number of establishments published by the U.S. Census Bureau in its County Business Patterns series.

institutional factors in determining the attractiveness of foreign direct investment in a given country. Countries with relatively lower corporate taxes and relatively better institutional environments should, in principle, be more prone to hosting affiliates of U.S. firms. In the regressions below, I use data on average corporate tax rates from a Price Waterhouse survey, as well as the index of institutional quality for the year 1990 reported in Gwartney et alt. (2002). Within the institutional factors, I also attempt to distinguish between the effect of a country's degree of openness to FDI and that of its degree of openess to international trade. Indices of openness to FDI and to trade are obtained from survey data reported in the World Competitiveness Report (1992). Table 2.3 reports descriptive statistics for all variables included in the regressions.

### 2.4.3  Results

**Intrafirm Trade and Factor Intensity**   The top panel of Table 2.4 presents random effects estimates of equation (2.30). Column I includes no controls in the regression and is therefore the econometric analog to Figure 2-1. The coefficient on $\ln(K/L)_m$ is positive and significantly different from zero at the 1%-significance level. The estimated elasticity of the share of intrafirm imports with respect to the capital-labor ratio in production implies that a 1% increase in $K/L$ increases the share of intrafirm imports by around 0.95%. Column II includes human-capital intensity in the regression. As expected, this leads to a reduction of the estimate of $\theta_2$, which, however, remains highly significant. The coefficient on $\ln(H/L)_m$ is positive but not statistically significant.

In column III, the ratio of R&D expenditures to sales is also included in the regression and is found to have a very significant effect on the share of intrafirm imports. The estimate of $\theta_2$ in column III is lower than that implied by Figure 2-1, but it still implies that a 1% increase in $K/L$, should lead to a 0.78% increase in the share of intrafirm imports. The inclusions of advertising intensity in column IV, of the size of economies of scale in column V, and of value-added intensity in column VI do not overturn any of the qualitative results. None of these variables seems to affect significantly the share of intrafirm imports, while capital intensity and R&D intensity remain significant at the 1% level.

Table 2.3. Descriptive Statistics

|  | Obs | Mean | St. dev. | Min | Max |
|---|---|---|---|---|---|
| $\ln\left(S_{i-f}^{US,ROW}\right)_m$ | 92 | -1.90 | 0.92 | -4.74 | -0.19 |
| $\ln(K/L)_m$ | 92 | 4.26 | 0.57 | 3.21 | 5.73 |
| $\ln(H/L)_m$ | 92 | -0.69 | 0.60 | -1.78 | 0.60 |
| $\ln(R\&D/Sales)_m$ | 92 | -4.20 | 1.00 | -6.07 | -2.47 |
| $\ln(ADV/Sales)_m$ | 92 | -4.27 | 1.10 | -6.63 | -2.24 |
| $\ln(Scale)_m$ | 92 | 1.63 | 0.92 | 0.06 | 3.48 |
| $\ln(VAD/Sales)_m$ | 92 | -0.66 | 0.18 | -1.13 | -0.32 |
| $\ln\left(S_{i-f}^{US,j}\right)$ | 28 | -2.08 | 1.44 | -6.67 | -0.45 |
| $\ln\left(K/L\right)_j$ | 28 | 10.54 | 0.86 | 8.13 | 11.59 |
| $\ln\left(L\right)_j$ | 28 | 16.03 | 1.20 | 13.63 | 18.16 |
| $\ln\left(H/L\right)_j$ | 28 | 0.82 | 0.19 | 0.47 | 1.10 |
| $CorpTax_j$ | 28 | 0.32 | 0.08 | 0.15 | 0.44 |
| $EconFreedom_j$ | 28 | 6.36 | 1.22 | 4.19 | 8.24 |
| $OpFDI$ | 26 | 7.83 | 1.23 | 4.73 | 9.57 |
| $OpTrade$ | 26 | 6.70 | 1.22 | 3.52 | 8.67 |
| $\ln\left(M_{i-f}^{US,j}\right)$ | 28 | 6.36 | 2.64 | -1.39 | 10.49 |

Table 2.4. Factor Intensity and the Share $S_{i-f}^{US,ROW}$

| Dep. var. is | Random Effects Regressions | | | | | |
|---|---|---|---|---|---|---|
| $\ln\left(S_{i-f}^{US,ROW}\right)_m$ | I | II | III | IV | V | VI |
| $\ln(K/L)_m$ | 0.947*** | 0.861*** | 0.780*** | 0.776*** | 0.703*** | 0.723*** |
| | (0.187) | (0.190) | (0.160) | (0.162) | (0.249) | (0.253) |
| $\ln(H/L)_m$ | | 0.369 | -0.002 | -0.038 | -0.037 | -0.081 |
| | | (0.213) | (0.188) | (0.200) | (0.206) | (0.221) |
| $\ln(R\&D/Sales)_m$ | | | 0.451*** | 0.470*** | 0.452*** | 0.421*** |
| | | | (0.107) | (0.114) | (0.128) | (0.140) |
| $\ln(ADV/Sales)_m$ | | | | 0.055 | 0.059 | 0.035 |
| | | | | (0.094) | (0.097) | (0.107) |
| $\ln(Scale)_m$ | | | | | 0.068 | 0.100 |
| | | | | | (0.179) | (0.190) |
| $\ln(VAD/Sales)_m$ | | | | | | 0.403 |
| | | | | | | (0.657) |
| $R^2$ | 0.50 | 0.55 | 0.72 | 0.73 | 0.73 | 0.73 |
| No. of obs. | 92 | 92 | 92 | 92 | 92 | 92 |

| | Fixed Effects Regressions | | | | | |
|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI |
| $\ln(K/L)_m$ | 0.599** | 0.610** | 0.610** | 0.610** | 0.943** | 1.058** |
| | (0.299) | (0.300) | (0.300) | (0.300) | (0.412) | (0.410) |
| p-value Wu-Hausman test | 0.14 | 0.27 | 0.62 | 0.64 | 0.52 | 0.19 |

Note: Standard errors in parenthesis (*, **, and *** are 10, 5, and 1% significance levels)

Consistency of the random effects estimates requires the industry effects to be uncorrelated with the other explanatory variables. One might worry that the omission of some relevant industry variables might lead to biases in the random effects estimates. As a robustness check, the bottom panel of Table 4 reports the fixed effects estimates of $\theta_2$ together with the p-value of a Wu-Hausman test for exogeneity of the industry effects.[43] The fixed effects estimates of $\theta_2$ are all significantly different from zero at the 5%-significance level. Furthermore, the point estimates are not too different from their random effects counterparts and the null hypothesis of exogeneity of the industry effects cannot be rejected at reasonable significance levels.

**Intrafirm Trade and Factor Abundance**   Table 2.5 reports OLS estimates of equation (2.31) for the cross-section of 28 countries. The estimates in column I correspond to the simple correlation depicted in Figure 2-2. The elasticity of the share of intrafirm imports with respect to the capital-labor ratio of the exporting country is significantly different from zero and, as predicted by the theory, the point estimate of the elasticity is necessarily higher than the average labor share in the economy. Column II confirms the claim in Proposition 2 that, for a given $K^j/L^j$, the size of the exporting country should not affect the share $S_{i-f}^{US,j}$. The coefficient of $\ln(L)_j$ is actually negative but statistically indistinguishable from zero. Column III introduces the measure of human-capital abundance in the regression. Contrary to what might have been expected (c.f., Zhang and Markusen, 2001), the estimated coefficient on $\ln(H/L)_j$ is negative, although again insignificantly different from zero. Conversely, the effect of physical-capital abundance remains significantly positive at the 1% level. As shown in column IV and V, controlling for the average corporate tax rate and the index of institutional quality does not overturn the results. The coefficients on both $CorpTax_j$ and on $EconFreedom_j$ are not significantly different from zero, while the estimate of $\gamma_2$ remains significantly positive at the 5% level. Finally, column VI suggests that the insignificance of the institutional variable in column V might be due to the counterbalancing effects of different policies. In particular, the share of intrafirm trade is negatively affected by the degree of openess to

---

[43]The R&D and advertising intensity variables are purely cross-sectional and are thus dropped in the estimation. This explains that the estimates in columns II, III and IV are all identical.

FDI but positively (although insignificantly) affected by the degree of openness to trade. Overall, the significant effect of the capital-labor ratio of the exporting country on the share of intrafirm imports appears to be very robust.[44]

Table 2.6 presents the OLS estimates of equation (2.32). Columns I and II confirm that the theoretical predictions in Lemma 4 are borne by the data. Both the capital-labor ratio of the exporting country and its size seem to have a significant positive effect on the volume of U.S. intrafirm imports. Consistently with the theory, the elasticity of $M_{i-f}^{US,j}$ with respect to $K^j/L^j$ is estimated to be higher than the elasticity of $S_{i-f}^{US,j}$ with respect to $K^j/L^j$. Furthermore, the elasticity of $M_{i-f}^{US,j}$ with respect to $L^j$ is, as predicted, not significantly different from one. As reported in columns III and IV, controlling for human capital abundance and for the average corporate tax rate has a negligible effect on the coefficients. The inclusion of the institutional index in column V leads to a substantial fall in the estimated elasticity of intrafirm imports to the capital-labor ratio, but the effect remains significant at the 5% level. Finally, column VI includes separate measures of openness to FDI and openness to trade. The results indicate that controlling for the capital-labor ratio of the exporting country, intrafirm imports are negatively affected by its openness to FDI.[45] More importantly, the effect of the capital-labor continues to be significant at the 1% level, while the effect of size is very close to being significant at the 10% level.

---

[44]Including $OpFDI$ and $OpTrade$ reduces the number of observation to 26, since no data on these variables is available for Egypt and Panama. I re-ran the regressions in columns I, through V, without these two countries and obtained very similar results.

[45]This may seem puzzling, but the model can shed light on this finding. Recall from section 2.2.3 that the attractiveness of integration is decreasing in the share $\phi$ of ex-post surplus accruing to final-good producers. If a higher openness to FDI corresponds to a larger bargaining power for foreign final-good producers, then *on this account* the model is consistent with the coefficient on $OpFDI$ being significantly negative. Note, however, that this is not the only effect of $\phi$ on the volume of intrafirm imports. From equation (2.25), $\phi$ also affects $M_{i-f}^{US,j}$ through the terms in $\widetilde{\beta_Y}$ and $\widetilde{\beta_Z}$ which are increasing in $\phi$, and through $\sigma_L$, which is decreasing in $\phi$. The overall effect of $\phi$ is in general ambiguous.

Table 2.5. Factor Endowments and the Share $S_{i-f}^{US,j}$

| Dep. var. is ln $\left(S_{i-f}^{US,j}\right)$ | I | II | III | IV | V | VI |
|---|---|---|---|---|---|---|
| ln $(K/L)_j$ | 1.141*** | 1.110*** | 1.244*** | 1.239*** | 1.097** | 1.119** |
| | (0.289) | (0.299) | (0.427) | (0.415) | (0.501) | (0.399) |
| ln $(L)_j$ | | -0.133 | -0.159 | -0.158 | -0.142 | 0.017 |
| | | (0.168) | (0.164) | (0.167) | (0.170) | (0.220) |
| ln $(H/L)_j$ | | | -1.024 | -0.890 | -1.273 | -0.822 |
| | | | (1.647) | (1.491) | (1.367) | (1.389) |
| $CorpTax_j$ | | | | -0.601 | 0.068 | 1.856 |
| | | | | (3.158) | (3.823) | (2.932) |
| $EconFreedom_j$ | | | | | 0.214 | |
| | | | | | (0.213) | |
| $OpFDI_j$ | | | | | | -0.384* |
| | | | | | | (0.218) |
| $OpTrade_j$ | | | | | | 0.292 |
| | | | | | | (0.273) |
| $R^2$ | 0.46 | 0.47 | 0.48 | 0.50 | 0.50 | 0.43 |
| No. of obs. | 28 | 28 | 28 | 28 | 28 | 26 |

Note: Robust standard errors in parenthesis (*, **, and *** are 10, 5, and 1% sig. levels)

Table 2.6. Factor Endowments and the volume $M_{i-f}^{US,j}$

| Dep. var. is $\ln\left(M_{i-f}^{US,j}\right)$ | I | II | III | IV | V | VI |
|---|---|---|---|---|---|---|
| $\ln\left(K/L\right)_j$ | 2.048*** | 2.192*** | 2.188*** | 2.154*** | 1.650** | 2.096*** |
| | (0.480) | (0.458) | (0.716) | (0.663) | (0.762) | (0.695) |
| $\ln\left(L\right)_j$ | | 0.607** | 0.608** | 0.614** | 0.670** | 0.700 |
| | | (0.229) | (0.268) | (0.271) | (0.243) | (0.419) |
| $\ln\left(H/L\right)_j$ | | | 0.031 | 0.953 | -0.406 | 0.708 |
| | | | (3.289) | (3.316) | (2.992) | (3.052) |
| $CorpTax_j$ | | | | -4.135 | -1.763 | -0.647 |
| | | | | (5.294) | (5.955) | (5.295) |
| $EconFreedom_j$ | | | | | 0.795 | |
| | | | | | (0.443) | |
| $OpFDI_j$ | | | | | | -1.006** |
| | | | | | | (0.474) |
| $OpTrade_j$ | | | | | | 0.674 |
| | | | | | | (0.560) |
| $R^2$ | 0.44 | 0.52 | 0.52 | 0.53 | 0.60 | 0.49 |
| No. of obs. | 28 | 28 | 28 | 28 | 28 | 26 |

Note: Robust standard errors in parenthesis (*, **, and *** are 10, 5, and 1% sig. levels))

## 2.5    Conclusions

This paper began by unveiling two systematic patterns in the intrafirm component of international trade. Traditional trade theory is silent on the boundaries of firms. Existing contributions to the theory of the firm tend to be partial-equilibrium in scope and have ignored the international dimensions of certain intrafirm transactions. Building on two workhorse models in international trade and the theory of the firm, I have developed a simple model that can account for the novel facts identified in the introduction. By combining a Grossman-Hart-Moore view of the firm with a Helpman-Krugman view of international trade, I have constructed a model that determines both the pattern of international trade and the boundaries of firms in a unified framework.

Nevertheless, much remains to be done. Future empirical investigations are likely to unveil new distinct features of the volume of intrafirm trade that cannot be accounted for by the simple model developed here. On the one hand, the Grossman-Hart-Moore theory enhances our understanding of only a subset of the determinants of ownership structure. Holmström and Milgrom (1994) have emphasized that, in many situations, issues related to job design and the cost of measuring performance are more relevant when choosing between inside or outside procurement.[46] It would be interesting to investigate the implications of such a view of the firm for the volume of intrafirm trade. On the other hand, in determining trade patterns, I have resorted to a very simple trade model. Future work should help us to understand potential channels by which technological differences, transport costs or international factor-price differences can affect the organization and location of international production.

---

[46]Baker and Hubbard (2002) find that ownership patterns in the trucking industry reflect the relevance of each of these two strands in the literature.

## 2.6 Appendix

### 2.6.1 Proof of Lemma 1

Combining equations (2.9) and (2.10), it follows that regardless of the level of demand $A_Y$ the final-good producer in a pair of stand-alone firms will decide to incur the capital expenditures itself whenever:

$$(1 - \alpha(1 - \beta_Y) + \alpha\phi(1 - 2\beta_Y)) \left(\frac{\phi}{1 - \phi}\right)^{\frac{\alpha\beta_Y}{1-\alpha}} > \phi + (1 - \alpha)(1 - \phi),$$

which holds whenever $\phi > 1/2$. To see this, define the function

$$H(\phi) = (1 - \alpha(1 - \beta) + \alpha\phi(1 - 2\beta)) \left(\frac{\phi}{1 - \phi}\right)^{\frac{\alpha\beta_Y}{1-\alpha}} - \phi - (1 - \alpha)(1 - \phi),$$

and notice first that $H(1/2) = 0$. Next note that

$$H'(\phi) = \alpha\left(\left(\frac{\phi}{1 - \phi}\right)^{\frac{\alpha\beta}{1-\alpha}} - 1\right) + \alpha\beta\left(\frac{\phi}{1 - \phi}\right)^{\frac{\alpha\beta}{1-\alpha}}\left(\frac{1 - \alpha(1 - \beta) + \alpha\phi(1 - 2\beta)}{(1 - \alpha)(1 - \phi)\phi} - 2\right).$$

The first term is clearly positive when $\phi > 1/2$. Furthermore, since $\frac{(1-\alpha(1-\beta)+\alpha\phi(1-2\beta))}{(1-\alpha)(1-\phi)\phi}$ increases with $\alpha$, it follows that $\frac{(1-\alpha(1-\beta)+\alpha\phi(1-2\beta))}{(1-\alpha)(1-\phi)\phi} - 2 \geq \frac{1}{(1-\phi)\phi} - 2 > 0$ and the second term is also positive. Hence, $H(\phi) > 0$ for all $\phi > 1/2$. Since $\overline{\phi} > \phi$, as long as $\phi > 1/2$, final-good producers in integrated pairs will also decide to rent the capital stock and hand it to the supplier. QED.

### 2.6.2 Proof of Lemma 2

From simple differentiation of (2.11), it follows that $\partial\Theta(\cdot)/\partial\beta_k > 0$ if and only if

$$\Omega(\beta_k)\ln\left(1 + \frac{\delta^\alpha}{\phi(1 - \delta^\alpha)}\right) > (2 - \alpha)(1 - \alpha)(1 - \phi)\delta^\alpha$$

where $\Omega(\beta_k) = \left(1 - \alpha(1 - \overline{\phi}) + \alpha\beta_k(1 - 2\overline{\phi})\right)(1 - \alpha(1 - \phi) + \alpha\beta_k(1 - 2\phi))$ and remember that $\overline{\phi} = \delta^\alpha + \phi(1 - \delta^\alpha)$. Now notice that if $\overline{\phi} > \phi \geq 1/2$ then $\Omega'(\beta_k) < 0 \ \forall \beta_k \in [0, 1]$, and if $\phi < \overline{\phi} \leq 1/2$, then $\Omega'(\beta_k) > 0 \ \forall \beta_k \in [0, 1]$. Furthermore, if $\overline{\phi} > 1/2 > \phi$, then $\Omega''(\beta_k) < 0 \ \forall \beta_k \in [0, 1]$. It thus follows that $\Omega(\beta_k) \geq \min\{\Omega(0), \Omega(1)\}$. Without loss of generality, assume that $\Omega(1) = (1 - \alpha\phi)(1 - \alpha(\phi + (1 - \phi)\delta^\alpha)) < \Omega(0)$ (the case $\Omega(1) > \Omega(0)$ is entirely symmetric). We need to show that $\vartheta(\delta) > 0$ for all $\delta \in (0, 1)$ where

$$\vartheta(\delta) = \ln\left(1 + \frac{\delta^\alpha}{\phi(1 - \delta^\alpha)}\right) - \frac{(2 - \alpha)(1 - \alpha)(1 - \phi)\delta^\alpha}{(1 - \alpha\phi)(1 - \alpha(\phi + (1 - \phi)\delta^\alpha))}$$

From simple differentiation of this expression, it follows that $\vartheta'(\delta) > 0$ if and only if $(1 - \alpha\rho)^2 - (2 - \alpha)(1 - \alpha)(1 - \rho)\rho > 0$ for some $\rho \in (0, 1)$. But it is simple to check that this is in fact true all $\alpha, \rho \in (0, 1)$, and therefore $\vartheta(\delta) > \vartheta(0) = 0$. Notice that Assumption 1 is not necessary for this result. QED.

### 2.6.3   Proof of Proposition 1

From equation (2.11) and the definition of $\overline{\phi}$, note that we can write

$$\Theta(0, \cdot) = \frac{1 - \alpha(1 - \overline{\phi})}{1 - \alpha(1 - \phi)}\left(\frac{1 - \overline{\phi}}{1 - \phi}\right)^{\frac{\alpha}{1 - \alpha}} < 1$$

and

$$\Theta(1, \cdot) = \frac{1 - \alpha\overline{\phi}}{1 - \alpha\phi}\left(\frac{\overline{\phi}}{\phi}\right)^{\frac{\alpha}{1 - \alpha}} > 1.$$

The inequalities follow from $\overline{\phi} > \phi$ and the fact that $(1 - \alpha x)x^{\frac{\alpha}{1 - \alpha}}$ is an increasing function of $x$ for $\alpha \in (0, 1)$ and $x \in (0, 1)$. The rest of the Proposition is a direct implication of Lemma 2. QED.

### 2.6.4   Data Appendix

This Appendix discusses in more detail the construction of the share of intrafirm imports in total U.S. imports. Intrafirm imports were obtained from the "Financial and Operating Data" on multinational firms downloadable from the BEA website. Since in the model

70

ownership is associated with control, I restricted the sample to majority-owned affiliates. As discussed in the main text, the BEA suppresses data cells in order to avoid disclosure of individual firm data. The unsuppressed data is only available to researchers affiliated to the BEA. Unfortunately, one of the requirements for affiliation is being a U.S. citizen (which I am not).

To construct intrafirm imports by industry, I combine data from foreign affiliates of U.S. firms and U.S. affiliates of foreign firms. Intrafirm imports comprise (i) imports shipped by overseas affiliates to their U.S. parents, by industry of affiliate, and; (ii) imports shipped to U.S. affiliates by their foreign parent group, by industry of affiliate.[47,48]. The sum of these two elements was constructed at the finest level of disaggregation available, focusing on manufacturing industries and excluding natural-resource industries (in particular, petroleum, ferrous metals and non-ferrous metals).[49] I also restricted the sample to years in which benchmark surveys were conducted. Overall, I end up with 23 industries and four years: 1987, 1989, 1992 and 1994.

To construct intrafirm imports by country, I add up (i) imports shipped by overseas affiliates to their U.S. parents, by country of origin, and (ii) imports shipped to U.S. affiliates by their foreign parent group, by country of origin. In both cases, I restrict the analysis to manufacturing industries, although in this case it was impossible to remove those transactions involving natural resources (this might explain why intrafirm imports from Chile and Venezuela are lower than predicted in Figure 2-2). The BEA performs two types of manipulations to the data. Apart from suppressing cells to avoid disclosure of data of individual companies, it also assigns a unique symbol to trade flows inferior in value to $500,000. I assign a value of $250,000 to these cells.[50] Overall, I end up with a

---

[47] The BEA defines a foreign parent group as consisting of (1) the foreign parent, (2) any foreign person, proceeding up the foreign parent's ownership chain, that owns more than 50 percent of the person below it, up to and including the ultimate beneficial owner, and (3) any foreign person, proceeding down the ownership chain(s) of each of these members, that is owned more than 50 percent by the person above it.

[48] The conceptually correct disaggregation for case (ii) would have been by the industry of the exporter (i.e. of the foreign parent group). Unfortunately, these series are not available. Intrafirm imports of type (i) constitute, however, more than two-thirds of all intrafirm imports. More importantly, a similar pattern to that in Figure 2-1 emerges when the analysis is restricted to intrafirm imports of type (i).

[49] Patterns of ownership in natural-resource sectors are likely to be determined by factors, such as national sovereignty, from which I abstract in the model.

[50] This is only done for two observations. The results are robust to imputing alternative values between

single cross-section with 28 countries in 1992. All the other benchmark survey years lack at least one of the components of intrafirm imports.

Finally, in order to compute the share of intrafirm imports, I construct total U.S. imports by industry and year, and then by country of origin, using data put together by Robert Feenstra and available from the NBER website. Import figures correspond to their c.i.f. values. Feenstra's four-digit industry classification was matched to the 23 BEA industries using a conversion table available from BEA and reproduced in Table A.1.

As pointed out by a referee, a significant portion of intrafirm trade involves final goods that are shipped from a manufacturing plant to an overseas wholesale affiliate which then distributes the good in the foreign country. Unfortunately, the BEA dataset does not distinguish between imports of intermediate inputs and imports of final goods, so that the latter cannot be substracted from intrafirm imports. As pointed out by the same referee, however, this is not necessarily a problem for the empirical work as an analogous theoretical model can also be interpreted in the context of a supplier-distributor relationship.

---

0 and \$500,000.

## 2.6.5 Additional Tables

Table A.1. Industry Description and Classification

| Code | Description | Corresponding Industry SIC Classification |
|------|-------------|-------------------------------------------|
| BEV | Beverages | 208 |
| FOO | Other food and kindred products | 201-207, 209 |
| CHE | Industrial chemicals and synthetics | 281, 282, 286 |
| DRU | Drugs | 283 |
| CLE | Soap, cleaners and toilet goods | 284 |
| OCH | Other chemical products | 285, 287, 289 |
| FME | Fabricated metal products | 341-349 |
| COM | Computer and office equipment | 357 |
| IMA | Other industrial machinery and equipment | 351-356, 358, 359 |
| AUD | Audio, video and communications equipment | 365, 366 |
| ELE | Electronic components and accessories | 367 |
| OEL | Other electronic and electrical machinery | 361-364, 369 |
| TEX | Textile products and apparel | 221-229, 231-39 |
| LUM | Lumber, wood, furniture and fixtures | 241-49, 251-59 |
| PAP | Paper and allied products | 261-263, 265, 267 |
| PRI | Printing and publishing | 271-279 |
| RUB | Rubber products | 301, 302, 305, 306 |
| PLA | Miscellaneous plastics products | 308 |
| STO | Stone, clay, and glass product | 321-29 |
| VEH | Motor vehicles and equipment | 371 |
| TRA | Other transportation equipment | 372-376, 379 |
| INS | Instruments and related products | 381, 382, 384-387 |
| OMA | Other manufacturing | 211-19, 311-19, 391-99 |

Table A.2. Country Codes

| Code | Country | Code | Country |
|------|---------|------|---------|
| ARG | Argentina | IDN | Indonesia |
| AUS | Australia | IRL | Ireland |
| BEL | Belgium | ISR | Israel |
| BRA | Brazil | ITA | Italy |
| CAN | Canada | JPN | Japan |
| CHE | Switzerland | OAN | Taiwan |
| CHL | Chile | PAN | Panama |
| COL | Colombia | PHL | Philippines |
| DEU | Germany | MEX | Mexico |
| EGY | Egypt | MYS | Malaysia |
| ESP | Spain | NDL | Netherlands |
| FRA | France | SGP | Singapore |
| GBR | United Kingdom | SWE | Sweden |
| HKG | Hong Kong | VEN | Venezuela |

# Chapter 3

# Incomplete Contracts and the Product Cycle

## 3.1  Introduction

In an enormously influential article, Vernon (1966) described a natural life cycle for the typical commodity. Most new goods, he argued, are initially manufactured in the country where they are first developed, with the bulk of innovations occurring in the industrialized, high-wage North. Only when the appropriate designs have been worked out and the production techniques have been standardized is the locus of production shifted to the less developed South, where wages are lower. Vernon emphasized the role of multinational firms in the international transfer of technology. In his formulation of a product's life cycle, the shift of production to the South is a profit-maximizing decision from the point of view of the innovating firm.

The "product cycle hypothesis" soon gave rise to an extensive empirical literature that searched for evidence of the patterns suggested by Vernon.[1] The picture that has emerged from this literature has turned out to be much richer than originally envisioned by Vernon. The evidence indeed supports the existence of product cycles, but it has become clear that foreign direct investment by multinational firms is not the only vehicle of production

---

[1]See Gruber et alt. (1967), Hirsch (1967), Wells (1969), and Parry (1975) for early tests of the theory.

transfer to the South. In particular, the literature has identified several instances in which technologies have been transferred to the South through licensing, subcontracting, and other similar arm's length arrangements. More interestingly, several studies have pointed out that the choice between intrafirm and market transactions is significantly affected by both the degree of standardization of the technology and the transferor's resources devoted to product development.[2] In particular, overseas assembly of relatively new and unstandardized products tends to be undertaken within firm boundaries, i.e. through foreign direct investment, while innovators seem more willing to resort to licensing and subcontracting in standardized goods with little product development requirements.

The product cycle hypothesis has also attracted considerable attention among international trade theorists eager to explore the macroeconomic and trade implications of Vernon's insights. Krugman (1979) developed a simple model of trade in which new goods are produced in the industrialized North and exchanged for old goods produced in the South. In order to concentrate on the effects of the product cycle on trade flows and relative wages, Krugman (1979) specified a very simple form of technological transfer, with new goods becoming old goods at an exogenously given Poisson rate. This "imitation lag," as he called it, was later endogenized by Grossman and Helpman (1991a,b) using the machinery developed by the endogenous growth literature. In particular, Grossman and Helpman (1991a,b) developed a model in which purposeful innovation and imitation gave rise to endogenous product cycles, with the timing of production transfer being a function of the imitation effort exerted by firms in the South. As the empirical literature on the product cycle suggests, however, the bulk of technology transfer is driven by *voluntary* decisions of Northern firms, which choose to undertake offshore production within firm boundaries or transact with independent subcontractors or licensees.[3]

---

[2]See, for instance, Davidson and McFetridge (1984, 1985), Mansfield et alt. (1979), Mansfield and Romeo (1980), Vernon and Davidson (1979), and Wilson (1977). These studies will be discussed in more detail later in the paper.

[3]Grossman and Helpman (1991b) claimed that their model generated realistic predictions about the evolution of market shares and the pattern of trade in a given industry. Indeed, as they point out, purposeful imitation by low-wage countries has been an important driving force in the transfer of production of microprocessors from the United States and Japan to Taiwan and Korea. Based on recent studies, I will argue below, however, that even in the case of the electronics industry, the spectacular increase in the market share of Korean exports is better explained by technology transfer by foreign-based firms than by simple imitation by domestic Korean firms.

In this paper, I provide a theory of the product cycle that is much more akin to Vernon's (1966) original formulation and that delivers implications that are very much in line with the findings of the empirical literature discussed above. In the model, goods are produced combining a hi-tech input, which I associate with product development, and a low-tech input, which is meant to capture the simple assembly or manufacturing of the good. As in Grossman and Helpman (1991a,b), the North is assumed to have a high enough comparative advantage in product development so as to ensure that this activity is always undertaken there. My specification of technology differs, however, in one important respect from that in Grossman and Helpman (1991a,b). In particular, I treat product development as a continuously active sector along the life cycle of a good. The concept of product development used here is therefore quite broad and is meant to include, among others, the development of ideas for improving existing products, marketing and advertising. Following Vernon (1966), this specification of technology enables me to capture the standardization process of a good along its life cycle. More specifically, I assume that the contribution of product development to output (as measured by the output elasticity of the hi-tech input) is inversely related to the age or maturity of the good. Intuitively, the initial phases of a product life cycle entail substantial testing and re-testing of prototypes as well as considerable marketing efforts to make consumers aware of the existence of the good. As the good matures and production techniques become standardized, the mere assembly of the product becomes a much more significant input in production.

Following Vernon (1966) and contrary to Grossman and Helpman (1991a,b), I allow Northern firms to split the production process internationally and transact with manufacturing plants in the South.[4] With no frictions to the international fragmentation of the production process, I show that the model fails to deliver a product cycle. Intuitively,

---

[4]There is a recent literature in international trade documenting an increasing international disintegration of the production process. A variety of terms have been used to refer to this phenomenon: the "slicing of the value chain", "international outsourcing", "vertical specialization", "global production sharing", and many more. Feenstra (1998) discusses the widely cite example of Nike. In 1994, Nike employed around 2,500 U.S. workers in management, design, sales, and promotion, while leaving manufacturing in the hands of some 75,000 workers in Asia. Interestingly, Nike subcontracts most parts of its production process, so that the production plants in Asia are not Nike affiliates and their 75,000 workers are not Nike employees.

provided that labor is paid a lower wage in the South than in the North, manufacturing will be shifted to the South even for the most unstandardized, *product-development intensive* goods. Vernon (1966) was well aware that his theory required some type of friction that delayed offshore assembly. In fact, he argued that in the initial phase of a product's life cycle, overseas production would be discouraged by a low price elasticity of demand, the need for a thick market for inputs, and the need for swift and effective communication between producers and suppliers.

This paper will instead push the view that what limits the international fragmentation of the production process is the incomplete nature of contracts governing international transactions. Building on the seminal work of Williamson (1985) and Grossman and Hart (1986), I show that the presence of incomplete contracts creates a hold-up problem, which in turn gives rise to suboptimal relationship-specific investments by the parties involved in a transaction. The product development manager of a Northern firm can alleviate this type of distortions by keeping the manufacturing process in the North, where contracts can be better enforced. In choosing between domestic or overseas manufacturing, the product development manager therefore faces a trade-off between the lower costs of Southern manufacturing and the distortions coming from incomplete contracting, which are shown to affect both the manufacturing *and* the product development inputs in production. This trade-off is shown to lead naturally to the emergence of product cycles: when the good is new and unstandardized, Southern production is very unattractive because it bears the full cost of incomplete contracting with little benefit from the lower wage in the South. Conversely, when the good is mature and requires very little product development, the benefits from lower wages in the South fare much better against the distortions from incomplete contracting, and the good is produced in the South.

The model focuses first on the profit-maximizing choice of location by a single Northern product development manager. I then embed this choice in a general-equilibrium, dynamic Ricardian model of North-South trade with a continuum of industries that standardize at different rates. The model solves for the timing of production transfer for any given industry, as well as for the time path of the relative wage in the two countries. In spite of the rich heterogeneity in industry product-cycle dynamics, the cross-sectional picture that emerges from the model is very similar to that in the Ricardian model

with a continuum of goods of Dornbusch, Fischer and Samuelson (1977). In contrast to the exogenous cross-industry and cross-country productivity differences in their model, comparative advantage arises here from a combination of the Northern productivity advantage in product development, the continuous standardization of goods, and the fact that contracts are incomplete.

Following the lead of Krugman (1979), I study the effect of an acceleration of technological change in the North on the world distribution of income. I show that Krugman's result that increased technological change widens the wage differential greatly depends on technological change taking the form of the introduction of new products into the economy. If, instead, technological change takes the form of an increase in the rate at which goods standardize, I show that the converse result is true and relative wages move in favor of the South. I next analyze the welfare implications of a shift from a steady-state equilibrium with incomplete contracts to a steady-state equilibrium with complete contracts. This improvement in the contracting environment in international transactions is shown to unambiguously increase welfare in the South, while having an ambiguous effect on Northern welfare.

Following the property-rights approach to the theory of the firm (Grossman and Hart, 1986, Hart and Moore, 1990), the same force that creates product cycles, i.e. incomplete contracts, opens the door to a parallel analysis of the determinants of ownership structure. As in Grossman and Hart (1986), I associate ownership with the entitlement of some residual rights of control. When parties undertake noncontractible, relationship-specific investments, the allocation of residual rights has a critical effect on each party's *ex-post* outside option, which in turn determines each party's *ex-ante* incentives to invest. Ex-ante efficiency (i.e., transaction-cost minimization) then dictates that residual rights should be controlled by the party whose investment contributes most to the value of the relationship. As a result, a new version of the product cycle emerges. If the threshold *product-development intensity* at which manufacturing is shifted to the South is high enough, production transfer will occur within the boundaries of the Northern firm by establishing a wholly-owned foreign affiliate. In such case, at a later stage in the product cycle, the product development manager finds it optimal to give away these residual rights of control, and the assembly will be assigned to an independent subcontractor in the

South, an arrangement which is analogous to the Northern firm licensing its technology (hi-tech input). For a lower *product-development intensity* threshold, i.e., for a higher maturity of the good at the time of the transfer, the model predicts that the transfer to the South will occur directly at arm's length, and multinationals will not arise.

Solving for the general equilibrium with multinational firms, I show that, relative to a world in which only arm's length or licensing transactions are allowed, the introduction of intrafirm production transfer by multinational firms accelerates the shift of production towards the South, while having an ambiguous effect on relative wages. Furthermore, provided that its effect on relative wages is small enough, the emergence of multinational firms is shown to be welfare improving for both countries. I finally discuss how the equilibrium timing of production transfer, as well as the choice between intrafirm and arm's length production transfer, are affected by several parameters in the model and discuss the empirical validity of these predictions.

The model is useful for understanding the evolution of the Korean electronics industry after the Korean War.[5] In the early 1960s, Korean electronic firms were producing mostly low-quality consumer electronics for their domestic market. The industry took off in the late 1960s with the establishment of a few large U.S. assembly plants, almost all wholly owned, followed in the early 1970s by substantial Japanese investments. These foreign subsidiaries tended to assemble components exclusively for export using imported parts.[6] In this initial phase, foreign affiliates were responsible for 71% of exports in electronics, with the percentage reaching 97% for the case of exports of integrated circuits and transistors, and 100% for memory planes and magnetic heads. In the 1970s and 1980s domestic Korean firms progressively gained a much larger market share, but the strengthening of domestic electronic companies was accompanied by a considerable expansion of technology licensing from foreign firms. Indeed, as late as 1988, 60% of Korean electronic exports were recorded as part of an Original Equipment Manufacturing (OEM) transaction.[7] The percentage approached 100% in the case of exports of computer ter-

---

[5] The following discussion is based on Bloom (1992), UNCTAD (1995, pp. 251-253), and Cyhn (2002).

[6] Motorola established a production plant in Korea in 1968. Other U.S.-based multinationals establishing subsidiaries in Korea during this period include Signetics, Fairchild and Control Data.

[7] OEM is a form of subcontracting which as Cyhn's (2002) writes "occurs when a company arranges for an item to be produced with its logo or brand name on it, even though that company is not the

minals and telecommunications equipment. Korean giants such as Samsung or Goldstar were heavily dependent on foreign licenses and OEM agreements even up to the late 1980s.

There is also some evidence that Northern firms did not license their leading edge technologies to their Korean licensees. For instance, in 1986, Hitachi licensed to Goldstar the technology to produce the 1-megabyte Dynamic Random Access Memory (DRAM) chip, when at the same time it was shifting to the 4-megabyte DRAM chip. Similarly, Phillips licensed the production of compact disk players to ten Korean producers, while keeping within firm boundaries the assembly of their deck mechanisms.

The rest of the paper is structured as follows. Section 2 focuses on a simple dynamic model that shows how the presence of incomplete contracts gives rise to product cycles. For simplicity, I initially abstract from the choice of ownership structure by allowing only arm's length production transfer to the South. In section 3, I embed this simple model in a general-equilibrium model of North-South trade and study the effects of incomplete contracting on relative wages and the speed of technology transfer. In section 4, I allow for intrafirm technology transfer and describe the richer product life cycle that emerges from it, both in partial and in general equilibrium. I discuss several implications of the model and relate them to the empirical literature on the product cycle. Section 5 concludes. The proofs of the main results are relegated to the Appendix.

## 3.2   Incomplete Contracts and the Life Cycle of a Product

This section describes a simple framework in which a product development manager has to decide how to organize production of a particular good, taking the behavior of other producers as well as wages as given. I will first analyze the static problem, and then show how a product cycle emerges in a dynamic extension in which the good gets standardized over time.

producer".

### 3.2.1 Set-up

**Environment**   There are two countries, the North and the South, and a single good $y$ produced only with labor. I denote the wage rate in the North by $w^N$ and that in the South by $w^S$. Consumer preferences are such that the unique producer of good $y$ faces the following iso-elastic demand function:

$$y = \lambda p^{-1/(1-\alpha)}, \qquad 0 < \alpha < 1 \tag{3.1}$$

where $p$ is the price of the good and $\lambda$ is a parameter that the producer takes as given.[8]

Production of good $y$ requires the development of a special and distinct hi-tech input $x_h$, as well as the production of a special and distinct low-tech input $x_l$. As discussed in the introduction, the hi-tech input is meant to comprise research and product development, marketing, and other similar skill-demanding tasks. The low-tech input is instead meant to capture the mere manufacturing or assembly of the good. The specialized inputs can be of good or bad quality. If any of the two inputs is of bad quality, the output of the final good is zero. If both inputs are of good quality, production of the final good requires no additional inputs and output is given by:

$$y = \zeta_z x_h^{1-z} x_l^z, \qquad 0 \le z \le 1, \tag{3.2}$$

where $\zeta_z = z^{-z} (1-z)^{-(1-z)}$.

The unit cost function for producing the hi-tech input varies by country. In the North, production of each unit of a good-quality, hi-tech input requires the employment of one unit of Northern labor. The South is much less efficient at producing the hi-tech input. In fact, the productivity advantage of the North is assumed large enough to ensure that $x_h$ is only produced in the North. Meanwhile, production of one unit of good-quality low-tech input also requires labor, but the unit input requirement is assumed to be equal to 1 for both countries. Finally, production of any type of bad-quality input can be undertaken at a positive but negligible cost. Both the hi-tech and the low-tech inputs

---

[8]This demand function will be derived from preferences in the general-equilibrium model.

are freely tradable.

**Firm Structure:** There are two types of producers: a research center and a manufacturing or assembly plant. A research center is defined as the producer of the hi-tech input. It follows that the research center will always locate in the North. The research center needs to contract with an independent manufacturing plant for the provision of the low-tech input. I allow for an international fragmentation of the production process, so that the research center can choose to transact with an independent assembly plant in the North or with one in the South. The timing of events is as follows. Before any investment is made, a research center decides whether to produce a hi-tech input, and if so, whether to match with an assembly plant in the North, or match with an assembly plant in the South. The manufacturer makes, upon entry, a lump-sum transfer $T$ to the research center.[9] I assume that ex-ante there is large number of identical, potential manufacturers of the good, so that competition among them makes $T$ adjust so as to make the chosen manufacturer break even. The research center chooses the location of manufacturing to maximize its ex-ante profits.

Investments are assumed to be relationship-specific. The research center tailors the hi-tech input specifically to the assembly plant, while the low-tech input is customized according to the specific needs of the research center. In sum, the investments in labor needed to produce $x_h$ and $x_l$ are incurred upon entry and are useless outside the relationship.

**Contracting environment:** The setting is one of incomplete contracts in situations of international production sharing. In particular, it is assumed that only when both inputs are produced in the same country can an outside party distinguish between a good-quality and a bad-quality intermediate input.[10] Hence, the manager of the research center and

---

[9]If the final good is produced by the manufacturing plant, this transfer $T$ can be interpreted as licensing fee for the use of the hi-tech input.

[10]This can be interpreted as a physical constraint imposed on the outside party if, for instance, in order to observe the quality of the input, the outside party needs to supervise its actual production process. More generally, this assumption is meant to capture broader contractual difficulties in international transactions, such as ambiguous jurisdiction, language conflicts, or, more simply, weak protection of property rights in low-wage countries.

that of a Southern assembly plant cannot sign an enforceable contract specifying the purchase of a certain type of intermediate input for a certain price. If they did, the party receiving a positive payment would have an incentive to produce the bad-quality input at the lower cost.
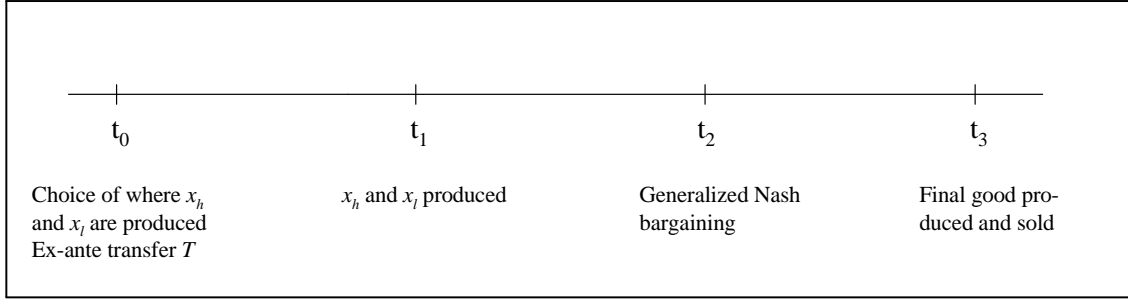
It is equally assumed that no outside party can verify the amount of ex-ante investments in labor. If these were verifiable, the managers could contract on them, and the cost-reducing benefit of producing a bad-quality input would disappear. For the same reason, it is assumed that the parties cannot write contracts contingent on the volume of sale revenues obtained when the final good is sold. The only contractible ex-ante is the transfer $T$ between the parties.[11]

When the research center chooses to transact with an assembly plant in the North, the fact that labor investments are not contractible is irrelevant because the parties can always appeal to an outside party to enforce quality-contingent contracts. In contrast, when manufacturing is done by a plant in the South, no enforceable contract will be signed ex-ante and the two parties will bargain over the surplus of the relationship after the inputs have been produced. At this point, the quality of the inputs (and therefore also the ex-ante investments) are observable to both parties and thus the costless bargaining will yield an ex-post efficient outcome. I assume that Generalized Nash Bargaining leaves the research center with a fraction $\phi \in (0,1)$ of the surplus. Since the inputs are tailored specifically to the other party in the transaction, if the two parties fail to agree on a division of the surplus, both are left with nothing.

As I will show below, the fact that contracts are incomplete will lead to underinvestment in both product development and manufacturing when the production process is fragmented internationally. Furthermore, the underprovision of product development will be more severe the lower is $\phi$, i.e., the lower the bargaining power of the research center manager. For some of the results below, it is important that product development

---

[11] I take the fact that contracts are incomplete as given. From the work of Aghion, Dewatripont and Rey (1994), Nöldeke and Schmidt (1995) and others, it is well known that allowing for specific-performance contracts can lead, under certain circumstances, to efficient ex-ante relationship-specific investments. Che and Hausch (1997) have shown, however, that when ex-ante investments are cooperative (in the sense, that one party's investment benefits the other party), specific-performance contracts may not lead to first-best investment levels and may actually have no value.

Figure 3-1: Timing of Events



| $t_0$ | $t_1$ | $t_2$ | $t_3$ |

Choice of where $x_h$ and $x_l$ are produced Ex-ante transfer $T$ — $x_h$ and $x_l$ produced — Generalized Nash bargaining — Final good produced and sold

is sufficiently distorted. This is ensured by the following assumption:

**Assumption 1:** $\phi \leq 3/4$.

If the transaction between the research center and the manufacturing plant is interpreted as a licensing arrangement, this assumption is consistent with available evidence. As Caves (1996) writes, *"the empirical evidence on licensing convincingly shows that licensors on average can appropriate less than half of the surplus associated with the license transaction"* (p. 167).

This completes the description of the model. The timing of events is summarized in Figure 3-1.

## 3.2.2 Firm Behavior

As discussed above, the North has a sufficiently high productivity advantage in producing the hi-tech input to ensure that $x_h$ is produced there. On the other hand, the decision of where to produce the low-tech input is nontrivial. In his choice, the manager of the research center compares ex-ante profits from two options, which I analyze in turn.

**A. Assembly in the North by an Independent Manufacturing Plant**

Consider first the case of a research center who decides to deal with an independent manufacturing plant in the North. In such case, the two parties can write an ex-ante quality-contingent contract that will not be renegotiated ex-post. The initial contract

stipulates production of good-quality inputs in an amount that maximizes the research center's ex-ante profits, which from equations (3.1) and (3.2), and taking account of the transfer $T$, are given by $\pi^N = \lambda^{1-\alpha} \zeta_z^\alpha x_h^{\alpha(1-z)} x_l^{\alpha z} - w^N x_h - w^N x_l$. It is straightforward to check that this program yields an optimal price of

$$p^N(z) = \frac{w^N}{\alpha}. \tag{3.3}$$

Because the research center faces a constant elasticity of demand, the optimal price is equal to a constant mark-up over marginal cost. Ex-ante profits for the research center are equal to

$$\pi^N(z) = (1-\alpha)\lambda \left(\frac{w^N}{\alpha}\right)^{-\alpha/(1-\alpha)}. \tag{3.4}$$

### B. Assembly in the South by an Independent Manufacturing Plant

Consider next the problem faced by a research center that decides to transact with an assembly plant in the South. As discussed above, in this case the initial contract stipulates only the transfer $T$. The game played by the manager of the research center and that of the assembly plant is solved by starting at $t_3$ and moving backwards. If both producers make good-quality intermediate inputs and the firms agree in the bargaining, the potential revenues from the sale of the final good at $t_3$ are $R = \lambda^{1-\alpha} \zeta_z^\alpha x_h^{\alpha(1-z)} x_l^{\alpha z}$. In contrast, if the parties fail to agree in the bargaining, both are left with nothing. The quasi-rents of the relationship are therefore equal to sale revenues, i.e., $R$.

The contract incompleteness will give rise to renegotiation at $t_2$. In the bargaining, Generalized Nash bargaining leaves the research center manager with a fraction $\phi$ of the quasi-rents, while the assembly plant manager in the South receives the remaining fraction $1 - \phi$ of the quasi-rents. Since $\phi$ is assumed to be in $(0, 1)$, both parties' ex-post revenues from producing a good-quality input are always strictly positive. Bad-quality inputs will therefore never be produced.

Rolling back to $t_1$, the research center manager will therefore set $x_h$ to maximize $\phi R - w^N x_h$. The assembly plant will simultaneously choose $x_l$ to maximize $(1 - \phi) R - w^S x_l$. Combining the first-order conditions of these two programs yields the following optimal

price for the final good:

$$p^S(z) = \frac{(w^N)^{1-z}(w^S)^z}{\alpha \phi^{1-z}(1-\phi)^z}.$$  (3.5)

It is straightforward to show that if contracts were also complete for international transactions, the research center would instead set a price equal to $(w^N)^{1-z}(w^S)^z/\alpha$. Incomplete contracting therefore inflates the optimal mark-up by a factor that is inversely related to $\phi^{1-z}(1-\phi)^z$. If $z$ is low, the distortion is relatively higher, the lower is $\phi$. Conversely, if production of $y$ requires mostly the low-tech input ($z$ high), the distortion is relatively higher, the higher is $\phi$.

Setting $T$ so as to make the assembly plant break even leads to the following expression for the research center's ex-ante profits:

$$\pi^S(z) = (1 - \alpha(\phi(1-z) + (1-\phi)z))\lambda \left( \frac{(w^N)^{1-z}(w^S)^z}{\alpha \phi^{1-z}(1-\phi)^z} \right)^{-\alpha/(1-\alpha)}.$$  (3.6)

Comparing equations (3.4) and (3.6), it is possible to show that the low-tech input will be produced in the South only if $\omega \equiv w^N/w^S > 1$. Intuitively, Southern assembly becomes profitable only when wages in the South are low enough to offset the distortions coming from incomplete contracting. By a continuity argument, if $\omega$ is close enough to one, it is also necessarily the case that $\pi^N(z) > \pi^S(z)$ for all $z \in [0,1]$, and production in the South would never be chosen either. To make matters more interesting, I assume that $\omega$ is high enough so that $\pi^N(z_c) < \pi^S(z_c)$ for some $z_c \in (0,1)$. Comparing equations (3.4) and (3.6) this condition can be rewritten as:

**Condition 1:** There exists a $z_c \in (0,1)$ such that $A(z_c) < \omega$, where

$$A(z) \equiv \left( \frac{1-\alpha}{1 - \alpha(\phi(1-z) + (1-\phi)z)} \right)^{(1-\alpha)/\alpha z} \frac{\phi^{-(1-z)/z}}{1-\phi}.$$  (3.7)

This condition will in fact be shown to necessarily hold in the general-equilibrium model and this is why I avoid labelling it as an assumption. The intuition for why Condition 1 will hold true in the general equilibrium is that the relative wage in the North will necessarily adjust to ensure positive labor demand in the South.

### 3.2.3  The Equilibrium Choice

The main result from this section will be that production of the low-tech input (i.e., manufacturing) will be kept in the North if and only if production of the final-good is sufficiently *product-development intensive*. To gain intuition on the result, consider first the case in which the mode of organization were chosen to minimize the optimal price (notice that revenues $py = \lambda p^{-\alpha/(1-\alpha)}$ are decreasing in $p$). In such case, assembly would be done by the research center in the North if $p^N(z) < p^S(z)$, a condition that can be rewritten as

$$(\omega(1-\phi))^z < \left(\frac{1}{\phi}\right)^{1-z}. \tag{3.8}$$

The left-hand side of the inequality corresponds to the benefits of locating assembly in the South. These benefits are higher, the higher the relative wage in the North and the lower is $\phi$. Intuitively, the higher is $\omega$, the larger will the cost reduction of manufacturing in the South be, while the lower is $\phi$, the larger will the Southern manager's bargaining power be, and therefore the less severe will his underinvestment in $x_l$ be. The right-hand side of the inequality captures instead the costs of Southern assembly and correspond to the distortions caused by incomplete contracting on the production of the hi-tech input in the North. Because the hold-up problem is two-sided, the research center will also underinvest in product development $(x_h)$, with the degree of underinvestment decreasing in the Northern manager's bargaining power $(\phi)$.

It is clear from equation (3.8) that if $\omega(1-\phi) < 1$, then $p^N(z) < p^S(z)$ will hold for all $z \in [0,1]$ and production in the South will never be chosen. This is the type of situations that Condition 1 rules out. On the other hand, if $\omega(1-\phi) > 1$, the benefits from Southern assembly are increasing in $z$, while the costs are decreasing in $z$. Intuitively, the more important is the hi-tech input in the production process (the lower $z$), the more costly are distortions in product development, and the less weight is carried by the lower assembly wages in the South. Furthermore, for $z = 0$, the costs necessarily outweigh the benefits, and conversely for $z = 1$. It follows that there exists a unique $z_p$ for which $p^N(z_p) = p^S(z_p)$, and such that $p^N(z) < p^S(z)$ if and only if $z < z_p$.

When the mode of organization is chosen to maximize profits rather than to minimize the price set (or maximize revenues), the analysis is complicated by the fact that

Figure 3-2: The Choice of Location
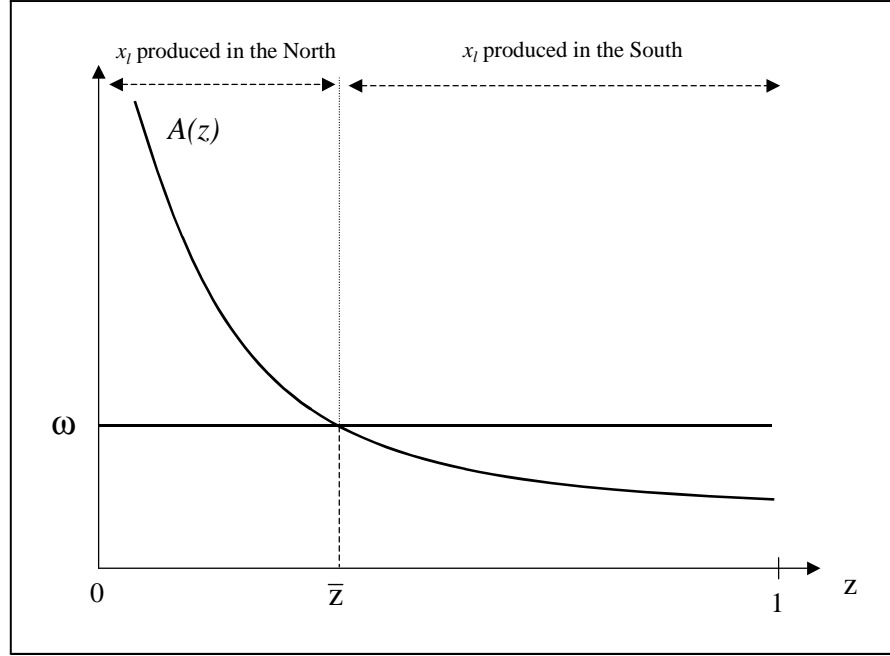


incomplete contracting not only distorts the mark-up over marginal cost, but also the fraction of revenues that producers are able to capture as operating profits. Provided that Assumption 1 ($\phi \leq 3/4$) holds, however, a result analogous to that under price minimization applies:

**Lemma 5** *Under Assumption 1 and Condition 1, there exists a unique threshold $\bar{z} \in (0,1)$ such that the low-tech input is produced in the North if $z < \bar{z} \equiv A^{-1}(\omega)$, while it is produced in the South if $z > \bar{z} \equiv A^{-1}(\omega)$, where $A(z)$ is given by equation (3.7) and $\omega$ is the relative wage in the North.*

    **Proof.** *See Appendix 3.6.1.* ∎

Figure 3-2 depicts the profit-maximizing choice of location as a function of $z$. Under Assumption 1, the $A(z)$ curve is decreasing in $z$ for all $z \in [0,1]$, while by Condition 1, $A(1) < \omega$. Therefore there exists a unique $\bar{z}$ such that $A(z) > \omega$ for $z < \bar{z}$, and $A(z) < \omega$ for $z > \bar{z}$. As shown in Figure 3-2, $x_l$ is produced in the North for $z < \bar{z}$, while it is produced in the South for $z > \bar{z}$.

From direct inspection of Figure 3-2, it is clear that an increase in the relative wage in the North reduces the threshold $\bar{z}$. Intuitively, an increase in $\omega$ makes Southern manufacturing relatively more profitable, and this leads to a reduction in the measure of hi-tech intensities for which the whole production process stays in the North.

An increase in the relative bargaining power of the research center can be shown to rotate the $A(z)$ curve in a counterclockwise direction about some $z_\phi \in (0, 1)$. It follows that for a high enough $\omega$, an increase in $\phi$ leads to a fall in $\bar{z}$, with the converse result applying for low enough $\omega$. The intuition behind this result will be discussed at length in section 3.4, when I introduce multinational firms.

### 3.2.4 Dynamics: The Product Cycle

As discussed in the introduction, one of the premises of Vernon's (1966) original product-cycle hypothesis is that as a good matures throughout its life cycle, it becomes more and more standardized.[12] Vernon believed that the unstandardized nature of new goods was crucial for understanding that they would first be produced in a high-wage country, such as the United States. To capture this standardization process in a simple way, consider the following simple dynamic extension of the static model developed above. Time is continuous, index by $t$, with $t \in [0, \infty)$. Consumers are infinitely lived and, at any $t \in [0, \infty)$, their preferences for good $y$ are captured by the demand function (3.1).

On the technology and contracting sides, the parameters $\alpha$ and $\phi$, as well as the relative wage $\omega$, are assumed to be time-invariant.[13] The output elasticity of the low-tech input is instead assumed to increase through time. In particular, this elasticity is given by

$$z(t) = h(t), \text{ with } h'(t) > 0, \ h(0) = 0, \text{ and } \lim_{t \to \infty} h(t) = 1. \tag{3.9}$$

---

[12]In discussing some previous empirical studies on the location of industry, Vernon wrote: "in the early stages of introduction of a new good, producers were usually confronted with a number of critical, albeit transitory, conditions. For one thing, the product itself may be quite unstandardized for a time; its inputs, its processing, and its final specifications may cover a wide range. Contrast the great variety of automobiles produced and marketed before 1910 with the thoroughly standardized product of the 1930's, or the variegated radio designs of the 1920's with the uniform models of the 1930's." (Vernon, 1966, p.197).

[13]The latter assumption will be relaxed in the general-equilibrium model, where $\omega$ will be endogenized.

I therefore assume that the product-development intensity of the good is inversely related to its maturity. This is meant to capture the above discussion indicating that most goods require a lot of R&D and product development in the early stages of their life cycle, whereas the mere assembling or manufacturing becomes a much more significant input in production as the good matures. I will take these dynamics as given, but it can be shown that, under Assumption 1 and Condition 1, profits for the Northern research center are weakly increasing in $z$. It follows that the smooth process of standardization specified here could, in principle, be derived endogenously in a richer framework that incorporated some costs of standardization.[14]

Finally, I assume that the firm structure is such that when Southern assembly is chosen, the game played by the two managers can be treated as a static one and we can abstract from an analysis of reputational equilibria. This is a warranted assumption when the separation rate for managers is high enough or when they discount enough future profit streams.[15]

With this simplified, dynamic set-up, the cut-off level $\bar{z} \equiv A^{-1}(\omega)$ is time-invariant, and the following result is a straightforward implication of Lemma 5:

**Proposition 3** *The model displays a product cycle. When the industry is young, i.e.* $t < h^{-1}(\bar{z})$, *the low-tech input is produced in the North. When the good is mature or standardized, i.e.* $t \geq h^{-1}(\bar{z})$, *manufacturing is shifted to the South.*

Consider, for instance, the following specification of the standardization process:

$$z(t) = h(t) = 1 - e^{-t/\theta},$$

where $1/\theta$ measures the rate at which $1-z$ falls towards zero, i.e. the rate of standardization. With this functional form, the whole production process remains in the North until the product reaches an age equal to $\theta \ln\left(\frac{1}{1-\bar{z}}\right)$, at which point manufacturing is shifted to the South. Naturally, production of the low-tech input is transferred to the South earlier,

---

[14]For instance, if such costs were increasing in $dz/dt$, then a discrete increase in $z$ would be infinitely costly.

[15]The general-equilibrium model developed below will in fact incorporate positive, endogenous separation rates.

the higher is the speed of standardization, $1/\theta$, and the lower is the threshold intensity $\bar{z}$. Furthermore, I showed above that $\bar{z}$ is itself a decreasing function of $\omega$, from which it follows that the higher is the relative wage in the North, the earlier will production transfer occur.[16]

**Comparison with the Case of Complete Contracts**   So far I have shown that the model delivers a product cycle. I now briefly demonstrate my claim in the introduction that contract incompleteness is in fact necessary for the product cycle to emerge. For this purpose, consider the case in which the quality of intermediate inputs was also verifiable by an outside court in international transactions, so that the manager of the research center and that of the Southern assembly plant could write enforceable contracts. It is straightforward to check that, in such case, profits for the research center would be:

$$\pi^S(z) = (1 - \alpha)\,\lambda \left( \frac{\left(w^N\right)^{1-z}\left(w^S\right)^z}{\alpha} \right)^{-\alpha/(1-\alpha)}$$

Comparing this expression with equation (3.4), it follows that labor demand in the South could be positive if and only if $\omega \geq 1$ (this is the analog of Condition 1 above). Next, notice that as long as $\omega > 1$, it will necessarily be the case that $\pi^N(z) \leq \pi^S(z)$ for all $z \in [0,1]$, with strict inequality for $z > 0$. It follows that the low tech-input will never be produced in the North, and the production process will be broken up from time 0. A product cycle therefore does not arise if $\omega > 1$. On the other hand, if $\omega = 1$, then it follows that $\pi^N(z) = \pi^S(z)$ for all $z \in [0,1]$ and the location of assembly remains indeterminate. In such case, product cycles emerge with probability zero.[17]

---

[16]Vernon (1966) hypothesized instead that before being transferred to low-wage countries, production would first be located in middle-income countries for a period of time. An important point to notice is that in doing the comparative statics with respect to $\omega$, I have held the contracting environment constant. In the real world, a country's degree of contract incompleteness seems to be negatively correlated with its wage rate. The empirical growth literature consistently finds that countries with better legal systems grow faster (Mauro, 1995, Knack and Keefer, 1995) and have higher per-capita income (Hall and Jones, 1999, Acemoglu et alt. 2001). If I allowed for this type of correlation in the model, production would not necessarily be transferred earlier the higher $\omega$.

[17]This shows that, in the otherwise frictionless framework I developed, incomplete contracts are necessary for product cycles to emerge. This is not to say that in incomplete contracting is the crucial driving force behind product cycles in the model. Arguably, the process of standardization is more important for

## 3.3    Incomplete Contracts and the Product Cycle in General Equilibrium

In this section, the partial-equilibrium model developed above is embedded in a dynamic, general-equilibrium framework with varieties in different sectors standardizing at different rates. I will first solve for the time-path of the relative wage in the two countries and show that the equilibrium wage in the North is necessarily higher than that in the South. Next, I will study some macroeconomic and welfare implications of this new view of the product cycle.

### 3.3.1    Set-up

The North is endowed with $L^N$ units of labor at any $t \in (0, \infty)$, while the Southern endowment is also constant and equal to $L^S$. At each period $t$, there is a measure $N(t)$ of industries indexed by $j$, each producing an endogenously determined measure $n_j(t)$ of differentiated goods. Preferences of the infinitely-lived representative consumer in each country are given by:

$$U = \int_0^\infty e^{-\rho t} \int_0^{N(t)} \log \left( \int_0^{n_j(t)} y_j \left( i, t \right)^\alpha di \right)^{1/\alpha} dj dt, \qquad (3.10)$$

where $\rho$ is the rate at which the consumer discounts future utility streams. Notice that all industries are viewed as symmetric with a unitary elasticity of substitution between them. The varieties of differentiated goods also enter symmetrically into (3.10), but with an elasticity of substitution equal to $1/(1 - \alpha) > 1$. Because the economy has no means of saving and preferences are time-separable, the consumer maximizes utility period by period and the discount rate plays no role in the model (other than to make the problem bounded). As is well known, the instantaneous utility function in (3.10) gives rise to a

---

obtaining the results. Incomplete contracting is just one of the many potential frictions that would make manufacturing stay in the North for a period of time. I choose to emphasize here the role of incomplete contracts because I believe that they are an important source of frictions in the real world and, more importantly, because they are a very useful theoretical tool for understanding firm boundaries, which are the focus of section 3.4 below.

constant price-elasticity of demand for any variety $i$ in any industry $j$:

$$y_j\left(i,t\right) = \lambda_j(t)p_j\left(i,t\right)^{-1/(1-\alpha)}, \tag{3.11}$$

where

$$\lambda_j(t) = \frac{1}{N(t)}\frac{E\left(t\right)}{\int_0^{n_j(t)} p_j(i',t)^{-\alpha/(1-\alpha)}di'} \tag{3.12}$$

and $E\left(t\right)$ is total world spending in period $t$. Because firms take $\lambda_j(t)$ as given, each producer of a final-good variety faces a demand function analogous to that in equation (3.1) above.

At each point in time, production of each final-good variety is also as described in the partial-equilibrium model above, with the additional assumption that, at every period $t$, production of each variety also requires a fixed cost of $f$ units of labor in the country where the hi-tech input is produced (i.e., the North). It is assumed that all producers in a given industry share the same technology as specified in (3.2), with a common time-varying elasticity $z_j(t - t_{0j})$, where $t_{0j}$ is the date at which industry $j$ was born. As before, I assume $z_j'(\cdot) > 0$, $z_j(0) = 0$, and $\lim_{t-t_{0j}\to\infty} z_j(t - t_{0j}) = 1$. In words, varieties in a given industry $j$ are produced for the first time at $t_{0j}$ using only the hi-tech input, and then all standardize at a common rate.

Notice that industries may vary not only in their birth dates, but also in the shape of their specific $z_j(\cdot)$ functions. To isolate the effect of cross-industry differences in maturity and in standardization rates, I assume that the technology for producing intermediate inputs, as well as fixed costs, are identical across industries and varieties.

Firm structure is as described above, with the additional assumption that there is free entry at every period $t$, so that the measure $n_j\left(t\right)$ of varieties in each industry always adjusts so as to make research centers break even. The lack of profits in equilibrium is implied by the fact that technology is a function of the industry's age and not of the age of the producer of a particular variety. Furthermore, as in section 3.2.4, I assume that firm structure is such that when Southern procurement is chosen, the game played by the two managers can be treated as a static one and we can abstract from an analysis of reputational equilibria. This assumption, coupled with the absence of means of saving

94

in the model, permits a period-by-period analysis of the dynamic, general-equilibrium model.

The contracting environment is also analogous to that of the partial-equilibrium model and, in particular, the parameter $\phi$ is time-invariant and common for all varieties and industries.

I consider an economy in which exogenous inventions continuously adds to the stock of existing industries. In particular, I let $\dot{N}(t) = gN(t)$ and $N(0) = N_0 > 0$. Hence, in any period $t$ there are $N(t) = N_0 e^{gt}$ industries producing varieties of final goods.

### 3.3.2 Industry Equilibrium

Consider now the industry equilibrium at any $t \in [0, \infty)$. The unit elasticity of substitution between varieties in different industries implies that we can analyze firm behavior in each industry independently. Consider then any industry $j$. Facing the same technology and contracting environment, all research centers in the same industry will necessarily set the same price and therefore will earn the same profits. It follows that letting again $\bar{z}(t) \equiv A^{-1}(\omega(t))$, the low-tech input will be produced in the North if $z_j(t - t_{0j}) < \bar{z}(t)$, and in the South if $z_j(t - t_{0j}) > \bar{z}(t)$, with the choice remaining indeterminate for $z_j(t - t_{0j}) = \bar{z}(t)$.

In order to characterize the industry equilibrium, notice first that since all firms in a given industry $j$ will charge an identical optimal price $p(z_j(\cdot))$, equation (3.12) simplifies to:

$$
\lambda_j(t) = \begin{cases} \dfrac{E(t)}{N(t)n_j(t)} \left( \dfrac{w^N(t)}{\alpha} \right)^{\alpha/(1-\alpha)} & \text{if } z_j(t - t_{0j}) < \bar{z}(t) \\[2em] \dfrac{E(t)}{N(t)n_j(t)} \left( \dfrac{\left(w^N(t)\right)^{1-z_j(t-t_{0j})}\left(w^S(t)\right)^{z_j(t-t_{0j})}}{\alpha\phi^{1-z_j(t-t_{0j})}(1-\phi)^{z_j(t-t_{0j})}} \right)^{\alpha/(1-\alpha)} & \text{if } z_j(t - t_{0j}) > \bar{z}(t) \end{cases}
$$

(3.13)

The equilibrium number of varieties produced in industry $j$ at time $t$ is computed by plugging equation (3.12) in the profit functions (3.4) and (3.6), and setting operating

profits equal to fixed costs, as dictated by free entry. This yields:

$$n_j(t) = \begin{cases} \frac{(1-\alpha)E(t)}{N(t)w^N(t)f} & \text{if } z_j(t-t_{0j}) < \bar{z}(t) \\ \\ \frac{(1-\alpha(\phi(1-z_j(t-t_{0j}))+(1-\phi)z_j(t-t_{0j})))E(t)}{N(t)w^N(t)f} & \text{if } z_j(t-t_{0j}) > \bar{z}(t) \end{cases} \quad . \quad (3.14)$$

Notice that in a stationary equilibrium in which both $E(t)$ and $w^N(t)$ are constant through time and $N(t)$ grows at rate $g$, the measure of varieties in a given industry falls at rate $g$ for $z_j(t-t_{0j}) < \bar{z}(t)$. The evolution of the number of varieties when $z_j(t-t_{0j}) > \bar{z}(t)$ is more complicated and depends on the relative growth rates of $z_j(t-t_{0j})$ and $N(t)$, as well as on whether $\phi$ is greater or lower than $1/2$.

### 3.3.3 General Equilibrium

Having described the equilibrium for a particular industry at time $t$, we can now move to the general equilibrium, in which income equals spending

$$w^N(t)L^N + w^S(t)L^S = E(t) \tag{3.15}$$

and the labor market in each country clears. By Walras' law, we can focus on the equilibrium of the labor market in the South. Southern labor will only be demanded by those assembly plants belonging to an industry with $z_j(t-t_{0j}) > \bar{z}(t)$. Labor demand by each assembly plant in the South can be shown to equal $L_l^S = \alpha(1-\phi)z_j(\cdot)\lambda_j(t)p_j^S(t)^{-\alpha/(1-\alpha)}/w^S(t)$. Denote by $F_{z,t}(z)$ the *fraction* of industries with $z_j(t-t_{0j}) < \bar{z}(t)$ at time $t$. Letting $f_{z,t}(z)$ be the corresponding probability density function, the Southern labor-market clearing condition can be expressed as:

$$\int_{\bar{z}(t)}^1 \alpha(1-\phi)zE(t)f_{z,t}(z)dz = w^S(t)L^S. \tag{3.16}$$

Defining $\xi_t(a,b) \equiv \int_a^b z f_{z,t}(z)dz$ and using (3.15), equation (3.16) can be then rewritten as:

$$\omega(t) = B_t(\bar{z}(t)) \equiv \frac{1-\alpha(1-\phi)\xi_t(\bar{z}(t),1)}{\alpha(1-\phi)\xi_t(\bar{z}(t),1)}\frac{L^S}{L^N}. \tag{3.17}$$

96

$B_t(\bar{z}(t))$ is an increasing function of $\bar{z}(t)$ satisfying $B_t(0) > 0$ and $\lim_{\bar{z}(t) \to 1} B_t(\bar{z}(t)) = +\infty$. Intuitively, the higher is $\bar{z}(t)$, the lower is labor demand in the South for a given $\omega(t)$. An increase in $\omega(t)$ is necessary to bring us back to equilibrium. When $\bar{z}(t)$ goes to 1, labor demand in the South goes to 0, and the required relative wage goes to $+\infty$. On the other hand, since the North always produces the hi-tech input, even when $\bar{z}(t)$ goes to 0 labor demand in the North is positive and $B_t(0)$ is greater than 0. Figure 3-3 depicts the curve $B_t(\cdot)$ in the $(z, \omega)$ space.

The other equilibrium condition that pins down $\bar{z}(t)$ and $\omega(t)$ comes from the partial equilibrium. In particular, since $\alpha$ and $\phi$ are common across industries, $\bar{z}(t)$ is also implicitly defined by

$$\omega(t) = A(\bar{z}(t)) \equiv \left( \frac{1 - \alpha}{1 - \alpha \left( \phi \left( 1 - \bar{z}(t) \right) + \left( 1 - \phi \right) \bar{z}(t) \right)} \right)^{(1-\alpha)/\alpha \bar{z}(t)} \frac{\phi^{-(1-\bar{z}(t))/\bar{z}(t)}}{1 - \phi}. \quad (3.18)$$

As discussed above, under Assumption 1, $A(\bar{z}(t))$ is a decreasing function of $\bar{z}(t)$ satisfying $\lim_{\bar{z}(t) \to 0} A(\bar{z}(t)) = +\infty$ and $A(1) > 1$.[18] The function $A(\cdot)$ is also depicted in Figure 3-3. As it apparent in Figure 3-3, there exists a unique equilibrium pair $(\bar{z}(t), \omega(t))$ for each $t$. Furthermore, the fact that $A(1)$ is greater than 1, ensures that in equilibrium the wage in the North is higher than that in the South, or $\omega(t) > 1$.

The general equilibrium of the dynamic model is simply the sequence of period-by-period general equilibria. In addition, if the distribution function $F_{z,t}(z)$ is (or converges to) a time-invariant distribution, the equilibrium values for $\bar{z}(t)$ and $\omega(t)$ will also be time-invariant. In such case, all industries will necessarily follow product cycles, with varieties in those industries being manufactured first in the North and then later in the South.[19]
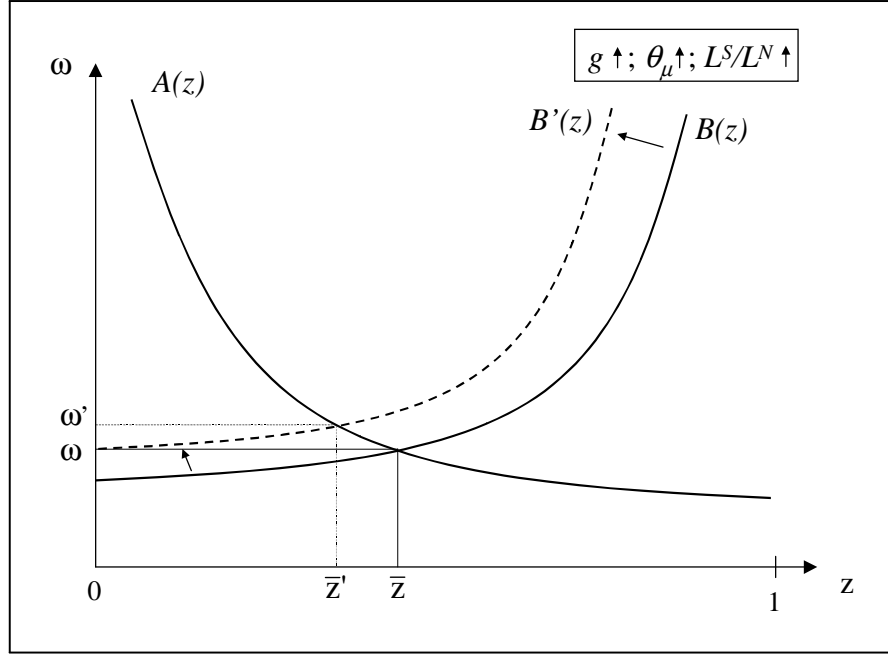
Consider again the particular functional form:

$$z_j(t, t_{0j}) = 1 - e^{-(t-t_{0j})/\theta_j}, \quad (3.19)$$

---

[18]In particular, $A(1) = \left( \frac{1-\alpha}{1-\alpha(1-\phi)} \left( 1 - \phi \right)^{-\alpha/(1-\alpha)} \right)^{(1-\alpha)/\alpha}$ which is greater than one because $(1 - \alpha x) x^{\alpha/(1-\alpha)}$ is increasing in $x$ for $\alpha \in (0, 1)$ and $x \in (0, 1)$.

[19]In fact, a much weaker assumption is need for product cycles to emerge, namely, that the growth rate of $\bar{z}(t)$ be lower than that of $z_j(t - t_{0j})$ for all industries $j$.

Figure 3-3: General Equilibrium



so that the elasticity of output with respect to $x_h$ falls at a constant rate $1/\theta_j$. As before, I will refer to $1/\theta_j$ as industry $j$'s specific rate of standardization. From the discussion in section 3.2.4, and given that the threshold $\bar{z}(t)$ is common across industries, the model predicts that industries with a higher rates of standardization will transfer manufacturing earlier to the South. In the general equilibrium, however, the cross-industry distribution of standardization rates will have an additional effect on the timing of production transfer, through its impact on the overall distribution of hi-tech intensities as given by $F_{z,t}(z)$. To see this, assume that $\theta_j$ is independent from $t_{0j}$ and is exponentially distributed with mean $\theta_\mu$, i.e., $F_\theta(\theta_j) = 1 - e^{-\theta_j/\theta_\mu}$. Notice that since $N(t)$ grows at constant rate $g$, the distribution of birth dates in the economy converges to the c.d.f. $F_{t_0}(t_0, t) \equiv N(t_0)/N(t) = e^{-g(t-t_0)}$. Under these assumptions, Appendix 3.6.2 shows that $F_{z,t}(z)$ converges to a time-invariant distribution function characterized by:

$$F_z(z) = \frac{g\theta_\mu \ln\left(\frac{1}{1-z}\right)}{1 + g\theta_\mu \ln\left(\frac{1}{1-z}\right)}. \tag{3.20}$$

98

Computing the corresponding density function, it follows that $\xi_t(a,b)$ converges to

$$\xi(a,b) \equiv \int_a^b z f_z(z) dz = \int_a^b \frac{zg\theta_\mu}{\left(1 + g\theta_\mu \ln\left(\frac{1}{1-z}\right)\right)^2 (1-z)} dz. \qquad (3.21)$$

In light of equations (3.17) and (3.21), the economy converges to a steady state in which the $B_t(\bar{z}(t))$ curve is time-invariant and increasing in $\bar{z}(t)$. In such a steady state, the general equilibrium values of $\bar{z}$ and $\omega$ are unique and time-invariant.

It is interesting to notice that in spite of the rich heterogeneity in industry product-cycle dynamics, the cross-sectional picture that emerges from the model is very similar to that in the classical Ricardian model with a continuum of goods of Dornbusch, Fischer and Samuelson (1977). In contrast to their model, however, comparative advantage arises here from a combination of the Northern productivity advantage in product development, the continuous standardization of goods, and the fact that contracts are incomplete.

### 3.3.4 Comparative Statics

Consider first an increase in the relative size of the South, i.e. an increase in $L^S/L^N$ (see Figure 3-3). As in Dornbusch, Fischer and Samuelson (1977), this shifts the $B(\cdot)$ curve up and to the left leading to an increase in the relative wage in the North and a fall in $\bar{z}$, which in turn implies that all industries will shift manufacturing to the South at an earlier phase of their life cycle. The intuition behind this result is as follows. At the initial relative wage, the increase in the Southern relative labor creates an excess supply of labor in the South and excess demand for labor in the North. An increase in $\omega$ and a fall in the measure of hi-tech intensities for which manufacturing is done in the North are necessary to restore equilibrium.

With the particular functional form in equation (3.19), an increase in the exogenous rate of invention $(g)$ or a fall in the rate of standardization $(1/\theta_\mu)$ also shift the $B(\cdot)$ curve up and to the left and thus have completely analogous effects on the equilibrium values of $\omega$ and $\bar{z}$.[20] As it is apparent from equation (3.20), an increase in $g$ or $\theta_\mu$ tends to

---

[20]In Appendix 3.6.2, I show that both $\partial\xi(\cdot)/\partial g < 0$ and $\partial\xi(\cdot)/\partial\theta_\mu < 0$.

transfer probability mass from high $z$'s to low $z$'s, which tends to create again an excess supply (demand) of labor in the South (North) at the initial $\omega$ and $\bar{z}$. In sum,

**Proposition 4** *Holding $\theta_j$ and $t_{0j}$ constant, the relative wage in the North is higher and the shift to Southern assembly occur earlier: (i) the higher the rate of invention $g$, (ii) the lower the average rate of standardization $1/\theta_\mu$, (iii) the larger the relative population size of the South $L^S/L^N$.*

The comparative statics on the relative wage are related to those obtained by Krugman (1979) in the first paper to explore the macroeconomic implications of the product cycle hypothesis. Krugman (1979) developed a simple model with two types of goods: new goods and old goods. In his model, exogenous innovation adds to the stock of new goods which, by assumption, can only be produced in the North. New goods become old goods at an exogenous "imitation" rate, at which point they start being manufactured in the low-wage South. As in Proposition 4 above, in his model too an increase in the rate at which new goods appear in the North or an increase in the relative size of the South lead an excess demand for Northern labor which needs to be matched by an increase in the relative wage in the North to restore equilibrium. Krugman (1979) concluded from his comparative static results that increased technological change in the North redistributed income from the South to the North. Proposition 4 demonstrates that the validity of this claim depends very much on the assumption that technological change takes the form of adding new products to the economy. If instead it takes the form of an increase in the rate at which goods standardize (an increase in $1/\theta_\mu$ in the model) the converse result applies.

By endogenizing the timing of technology transfer, the present model delivers additional implications of a shift in the parameters $g$ and $\theta_\mu$. For instance, an increase in the rate of invention also leads to a reduction in the time it takes for manufacturing of a particular good to be shifted to the South, which in Krugman's (1979) model depends only on the exogenous rate of imitation.

Furthermore, in light of Proposition 4, the industry-specific and the average rates of standardization have opposite effects on the timing of technology transfer. Although, the manufacturing of those goods that become standardized relatively faster will be

transferred to the South relatively earlier, the model predicts that for a given $\theta_j$, an increase in the average rate of standardization will in fact slow production transfer of varieties in industry $j$. This result may help shed light on the evidence of a fall in the interval of time between the introduction of a new product in the United States and its first production in a foreign location (c.f., Vernon, 1979). The partial equilibrium model would suggest a simple explanation for this pattern: the technologies transferred in the recent past (e.g., computer parts) get standardized much faster than those transferred in the 1950s and 1960s (e.g. television sets). Nevertheless, Proposition 4 indicates that an *average* decline in the transfer of technologies does not necessarily follow from an *average* increase in the speed of standardization. Moreover, Proposition 4 suggests two alternatives for explaining this fact: an increase in the rate of invention or an increase in the relative population size of the South.

### 3.3.5 Welfare: Comparison with the Case of Complete Contracts

At the end of section 3.2, I showed that in the one-good model, incomplete contracts were necessary for a product cycle to emerge. This remains valid in the general-equilibrium model. From the analysis of firm behavior, if contracts are complete, Northern and Southern assembly are equally profitable if and only if $\omega = 1$, from which it follows that the $A(\cdot)$ curve becomes flat at one. Solving for Southern labor demand under complete contracts, the analogous of the $B(\cdot)$ curve is given by:

$$\omega\left(t\right) = B_C(\bar{z}\left(t\right)) = \frac{1 - \alpha\xi_t\left(\bar{z}, 1\right)}{\alpha\xi_t\left(\bar{z}, 1\right)} \frac{L^S}{L^N}.$$

Relative to a world with incomplete contracts, the introduction of complete contracts shifts the $B(\bar{z})$ down and to the right (since $\alpha(1 - \phi) < 1$).

The general equilibrium under complete contracts is as follows. If $B_C(0) > 1$, then the relative wage in the North is necessarily greater than one and assembly is always done in the South. As before, production transfer occurs instantly and no product cycles emerge. If instead $B_C(0) < 1$, then the relative wage in the North will be equal to one in equi-

librium and the location of production of low-tech inputs will be indeterminate. Labor market clearing in the South will only require that the set of industries $\mathcal{Z} \subset [0,1]$ that choose to manufacture in the South satisfy $L^S \left(1 - \int_{z \in \mathcal{Z}} z f_{z,t}(z) dz\right) = L^N \int_{z \in \mathcal{Z}} z f_{z,t}(z) dz$. Clearly, in such case product cycles emerge with probability zero.

Consider next the welfare implications for each country of a shift from incomplete contracting to complete contracting.[21] It is useful to decompose the welfare change into three items: (a) terms of trade, (b) production efficiency, and (c) available products. The terms of trade effect relates to the fact that such a move towards complete contracts necessarily depresses the relative wage in the North. Intuitively, Northern wages include a rent on its comparative advantage stemming in part from incomplete contracting. In a world of complete contracts, this rent disappears and the relative wage in the North is lower. On account of this terms of trade effect, the South benefits from complete contracting, while the North is worse off.[22]

Second, the shift to complete contracts improves production efficiency on two accounts. On the one hand, it eliminates the direct distortions coming from incomplete contracting (e.g., inflation of mark-ups). On the other hand, there is also a production composition effect related to the fact that complete contracts lead to a shift of production towards the lower-wage country. On this account, both countries gain from a move to complete contracts.

The third effect relates to the effect on the endogenously determined measure of varieties in each industry. Because preferences of the representative consumer feature a love for variety, welfare is increasing in this measure . In a steady-state equilibrium with complete contracts, this measure would be given by $n^j = (1 - \alpha) \left(L^N + L^S/\omega_C\right)/Nf$, where $\omega_C$ is the equilibrium relative wage in such case. Comparing this expression with equation (3.14), it becomes clear that the overall effect of complete contracting on $n^j$ is only unambiguous for those industries with $z(j) < \bar{z}$. On the one hand, the fall in the

---

[21] To be more accurate, I will compare welfare in a steady-state equilibrium with incomplete contracts to that in a steady-state equilibrium with complete contracts.

[22] The terms of trade move in favor of the South for an additional reason which is related to the assumption that fixed costs are incurred in the North. In the model, incomplete contracts shift rents from the variable to the fixed cost sector, i.e., from the South to the North. A move to complete contracting eliminates this rent-shifting effect thereby reducing the relative wage in the North. Graphically, this corresponds to the $B(\cdot)$ curve shifting down and to the right.

relative wage ($\omega_C < \omega$) increases the measure of varieties because fixed costs in the North become relatively cheaper. On the other hand, the fraction of revenues that producers are able to capture as operating profits falls $(1 - \alpha < 1 - \alpha\,(\phi(1 - z + (1 - \phi)z))$, which translates into a lower $n^j$ for a given relative wage.

It follows that the first and second effects work in favor of the South, while the North also benefits from the second, but loses from the first. Although the third effect is in general ambiguous, it can be shown that it is always dominated by the production efficiency effect, in the sense that, on account of the second and third effects combined, both the South and the North are better off under complete contracts (see Appendix 3.6.3). From this result, it follows that:

**Proposition 5** *Relative to a world with incomplete contracting, a shift to complete contracts unambiguously increases welfare in the South, while having an ambiguous effect on welfare in the North.*

   **Proof.** *See Appendix 3.6.3.* ∎

If we interpret the shift towards better enforcement of international contracts as a tightening of intellectual property rights (IPRs) in the South, then Proposition 5 gives a result diametrically opposed to that in Helpman (1993).[23] He analyzed the welfare implications of a shift to better enforcement of IPRs in models with both exogenous and endogenous innovation and imitation. As he wrote in his concluding remarks:

> "Who benefits from tight intellectual property rights in less developed countries? My analysis suggests that if anyone benefits, it is not the South" (Helpman, 1993, p. 1274)

My analysis suggests instead that if anyone loses out, it is not the South. Our different conclusions stem from our different modelling of the vehicle of production transfer.[24]

---

[23] Admittedly, in the set-up described in section 3.2.1, an improvement in the contracting environment is more closely related to a tightening of trade-related intellectual property rights (TRIPS). Nevertheless, an alternative set-up in which contract incompleteness originated from a weak legal system in the South would yield similar results. In such a framework, the comparison between Proposition 5 and the results in Helpman (1993) would be more transparent.

[24] Lai (1998) makes a similar point in comparing the welfare implications of a tightening of IPRs in an endogenous-growth model with both foreign direct investment and imitation as vehicles of technology transfer.

When the South gains market share mostly by imitating, a tightening of IPRs will naturally hinder production transfer, reduce labor demand in the South, and make the South worse off. If production transfer is instead driven by voluntary decisions of Northern firms choosing to transact with offshore independent subcontractors or licensees, then a tightening of IPRs, to the extent that it alleviates contractual distortions in those international transactions, will increase relative labor demand in the South and improve welfare there.

## 3.4   Firm Boundaries and the Product Cycle

The purpose of this paper has been so far to develop a new theory of the product cycle based on the existence of incomplete contracts, and to analyze its macroeconomic implications. The recent literature on the theory of the firm has shown that incomplete contracts are also a useful vehicle to understand the determinants of ownership structure. I will build on this literature to draw firm boundaries in a simple extension of the model developed above. As a result, a much richer product cycle will emerge in which the transfer of production to the South may occur within firm boundaries or at arm's length.

The set-up of the model is the same as before with just one new feature: I allow the research center to vertically integrate the assembly plant and become a multinational firm. Following the property-rights approach of the theory of firm, vertical integration has the benefit of strengthening the ex-post bargaining power of the integrating party (the research center), but the cost of reducing the ex-post bargaining power of the integrated party (the assembly plant).

In particular, by integrating the production of the low-tech input, the manager of the assembly plant becomes an employee of the manager in the research center. This implies that if the assembly plant manager refuses to trade after the sunk costs have been incurred, the research center manager now has the option of firing him and seize the amount of $x_l$ produced. As in Grossman and Hart (1986), ownership is identified with the residual rights of control over certain assets. In this case, the low-tech input

plays the role of this asset.[25]

If there were no costs associated with firing the assembly plant manager, there would be no surplus to bargain over after production, and the assembly plant manager would ex-ante optimally set $x_l = 0$ (which of course would imply $y = 0$). In such case, integration would never be chosen. To make things more interesting, I assume that firing the assembly plant manager results in a negative productivity shock so that a fraction $1 - \delta$ of final-good production is lost. Under this assumption, the surplus of the relationship remains positive even under integration.[26] I take the fact that $\delta$ is strictly less than one as given, but this assumption could be rationalized in a richer framework.

For reasons analogous to those that make Assumption 1 necessary, I now assume that:

**Assumption 2:** $\delta \leq \left( \frac{3/4 - \phi}{1 - \phi} \right)^{1/\alpha}$.

The rest of this section is structured as follows. I will first revisit the static, partial-equilibrium model developed in section 3.2. Next, I will analyze the dynamics of the model and discuss the implications of vertical integration for this new view of the product cycle. Finally, I will solve for the general-equilibrium model with multinational firms.

### 3.4.1 Firm Behavior

In section 3.2.2, I computed ex-ante profits for the research center under two possible modes of organization: (A) assembly in the North by an independent manufacturing plant; and (B) assembly in the South by an independent manufacturing plant. The possibility of vertical integration introduces two additional options: assembly in the North by a vertically integrated manufacturing plant and assembly in the South by a vertically integrated manufacturing plant. Because contracts are assumed to be perfectly enforceable in transactions involving two firms located in the same country, it is straightforward to show that the first of these new options yields ex-ante profits identical to those in case (A). As is well known from the property-rights literature, in a world of complete

---

[25] See Antràs (2002a) for a related set-up.

[26] The fact that the fraction of final-good production lost is independent of $z$ greatly simplifies the analysis, but is not necesary for the qualitative results discussed below.

contracts, ownership structure is both indeterminate and irrelevant. In contrast, when Southern assembly is chosen, the assignment of residual rights is much more interesting.

## C. Assembly in the South by a Vertically-Integrated Manufacturing Plant

Consider next the problem faced by a research center and its integrated assembly plant in the South. If both managers decide to make good-quality intermediate inputs and they agree in the bargaining, the potential revenues from the sale of the final good are again $R = \lambda^{1-\alpha} \zeta_z^{\alpha} x_h^{\alpha(1-z)} x_l^{\alpha z}$. In contrast, if the parties fail to agree in the bargaining, the product-development manager will fire the assembly plant manager, who will be left with nothing. The research center will instead be able to sell an amount $\delta y(i)$ of output, which using equation (3.1) will translate into sale revenues of $\delta^{\alpha} R$. The quasi-rents of the relationship are therefore given by $(1 - \delta^{\alpha}) R$.

The contract incompleteness will give rise to renegotiation at $t_2$. In the bargaining, Generalized Nash bargaining leaves the research center with its default option plus a fraction $\phi$ of the quasi-rents. The manager of the integrated assembly plant receives the remaining fraction $1 - \phi$ of the quasi-rents. Since both $\phi$ and $\delta$ are assumed to be strictly less than one, ex-post revenues from producing a good-quality input are always strictly positive. Bad-quality inputs will therefore never be produced.

Rolling back to $t_1$, the research center will therefore set $x_h$ to maximize $\bar{\phi} R - w^N x_h$, where

$$\bar{\phi} = \delta^{\alpha} + \phi (1 - \delta^{\alpha}) > \phi.$$

On the other hand, the Southern assembly plant will simultaneously choose $x_l$ to maximize $(1 - \bar{\phi}) R - w^S x_l$. Combining the first-order conditions of these two programs yields the following optimal price for the final good:

$$p_M^S (z) = \frac{(w^N)^{1-z} (w^S)^z}{\alpha \bar{\phi}^{1-z} (1 - \bar{\phi})^z}.$$

Setting $T$ so as to make the integrated assembly plant break even leads to the following

expression for the research center's ex-ante profits:

$$\pi_M^S(z) = \left(1 - \alpha \left(\bar{\phi}(1-z) + \left(1 - \bar{\phi}\right) z\right)\right) \lambda \left(\frac{\left(w^N\right)^{1-z} \left(w^S\right)^z}{\alpha \bar{\phi}^{1-z} \left(1 - \bar{\phi}\right)^z}\right)^{-\alpha/(1-\alpha)}, \qquad (3.22)$$

which is identical to (3.6) with $\bar{\phi}$ replacing $\phi$, and where the subscript $M$ reflects the fact that the research center becomes a multinational firm under this arrangement.

**The Equilibrium Choice Revisited**

The product manager will now choose the manufacturing location *and* ownership structure that maximize profits for a given $z$. Consider first the choice between Northern assembly and Southern assembly by an independent firm. This was analyzed in section 3.2.3, where I showed that under Assumption 1 and Condition 1, there exists a unique $\bar{z}$ such that $\pi^N(z) > \pi^S(z)$ for $z < \bar{z}$, and $\pi^N(z) < \pi^S(z)$ for $z > \bar{z}$. Again, the cut-off level $\bar{z}$ is implicitly defined by $\bar{z} = A^{-1}(\omega)$, where $A(z)$ is given in equation (3.7).

Consider next the choice between Northern assembly and Southern assembly by an integrated firm. Comparing equations (3.4) and (3.22), it follows that $\pi^N(z) > \pi_M^S(z)$ if and only if

$$\omega < \bar{A}(z) \equiv \left(\frac{1 - \alpha}{1 - \alpha \left(\bar{\phi}(1-z) + \left(1 - \bar{\phi}\right) z\right)}\right)^{(1-\alpha)/\alpha z} \frac{\bar{\phi}^{-(1-z)/z}}{1 - \bar{\phi}}. \qquad (3.23)$$

Under Assumption 2, $\bar{\phi}$ is less or equal to $3/4$ and it can be shown that $\bar{A}'(z) < 0$ (see Appendix 3.6.1). Furthermore, when $z$ goes to zero, $\bar{A}(z)$ goes to $+\infty$, while when $z$ goes to 1, $\bar{A}(1) > 1$.[27] If the relative wage is high enough, namely, $\omega > \bar{A}(1)$, then there exists a unique cutoff $\bar{z}_{MN} = \bar{A}^{-1}(z) \in (0,1)$ such that $\pi^N(z) > \pi_M^S(z)$ for $z < \bar{z}_{MN}$, and $\pi^N(z) < \pi_M^S(z)$ for $z > \bar{z}_{MN}$.[28] The intuition for this result is essentially the same as the one discussed above. When the low-tech input is not very important in the production process, the cost-saving benefit of producing it in the South is outweighed

---

[27]This follows again from $(1 - \alpha x) x^{\alpha/(1-\alpha)}$ being an increasing function of $x$ for $\alpha \in (0,1)$ and $x \in (0,1)$.

[28]In contrast, if $\omega < \bar{A}(1)$, then $\pi^N(z) > \pi_M^S(z)$ for all $z$ and vertical integration is never chosen in equilibrium.

by the costs of incomplete contracts, which distort the marginal cost of production of both the hi-tech and the low-tech inputs. On the other hand, when very little product development is needed for production, then lower wages in the South are much more profit enhancing, and Southern assembly is more profitable. Crucial for this result is the fact that, following Grossman and Hart (1986), and contrary to the older transaction-cost literature, vertical integration does not eliminate the opportunistic behavior at the heart of the hold-up problem. Ownership, however, crucially affects the distribution of ex-post surplus through its effect on each party's outside option, and this explains that $\bar{z}$ and $\bar{z}_{MN}$ are not identical.

Finally, we need to compare profits under Southern assembly by an independent firm with those under Southern assembly by an integrated firm. Comparing equations (3.6) and (3.22), it follows that $\pi_M^S(z) > \pi^S(z)$ if and only if

$$1 < \Theta(z) \equiv \frac{1 - \alpha\left(\bar{\phi}(1-z) + \left(1-\bar{\phi}\right)z\right)}{1 - \alpha\left(\phi(1-z) + (1-\phi)z\right)} \left(\left(\frac{\bar{\phi}}{\phi}\right)^{1-z} \left(\frac{1-\bar{\phi}}{1-\phi}\right)^z\right)^{\frac{\alpha}{1-\alpha}} \qquad (3.24)$$

Appendix 3.6.4 demonstrates that $\Theta(z)$ is a decreasing function of $z$, and that both $\Theta(0) > 1$ and $\Theta(1) < 1$. This implies that there is unique cutoff $\bar{z}_{MS} = \Theta^{-1}(1)$ such that $\pi_M^S(z) > \pi^S(z)$ for $z < \bar{z}_{MS}$, and $\pi_M^S(z) < \pi^S(z)$ for $z > \bar{z}_{MS}$. The logic of this result also lies at the heart of Grossman and Hart's (1986) seminal contribution. Ex-ante efficiency dictates that residual rights should be controlled by the party undertaking a relatively more important investment. If production of the final good requires mostly product development (i.e., $z$ is low), then the investment made by the assembly plant manager will be relatively small, and thus it will be optimal to assign the residual rights of control to the research center. Conversely, when the low-tech input is important in production, the research center will optimally choose to tilt the bargaining power in favor of the assembly plant by giving away these same residual rights.

Figure 3-4 illustrates this point by depicting the amounts of inputs produced under each organizational mode, as well as those prevailing under complete contracts. Incomplete contracts leads to underproduction of both $x_h$ and $x_l$. The curves $M^*$ and $R^*$ represent the reaction functions $x_h^*(x_l)$ and $x_l^*(x_h)$ under complete contracts, with the

Figure 3-4: Underproduction and Ownership Structure
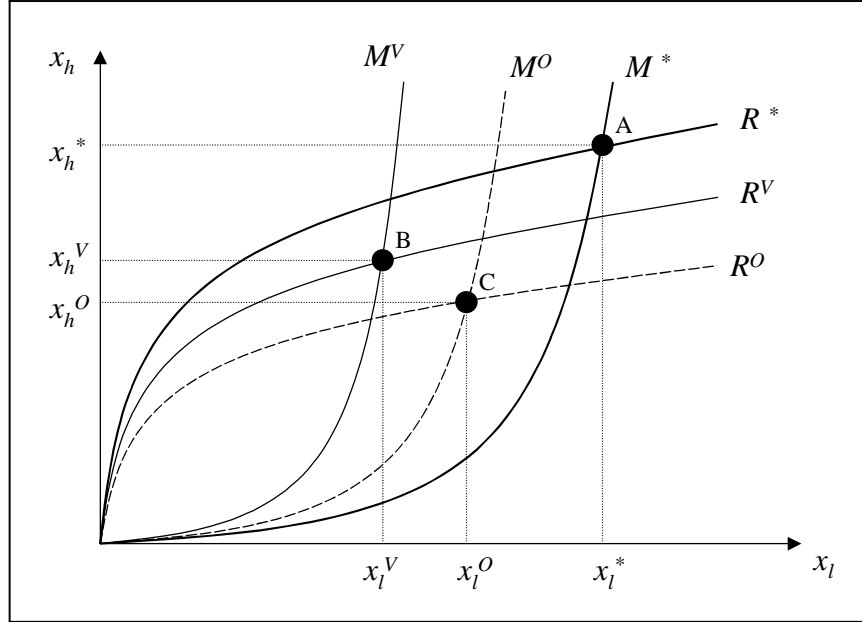


corresponding equilibrium in point A. Similarly, B and C depict the incomplete-contract equilibria corresponding to vertical integration and arm's length transacting, respectively. The crucial point to notice from Figure 3-4 is that the underproduction in $x_l$ is relatively higher under integration that under outsourcing. The more important is the low-tech input in production, the more value-reducing will the underinvestment in $x_l$ be. It thus follows that profits under integration relative to those under arm's length transacting will tend to be lower, the more important is the low-tech input in production (i.e., the higher $z$).

So far I have only discussed the existence of the three cut-off levels: $\bar{z}$, $\bar{z}_{MN}$, and $\bar{z}_{MS}$. Notice, however, that $\Theta(z) = \left(A(z)/\bar{A}(z)\right)^{\alpha z/(1-z)}$. This can be shown to imply that $\bar{z}$, $\bar{z}_{MN}$, and $\bar{z}_{MS}$ must satisfy one of these three: (i) $\bar{z}_{MS} = \bar{z} = \bar{z}_{MN}$, (ii) $\bar{z}_{MS} < \bar{z} < \bar{z}_{MN}$, or (iii) $\bar{z}_{MN} < \bar{z} < \bar{z}_{MS}$.[29] Figure 3-5 is instructive in understanding this result. The figure depicts the curves $A(z)$ and $\bar{A}(z)$, the latter being the thicker curve. Under Assumptions 1 and 2, both $A(z)$ and $\bar{A}(z)$ are decreasing in $z$. That these curves intersect just once is

---

[29]To see this, notice for instance that $\bar{z}_{MS} < \bar{z}$ if and only if both $A(\bar{z}_{MS}) > \omega$ and $\Theta(\bar{z}) < 1$. But the latter can only be true if $A(\bar{z})/\bar{A}(\bar{z}) = \omega/\bar{A}(\bar{z}) < 1$, which implies $\bar{z} < \bar{z}_{MN}$.

ensured by the fact that $A(z') = \bar{A}(z')$ is only consistent with $\Theta(z') = 1$, which necessarily implies $z' = \bar{z}_{MS}$. Furthermore, it is easy to show that $A(z) > \bar{A}(z)$ for low enough $z$, while $A(1) < \bar{A}(1)$.[30] For any relative wage $\omega$, it is clear that either $\bar{z}_{MS} < \bar{z} < \bar{z}_{MN}$ (top panel) or $\bar{z}_{MN} < \bar{z} < \bar{z}_{MS}$ (bottom panel). The case $\bar{z}_{MS} = \bar{z} = \bar{z}_{MN}$ will occur with probability zero and will be ignored hereafter.

As indicated in both panels in Figure 3-5, for a low enough value for $z$, the benefits from Southern assembly are too low relative to the distortions from incomplete contracting, and $x_l$ is produced in the North. Furthermore, for a sufficiently high value of $z$, a profit-maximizing research center will decide to outsource the manufacturing input to an independent assembly plant in the South. Whether, for intermediate values of $z$, the research center becomes a multinational firm or not depends on parameter values. If $\bar{z}_{MS} < \bar{z} < \bar{z}_{MN}$, then there exists no $z \in [0,1]$ for which $\pi^S_M > \max\{\pi^N, \pi^S\}$, and multinational firms do not arise in equilibrium. Conversely, if $\bar{z}_{MN} < \bar{z} < \bar{z}_{MS}$, multinational firms can arise provided that $z \in [\bar{z}_{MN}, \bar{z}_{MS}]$. To summarize the results of this section,

**Lemma 6** *If $\bar{z}_{MS} < \min\{\bar{z}, \bar{z}_{MN}\}$, the research center chooses to produce the low-tech input in the North for $z < \bar{z}$, and in the South by an* unaffiliated *party for $z > \bar{z}$. On the other hand, if $\bar{z}_{MS} > \min\{\bar{z}, \bar{z}_{MN}\}$, the low-tech input is produced in the North for $z < \bar{z}_{MN}$, in the South by an* affiliated *party if $\bar{z}_{MN} < z < \bar{z}_{MS}$, and in the South by an* unaffiliated *party if $\bar{z} > \bar{z}_{MS}$.*
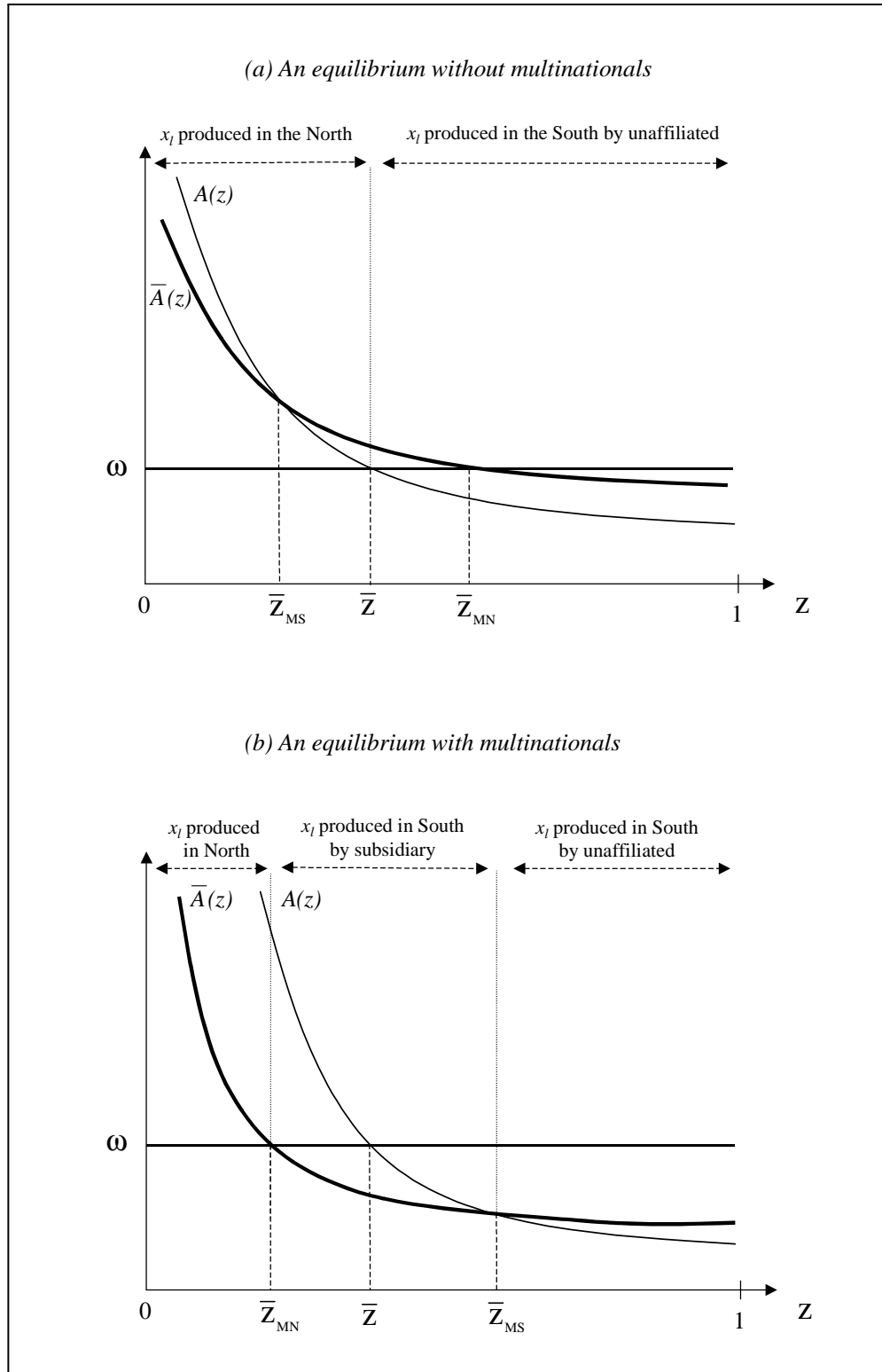
**Proof.** *See Appendix 3.6.4.* ∎

An equilibrium with multinational firms is more likely the higher is $\bar{z}_{MS}$ relative to the other two thresholds $\bar{z}$ and $\bar{z}_{MN}$. In section 3.2.3, I showed that $\bar{z}$ is a decreasing function of the relative wage $\omega$. By way of implicit differentiation, and making use of the fact that $\bar{A}'(z) < 0$ whenever $\bar{\phi} \leq 3/4$, it follows that $\bar{z}_{MN}$ is also decreasing in $\omega$. The choice between an independent and an integrated Southern supplier, as captured by the threshold $\bar{z}_{MS}$ is instead unaffected by the relative wage in the North.[31] From

---

[30]Notice the similarities between the curves $A(z)$ and $\bar{A}(z)$, and the comparative static exercise of increasing $\phi$ at the end of section 3.2.3.

[31]This follows directly from the assumption of Cobb-Douglas technology and isolates the partial-

# Figure 3-5: Partial Equilibrium and Firm Boundaries

### (a) An equilibrium without multinationals

$x_l$ produced in the North    $x_l$ produced in the South by unaffiliated

$A(z)$

$\overline{A}(z)$

$\omega$

$0$    $\overline{Z}_{MS}$    $\overline{Z}$    $\overline{Z}_{MN}$    $1$    $Z$

### (b) An equilibrium with multinationals

$x_l$ produced in North    $x_l$ produced in South by subsidiary    $x_l$ produced in South by unaffiliated

$\overline{A}(z)$    $A(z)$

$\omega$

$0$    $\overline{Z}_{MN}$    $\overline{Z}$    $\overline{Z}_{MS}$    $1$    $Z$

111

this it follows that the measure of hi-tech intensities for which multinationals exist, i.e., $\min\{\bar{z}_{MS} - \bar{z}_{MN}, 0\}$, is non-decreasing in the relative wage $\omega$ in the North.

## 3.4.2 Dynamics: The Product Cycle

Consider a dynamic set-up analogous to that in section 3.2.4. As before, assume that the output elasticity of the low-tech input evolves according to $z(t) = h(t)$, with $h'(t) > 0$, $h(0) = 0$, and $\lim_{t\to\infty} h(t) = 1$. The parameters $\phi$, $\alpha$, $\omega$, and $\delta$ are instead assumed to be time-invariant, implying that $\bar{z}$, $\bar{z}_{MN}$, and $\bar{z}_{MS}$ are also constant through time. The following is a straightforward corollary of Lemma 6:

**Proposition 6** *The model displays a product cycle. If $\bar{z}_{MS} < \min\{\bar{z}, \bar{z}_{MN}\}$, the product cycle is as described in Proposition 3. If instead $\bar{z}_{MS} > \min\{\bar{z}, \bar{z}_{MN}\}$, the following product cycle emerges. When the industry is young, i.e. $t < h^{-1}(\bar{z}_{MN})$, the low-tech input is produced in the North. For an intermediate maturity of the industry, $h^{-1}(\bar{z}_{MN}) < t < h^{-1}(\bar{z}_{MS})$, production of the low-tech input is shifted to the South but is undertaken within firm boundaries. When the good is highly standardized, i.e. $t > h^{-1}(\bar{z}_{MS})$, production is shifted to an unaffiliated party in the South.*

This is an important result of the paper. In words, it states that if the threshold product-development intensity at which manufacturing is shifted to the South is low enough, then production transfer will occur at arm's length and multinationals will not emerge in equilibrium. Conversely, if this threshold product-development intensity is high enough, manufacturing will be shifted to the South within the boundaries of the Northern firm by establishing a wholly-owned foreign affiliate. In such case, arm's length assembly in the South will only be observed at a later stage in the life cycle of the good.

As argued in the introduction, Proposition 6 is consistent with the evolution of the Korean electronics industry from the early 1960s to the late 1980s. The industry took off with the establishment of a few foreign affiliates of U.S. and Japanese-based corporations.

---

equilibrium decision to integrate or outsource from any potential general-equilibrium feedbacks. This implied block-recursiveness is a useful property for solving the model sequentially, but the main results should be robust to more general specifications for technology.

These foreign affiliates were manufacturing goods that were relatively new for Korean standards and their exports constituted a very large percentage of Korean exports in electronics at that time. Then in the 1970s and 1980s, when the industry had matured, Northern corporations changed their strategies and started resorting much more to OEM and similar licensing arrangement. Within the industry, the evidence also seems to suggest that the probability of internalization of the technology transfer was higher for relatively new goods.

At a more micro level, several cross-sectional implications of the model are consistent with the findings of the empirical literature on the product cycle. Mansfield and Romeo (1980) analyzed 65 technology transfers by 31 U.S.-based firms in a variety of industries. They found that, on average, U.S.-based firms tended to transfer technologies internally to their subsidiaries within 6 years of their introduction in the United States. The average lag for technologies that were transferred through licensing or through a joint venture was instead 13 years.[32] Similarly, after surveying R&D executives of 30 U.S.–based multinational firms, Mansfield, Romeo, and Wagner (1979) concluded that for young technologies (less than 5 years old), internal technology transfer tended to be preferred to licensing, whereas for more mature technologies (between 5 and 10 years), licensing became a much more attractive choice.

In much more detailed studies, Davidson and McFetridge (1984, 1985) looked at 1,376 internal and arm's-length transactions involving high-technology products carried out by 32 US.-based multinational enterprises between 1945 and 1975. Their logit estimates indicated that the probability of internalization was higher: (i) the newer and more radical was a technology; (ii) the fewer was the number of previous transfers of the same technology; and (iii) the larger was the fraction of the transferor's resources devoted to scientific R&D.

Considering again the specific functional form $z(t) = 1 - e^{-\frac{t}{\theta}}$, it is straightforward to check that another cross-sectional implication of the model is that an increase in the rate of standardization $1/\theta$ anticipates the shift of production towards the South but (weakly) reduces the length of time that the good is produced by a wholly-owned affiliate in the

---

[32] Their sample consisted of 39 internal transfer and 26 arm's length transfers.

South. Using a sample of 350 US firms, Wilson (1977) indeed concluded that licensing was more attractive the less complex was the good involved, with his measure of complexity being positively correlated with the amount of R&D undertaken for its production. In their study of the transfer of 35 Swedish innovations, Kogut and Zander (1993) similarly found that the probability of internalization was lower the more codifiable and teachable and the less complex was the technology.

As mentioned at the end section 3.4.1, another implication of the model is that an increase in $\omega$ not only makes the shift to Southern production earlier (as in section 3.2.4), but also increases the probability that this transfer will occur within firm boundaries. This is also consistent with the cross-sectional results of at least two contributions to the empirical literature on the product cycle. In their sample of 1,376 transfers, Davidson and McFetridge (1985) found that controlling for several characteristics of the recipient country, a higher GNP per capita of the recipient country (in the model a lower $\omega$) was associated with a lower probability of internalization. Using aggregate industry data from the U.S. Department of Commerce, Contractor (1984) found similar results.

### 3.4.3   General Equilibrium

In this section, I will adopt again the general-equilibrium framework in section 3.3 and solve for the equilibrium relative wage in the North in the model with multinational firms. By doing so, we can study the effect of some additional factors on the choice between foreign direct investment and arm's length arrangements, as well as the welfare implications of the emergence of multinational firms.

Consider first the industry equilibrium. As before, the unit elasticity of substitution between varieties in different industries allows an independent analysis of each industry. Furthermore, since all research centers share the same technology and contracting environment, they all behave in an identical manner. If $\bar{z}_{MS} < \min\{\bar{z}(t), \bar{z}_{MN}(t)\}$, there will be no active multinationals and the equilibrium number of varieties in any industry $j$ is again given by equation (3.14).[33] If instead $\bar{z}_{MS} > \min\{\bar{z}(t), \bar{z}_{MN}(t)\}$, then it can

---

[33]Notice that if $\phi$, $\alpha$, and $\delta$ are time-invariant, so will $\bar{z}_{MS}$. The other two thresholds will only be constant through time in a steady-state equilibrium with a constant relative wage.

be shown that $n_j(t)$ is given by:

$$n_j(t) = \begin{cases} \frac{(1-\alpha)E(t)}{N(t)w^N(t)f} & \text{if } z_j(t-t_{0j}) < \bar{z}_{MN}(t) \\[3mm] \frac{\left(1-\alpha\left(\bar{\phi}(1-z_j(t-t_{0j}))+\left(1-\bar{\phi}\right)z_j(t-t_{0j})\right)\right)E(t)}{N(t)w^N(t)f} & \text{if } \bar{z}_{MN}(t) < z_j(t-t_{0j}) < \bar{z}_{MS} \\[3mm] \frac{(1-\alpha(\phi(1-z_j(t-t_{0j}))+(1-\phi)z_j(t-t_{0j})))E(t)}{N(t)w^N(t)f} & \text{if } z_j(t-t_{0j}) > \bar{z}_{MS} \end{cases}$$
(3.25)

In order to solve for the equilibrium relative wage, we will focus again on labor-market clearing in the South. It is useful to define $\widetilde{z}(t) = \min\{\bar{z}_{MN}(t), \bar{z}(t)\}$, which according to Lemma 6 constitutes the threshold $z$ above which production is shifted to the South regardless of ownership structure. Notice that this is the only variable apart from $\omega(t)$ that we need to pin down in the general equilibrium. This is because $\bar{z}_{MS}$ does not depend on $\omega(t)$ and because the threshold $\bar{z}_{MN}(t)$ is only relevant when $\bar{z}_{MN}(t) < \bar{z}(t)$ (in which case $\bar{z}(t)$ is irrelevant). Computing labor demand by the Southern assembly plant yields a labor market clearing condition analogous to equation (3.16) above:

$$\int_{\widetilde{z}(t)}^{1} \alpha(1-\phi_z)zE(t)f_{z,t}(z)dz = w^S(t)L^S,$$
(3.26)

where $\phi_z = \bar{\phi}$ if $\bar{z}_{MN}(t) < z < \bar{z}_{MS}(t)$, and $\phi_z = \phi$ otherwise.

To save on notation, let the distribution function $F_{z,t}(z)$ converges to a time-invariant distribution, so that we can focus on the steady-state general equilibrium and safely drop time subscripts. Notice then that multinationals are active only if $\widetilde{z} \leq \bar{z}_{MS}$. Plugging equation (3.15) into (3.26), and defining again $\xi(a,b) \equiv \int_a^b zf_z(z)dz$, we find the following equilibrium condition relating $\omega$ and $\widetilde{z}$:

$$\omega = B_M(\widetilde{z}) \equiv \begin{cases} \frac{1-\alpha\left(1-\bar{\phi}\right)\xi(\widetilde{z},\bar{z}_{MS})-\alpha(1-\phi)\xi(\bar{z}_{MS},1)}{\alpha\left(1-\bar{\phi}\right)\xi(\widetilde{z},\bar{z}_{MS})+\alpha(1-\phi)\xi(\bar{z}_{MS},1)} \frac{L^S}{L^N} & \text{if } \widetilde{z} < \bar{z}_{MS} \\[3mm] \frac{1-\alpha(1-\phi)\xi(\widetilde{z},1)}{\alpha(1-\phi)\xi(\widetilde{z},1)} \frac{L^S}{L^N} & \text{if } \widetilde{z} > \bar{z}_{MS} \end{cases}.$$
(3.27)

Notice that if $\widetilde{z} > \bar{z}_{MS}$ the equilibrium is one without multinationals and therefore the equilibrium condition collapses back to the one in the previous model, i.e. $B_M(\widetilde{z}) = B(\widetilde{z})$.

115

If instead $\widetilde{z} < \bar{z}_{MS}$, multinationals will arise in equilibrium. Furthermore, from the standard logic in Grossman and Hart (1986), an integrated assembly plant manager will tend to underinvest relative to a non-integrated one. Hence, for a given $z$, Southern labor demand is relatively lower for vertically-integrated manufacturing plants. It follows that for $\widetilde{z} < \bar{z}_{MS}$, $B_M(\widetilde{z}) > B(\widetilde{z})$. This is depicted in Figure 3-6, where for comparison the curve $B(\widetilde{z})$ is plotted as a dotted curve. Note also that $B_M(\widetilde{z})$ is continuous and increasing function of $\widetilde{z}$. Furthermore it satisfies $B_M(0) > B(0) > 0$ and $\lim_{\widetilde{z} \to 1} B_M(\widetilde{z}) = +\infty$.

The other equilibrium condition that pins down $\widetilde{z}$ and $\omega$ comes from firm behavior. In particular, since $\bar{z}$ is implicitly defined by $\omega = A(\bar{z})$, and $\bar{z}_{MN}$ is implicitly defined by $\omega = \bar{A}(\bar{z}_{MN})$, it follows that $\widetilde{z} = \min\{\bar{z}_{MN}, \bar{z}\}$ is implicitly defined by:

$$
\omega = A_M(\widetilde{z}) \equiv
\begin{cases}
\left(\dfrac{1-\alpha}{1-\alpha\left(\bar{\phi}(1-\widetilde{z})+\left(1-\bar{\phi}\right)\widetilde{z}\right)}\right)^{(1-\alpha)/\alpha\widetilde{z}} \dfrac{\bar{\phi}^{-(1-\widetilde{z})/\widetilde{z}}}{1-\bar{\phi}} & \text{if } \widetilde{z} < \bar{z}_{MS} \\[4mm]
\left(\dfrac{1-\alpha}{1-\alpha(\phi(1-\widetilde{z})+(1-\phi)\widetilde{z})}\right)^{(1-\alpha)/\alpha\widetilde{z}} \dfrac{\phi^{-(1-\widetilde{z})/\widetilde{z}}}{1-\phi} & \text{if } \widetilde{z} > \bar{z}_{MS}
\end{cases}
\tag{3.28}
$$

Again, if $\widetilde{z} > \bar{z}_{MS}$, multinationals are not active in equilibrium and $A_M(\widetilde{z}) = A(\widetilde{z})$. On the other hand, if $\widetilde{z} < \bar{z}_{MS}$, then it must be the case that $\Theta(\widetilde{z}) > 1$, which of course implies $A_M(\widetilde{z}) = \bar{A}(\widetilde{z}) < A(\widetilde{z})$. Research centers choose to vertically integrate the Southern manufacturing plant only when, by doing so, they manage to alleviate the distortions stemming from incomplete contracting. It should therefore not be surprising that to match the profitability of Northern assembly, an integrated manufacturing plant requires a lower relative wage than an independent one, This is again depicted in Figure 3-6, where the curve $A(\widetilde{z})$ is the dotted one. Notice that $A_M(\widetilde{z})$ is continuous and, under Assumptions 1 and 2, decreasing in $\widetilde{z}$. Furthermore, $\lim_{\widetilde{z} \to 0} A_M(\widetilde{z}) = +\infty$ and $A_M(1) > 1$.

As illustrated in Figure 3-6, there exists a unique general-equilibrium $(\widetilde{z}, \omega)$ pair. Depending on parameter values, the equilibrium is one with no multinationals firms (top panel) or one with multinational firms (bottom panel). In the first case, the equilibrium is identical to that in the world with only arm's length transacting. As before, in the steady state, the general-equilibrium values of $\widetilde{z}$ and $\omega$ are time-invariant, and all industries will necessarily follow product cycles, with varieties in those industries first being

Figure 3-6: General Equilibrium and Firm Boundaries



*(a) An equilibrium without multinationals*

*(b) An equilibrium with multinationals*
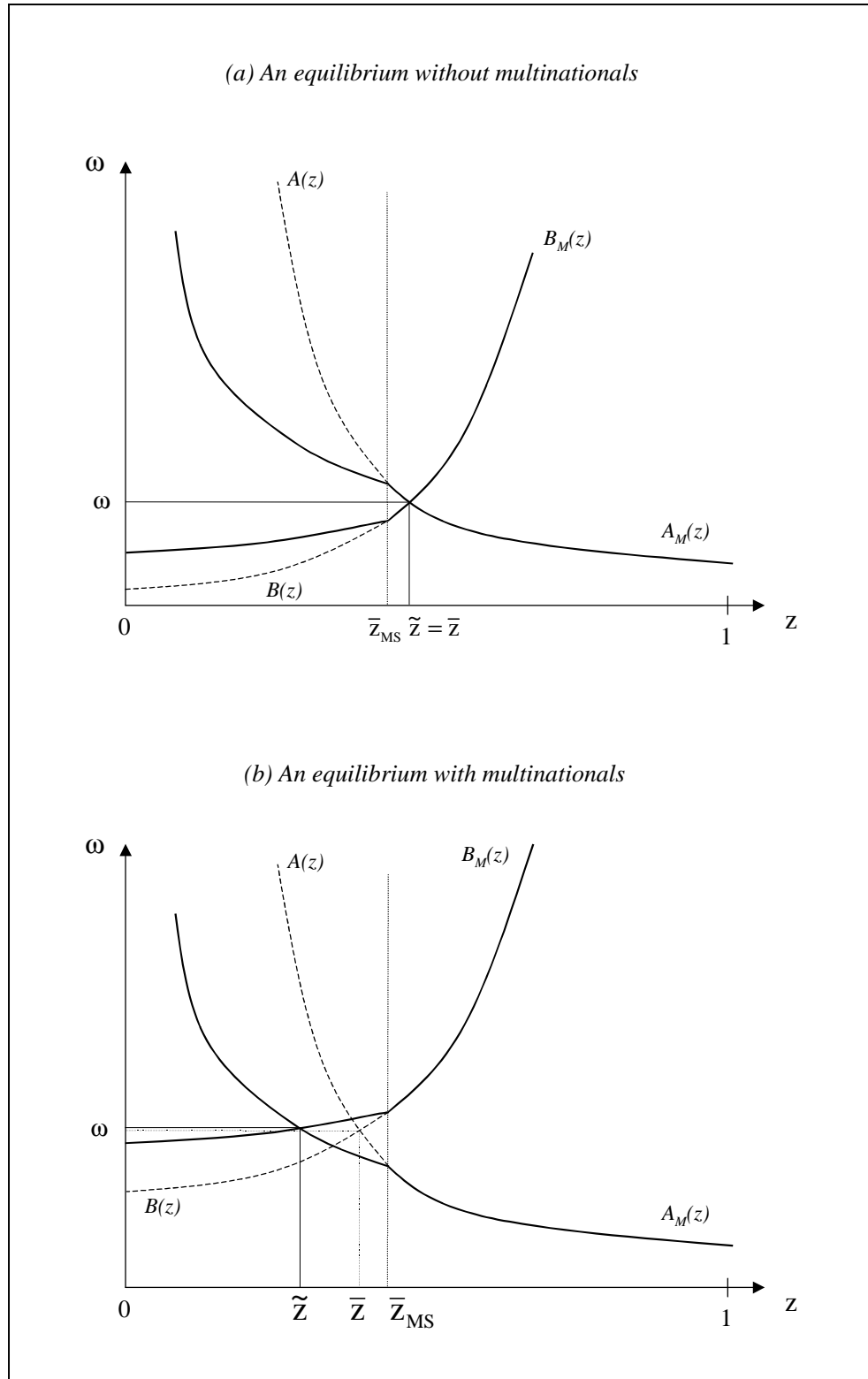
manufactured in the North and then later in the South. On top of this endogenous product cycles, the model may feature endogenous organizational cycles, with technologies first shifted to the South within firm boundaries and only later licensed to independent firms in the South.

The following interesting result follows from direct inspection of Figure 3-6:

**Proposition 7** *Relative to a world with only arm's length transacting, allowing for intrafirm technology transfer by multinational firms weakly accelerates the transfer of production to the South (lowers $\widetilde{z}$), while having an ambiguous effect on the relative wage $\omega$.*

The intuition for this result is the following. The introduction of multinational firms in a world with only arm's length transacting (weakly) alleviates contract incompleteness. This is because the research center is now given the possibility to have the low-tech input produced in the South, while facing less of a hold-up problem in their production of the hi-tech input. When this option is chosen, the threshold $z$ above which $x_l$ is produced in the South necessarily falls ($\widetilde{z}$ falls). Alternatively, the relative Northern wage above which $x_l$ is produced in the South necessarily falls (remember that this is the reason why $\bar{A}(\widetilde{z}) < A(\widetilde{z})$ to the left of $\bar{z}_{MS}$). On top of this partial-equilibrium effect, there is a general-equilibrium effect. As discussed above, integrated manufacturing plants demand relatively less Southern labor. To clear the Southern labor market requires either an increase in the measure of $z$'s for which production is done in the South (again a fall in $\widetilde{z}$) or a fall in the relative wage in the South (a higher $\omega$). In sum, both the partial and general-equilibrium effects work to reduce $\widetilde{z}$, while they have opposite effects on the equilibrium relative wage.

The result in Proposition 7 fits well Moran's (2001) recent study of the effects of domestic-content, joint-venture, and technology-sharing mandates on production transfer to developing countries. Plants in host countries that impose such restrictions, he writes, "*utilize older technology, and suffer lags in the introduction of newer processes and products in comparison to wholly owned subsidiaries without such requirements*" (p. 32). He also describes an interesting case study. In 1998, Eastman Kodak agreed to set up joint ventures with three designated Chinese partners. These joint ventures specialized

in producing conventional films under the Kodak name. When the Chinese government allowed Kodak to establish a parallel wholly owned plant, Kodak shifted to this affiliate the manufacturing of the latest digitalized film and camera products (Moran, 2001, p. 36).

### 3.4.4   Comparative Statics

Let us now return to the particular case analyzed before, so that $\xi_t(a,b)$ converges to

$$\xi(a,b) \equiv \int_a^b z f_{z,t}(z)dz = \int_a^b \frac{zg\theta_\mu}{\left(1 + g\theta_\mu \ln\left(\frac{1}{1-z}\right)\right)^2 (1-z)} dz$$

As before, an increase in $g$, $\theta_\mu$ or $L^S/L^N$ shifts the $B_M(\widetilde{z})$ curve up and to the left (see Appendix 3.6.2). This tends to increase the relative wage in the North and reduce $\widetilde{z}$, so that production is shifted to the South earlier. Hence, introducing multinational firms does not undermine the validity of the comparative static results results in Proposition 4.

More interestingly, an increase in the rate of invention, a slowdown in the rate of standardization, and an increase in the relative size of the South all tend to (weakly) increase the measure of product-development intensities for which multinational firms exist:

**Proposition 8** *The measure of product-development intensities for which multinationals exist, i.e.,* $\min\{\bar{z}_{MS} - \bar{z}_{MN}, 0\}$*, is non-decreasing in* $g$*,* $\theta_\mu$*, and* $L^S/L^N$*.*

Intuitively, by creating an excess supply of Northern labor, an increase in $g$, $\theta_\mu$, and $L^S/L^N$, lead to an increase in the relative wage and an increase in the threshold product-development intensity $(1 - \widetilde{z})$ below which manufacturing is transferred to the South. From the standard Grossman-Hart logic, it then becomes more likely that the Northern research center will decide to keep this transfer within firm boundaries.

Several authors have identified a recent surge in foreign direct investment flows to less developed countries (e.g., Feenstra, 1999). According to Proposition 8, this fact can be explained by the same forces that would have led to a fall in the interval of time

between the introduction of a new product in the North and its first production in a less developed country (see the discussion following Proposition 4).

### 3.4.5  Welfare

According to Proposition 7, a shift towards an equilibrium in which foreign direct investment flows are positive accelerates the transfer of production to the South, while having an ambiguous effect on the relative wages. I now analyze the welfare implications of such a shift. To do so, it is useful to decompose again the change in welfare into the following three components: (a) terms of trade, (b) production efficiency, and (c) available products.

The fact that the introduction of multinational firms has an ambiguous effect on relative wages implies an equally ambiguous welfare change on account of both the terms-of-trade and the available-products components. In contrast, it is easy to show that the production efficiency effect works to increase welfare in both countries. On the one hand, multinationals will only be active to the extent that they alleviate the distortions coming from incomplete contracting. On the other hand, when active, multinationals lead to a shift of production to the lower-wage country.

In general, the net effect on each country's welfare is ambiguous. It can be shown, however, that the available-products effect is again dominated by the production efficiency component. It thus follows that:

**Proposition 9** *Provided that its effect on relative wages is small enough, allowing for intrafirm production transfer by multinational firms is welfare improving for both countries.*

If the terms-of-trade effect is negligible, then the introduction of multinational firms has a net positive effect on each country's welfare. This result provides some support for the view pushed by Moran (2001), that the domestic-content, joint-venture, and technology-sharing requirements that certain host countries impose on foreign firms have a negative net contribution to welfare.

## 3.5    Conclusions

This paper presented a dynamic, general-equilibrium model featuring both endogenous product cycles and endogenous organizational cycles. It has been argued that the same forces that make firms choose to manufacture their new goods in high-wage countries can explain why, when they decide to transfer production to low-wage countries, they might choose to do so inside their firm boundaries. The model delivered a few macroeconomic implications that complement the work of Krugman (1979). For instance, I showed that increased technological change in rich countries will widen the world distribution of income only to the extent that technological change takes the form of an addition of new goods to the economy. When technological change takes the form of a continuous standardization of products, then an acceleration of technological change leads instead to a narrowing of the world distribution of income. Furthermore, it was shown that an improvement in the contractual environment in international transactions would necessarily benefit low-wage countries, with the net welfare effect being in generally ambiguous for high-wage countries.

In contrast to previous general-equilibrium theories of the multinational firm, firm boundaries were not drawn appealing to technological considerations, such as economies of scale or transport costs.[34] As in Antràs (2002a), I instead set forth a purely organizational, property-rights model of the multinational firm.[35] Multinational firms emerged in equilibrium whenever transaction-cost minimization dictated that certain goods would be transacted more efficiently within firm boundaries than at arm's length. Relative to a world with only arm's length transacting, I showed that multinational firms, by alleviating contract incompleteness, anticipated the transfer of production to low-wage countries and, under certain conditions, increased welfare in both rich and poor countries.

---

[34]This previous literature builds on the seminal work of Helpman (1984) and Markusen (1984). An exception is the work of Ethier (1986), who also dealt with the crucial issue of internalization, although in a very different framework. For an extensive review of previous theories of the multinational firm see Caves (1996).

[35]This paper is related to an emerging literature on general-equilibrium models of ownership structure (c.f., McLaren, 2000, Grossman and Helpman, 2002a, Antràs, 2002a). In Antràs (2002a), I unveiled two systematic patterns in the volume of intrafirm trade, which I then rationalized in a theoretical framework that combined a Grossman-Hart-Moore view of the firm with a Helpman-Krugman view of international trade.

The simple model developed here has proven to be a useful lens through which to interpret several findings in the international business literature. Nevertheless, much remains to be done. For instance, the present framework has abstracted from at least one important channel of production transfer, namely, imitation. Future efforts should also be directed at incorporating elements of alternative theories of the firm to the study of international patterns of specialization. Holmström and Milgrom (1994) have emphasized that, in many situations, issues related to job design and the cost of measuring performance are more relevant when choosing between inside or outside procurement. It would be interesting to investigate the implications of such a view of the firm for the way in which multinational firms organize their operations worldwide.

## 3.6 Appendix

### 3.6.1 Proof of Lemma 5

I will first show that under Assumption 1, $A'(z) < 0$ for all $z \in [0, 1]$.

**Lemma 7** *If $\phi \leq 3/4$, then $A'(z) < 0$ for all $z \in [0, 1]$.*

**Proof.** *Straightforward differentiation yields $A'(z) < 0$ if and only if*

$$r(z, \phi, \alpha) = \ln \left( \frac{1 - \alpha \left( \phi \left( 1 - z \right) + \left( 1 - \phi \right) z \right)}{1 - \alpha} \phi^{\alpha/(1-\alpha)} \right) - \frac{\left( 2\phi - 1 \right) \alpha z}{\left( 1 - \alpha \left( \phi \left( 1 - z \right) + \left( 1 - \phi \right) z \right) \right)} < 0$$

*It can be shown that both $\partial r(\cdot)/\partial z \geq 0$ for all $z \in [0, 1]$ (with strict inequality for $z > 0$, and $\partial r(\cdot)/\partial \phi \geq 0$ all $\phi \in (0, 1)$. We hence need only show that $r(1, 3/4, \alpha) < 0$. But this is true because $\partial r(1, 3/4, \alpha)/\partial \alpha < 0$ for $\alpha \in (0, 1)$ and $r(1, 3/4, 0) = 0$. QED.*[36] ∎

Notice next that $\omega < \lim_{z \to 0} A(z) = \infty$, where in computing the limit I make use of the fact that $(1 - \alpha x) x^{\alpha/(1-\alpha)}$ is increasing in $x$ for $\alpha \in (0, 1)$ and $x \in (0, 1)$. Finally, by Condition 1, there exists a $z_c \in [0, 1]$ such that $A(z_c) < \omega$. Since $A(z)$ is a continuous function of $z$ for $z \in [0, 1]$, and by Lemma 7 $A'(z) < 0$, it follows that there is a unique $\bar{z}$ such that $\omega = A(\bar{z})$. Furthermore, for all $z < \bar{z}$, $A(z) > \omega$, and for all $z > \bar{z}$, $A(z) < \omega$. QED.

### 3.6.2 Algebra of the Particular Case

I will first show that the distribution of $z$ converges to a c.d.f. characterized by equation (3.20). Note that from equation (3.19), $t_{0j} = t + \theta_j \ln \left( 1 - z_j(t, t_{0j}) \right)$. For a given $\theta_j$, the fraction of industries with $z_j(t - t_{0j}) < z$ converges to

$$F_z \left( z | \theta_j \right) = 1 - F_{t_0} \left( t + \theta_j \ln \left( 1 - z \right), t \right) = 1 - \left( 1 - z \right)^{g\theta_j},$$

---

[36] *To see that an upper bound on $\phi$ is required, notice that at $\phi = 1$, $r(z, 1, \alpha) = \ln \left( \frac{1-\alpha(1-z)}{1-\alpha} \right) - \frac{\alpha z}{1-\alpha(1-z)}$, which is positive for high enough $z$.*

123

where I have used the fact that the distribution of birth dates converges to $F_{t_0}(t_0, t) = e^{-g(t-t_0)}$, as well as the fact that $t_{0j}$ and $\theta_j$ are independent. If $F_\theta(\theta_j) = 1 - e^{-\theta_j/\theta_\mu}$, the limiting unconditional distribution is then given by

$$F_z(z) = 1 - \int_0^\infty (1-z)^{g\theta_j} \cdot f_\theta(\theta_j) \cdot d\theta_j = \frac{g\theta_\mu \left(\ln \left(\frac{1}{1-z}\right)\right)}{1 + g\theta_\mu \left(\ln \left(\frac{1}{1-z}\right)\right)},$$

as claimed in equation (3.20) above.

Letting $f_z(z)$ be the corresponding probability density function, it follows that

$$\xi(a, b) \equiv \int_a^b z f(z) dz = \int_a^b \frac{z g\theta_\mu}{\left(1 + g\theta_\mu \ln \left(\frac{1}{1-z}\right)\right)^2 (1 - z)} dz$$

Integrating by parts yields:

$$\xi(a, b) = \frac{a}{1 + g\theta_\mu \ln \left(\frac{1}{1-a}\right)} - \frac{b}{1 + g\theta_\mu \ln \left(\frac{1}{1-b}\right)} + \int_a^b \frac{1}{1 + g\theta_\mu \ln \left(\frac{1}{1-z}\right)} dz \qquad (3.29)$$

To proof the claim in section 3.3.4 that an increase in $g$ or $\theta_\mu$ shifts the $B(\cdot)$ curve up and to the left, we need only show that $\partial \xi(\bar{z}, 1)/\partial g < 0$ and $\partial \xi(\bar{z}, 1)/\partial \theta_\mu < 0$. But this is follows from straightforward differentiation of $\xi(\bar{z}, 1)$ in (3.29). Similarly, to proof the claim in section 3.4.3 that an increase in $g$ or $\theta_\mu$ shifts the $B_M(\cdot)$ curve up and to the left, notice that $\xi(\tilde{z}, \bar{z}_{MS}) + \xi(\bar{z}_{MS}, 1) = \xi(\tilde{z}, 1)$, and therefore $(1 - \bar{\phi}) \xi(\tilde{z}, \bar{z}_{MS}) + (1 - \phi) \xi(\bar{z}_{MS}, 1) = (1 - \bar{\phi}) \xi(\tilde{z}, 1) + (\bar{\phi} - \phi) \xi(\bar{z}_{MS}, 1)$. The result then follows again from $\partial \xi(z, 1)/\partial \theta_\mu < 0$ and $\partial \xi(z, 1)/\partial g < 0$ for all $z$.

### 3.6.3   Proof of Proposition 5

The results in Proposition 5 follow from comparing instantaneous welfare in a steady-state equilibrium with incomplete contracts to that in a steady-state equilibrium with complete contracts. Let us first derive a general expression for instantaneous welfare. Plugging equation (3.11) into (3.10), imposing the equality of income and spending, and

rearranging, welfare in country $c = \{N, S\}$ can be expressed as:

$$u^c = \frac{w^c L^c}{N} \int_0^N \left( \int_0^{n_j} p_j(i)^{-\alpha/(1-\alpha)} di \right)^{(1-\alpha)/\alpha} dj \qquad (3.30)$$

In an equilibrium with incomplete contracts, the North manufactures all varieties in industries with $z_j < \bar{z}$. Substituting the relevant values of $p_j(i)$ and $n_j(i)$ and using the definition of $A(z)$ in equation (3.7), welfare in the North can be expressed as

$$u_{IC}^N = \alpha L^N \left( \frac{(1-\alpha)\left(L^N + L^S/\omega\right)}{Nf} \right)^{(1-\alpha)/\alpha} \left( F(\bar{z}) + \int_{\bar{z}}^1 \left( \frac{\omega}{A(z)} \right)^z f_z(z) dz \right), \qquad (3.31)$$

while that in the South is simply $u_{IC}^S = u_{IC}^N L^S / \left( \omega L^N \right)$.

Under complete contracts and wage equalization (i.e., $B_C(0) < 1$), welfare in country $c$ reduces to

$$u_{CC}^c = \alpha L^c \left( \frac{(1-\alpha)\left(L^N + L^S\right)}{Nf} \right)^{(1-\alpha)/\alpha}. \qquad (3.32)$$

Comparing equations (3.31) and (3.32) and given that $\omega > A(z)$ for all $z > \bar{z}$, it is easy to see that the North may indeed be better off under incomplete contracts. The result is ambiguous because of the the second term in equation (3.31). In contrast, the South necessarily benefits from complete contracts because under incomplete contracts both $\omega > 1$ and $A(z) > 1$ for all $z \in [0,1]$.

Finally consider the case of complete contracts and $B_C(0) > 1$. Denote the relative wage in this equilibrium by $\omega_C > 1$. In this case, welfare is given by:

$$u_{CC'}^N = \alpha L^N \left( \frac{(1-\alpha)\left(L^N + L^S/\omega_C\right)}{Nf} \right)^{(1-\alpha)/\alpha} \int_0^1 \omega_C^z f_z(z) dz, \qquad (3.33)$$

while that in the South is simply $u_{CC'}^S = u_{CC'}^N L^S / \left( \omega_C L^N \right)$. The South is again necessarily better off under complete contracts because $\omega > \omega_C$ and $A(z) > 1$ for all $z \in [0,1]$. The effect on Northern welfare is again ambiguous.

### 3.6.4  Proof of Lemma 6

The only thing that was not proved in the main text is that $\Theta(z)$ is decreasing in $z$, and that both $\Theta(0) > 1$ and $\Theta(1) < 1$. Using $\bar{\phi} = \delta^\alpha + \phi(1 - \delta^\alpha)$, we can rewrite $\Theta(z)$ in equation (3.24) as

$$\Theta(z) = \left(1 + \frac{\alpha(1 - \phi)\,\delta^\alpha(2z - 1)}{1 - \alpha\,(\phi(1 - z) + (1 - \phi)\,z)}\right)\left(1 + \frac{\delta^\alpha}{\phi(1 - \delta^\alpha)}\right)^{\frac{\alpha(1 - z)}{1 - \alpha}}(1 - \delta^\alpha)^{\frac{\alpha}{1 - \alpha}}$$

From simple differentiation, it follows that $\partial\Theta(\cdot)/\partial z < 0$ if and only if

$$\Omega(z)\ln\left(1 + \frac{\delta^\alpha}{\phi(1 - \delta^\alpha)}\right) > (2 - \alpha)(1 - \alpha)(1 - \phi)\,\delta^\alpha$$

where $\Omega(z) = (1 - \alpha\,(\phi(1 - z) + (1 - \phi)\,z))\left(1 - \alpha\left(\bar{\phi}(1 - z) + (1 - \bar{\phi})\,z\right)\right)$ and remember that $\bar{\phi} = \delta^\alpha + \phi(1 - \delta^\alpha)$. Now notice that if $\bar{\phi} > \phi \geq 1/2$ then $\Omega'(z) > 0 \ \forall z \in [0, 1]$, and if $\phi < \bar{\phi} \leq 1/2$, then $\Omega'(z) < 0 \ \forall z \in [0, 1]$. Furthermore, if $\bar{\phi} > 1/2 > \phi$, then $\Omega''(z) < 0 \ \forall z \in [0, 1]$. It thus follows that $\Omega(z) \geq \min\{\Omega(0), \Omega(1)\}$. Without loss of generality, assume that $\Omega(0) = (1 - \alpha\phi)(1 - \alpha\,(\phi + (1 - \phi)\delta^\alpha) < \Omega(1)$ (the case $\Omega(1) < \Omega(0)$ is entirely symmetric). We need to show that $\vartheta(\delta) > 0$ for all $\delta \in (0, 1)$ where

$$\vartheta(\delta) = \ln\left(1 + \frac{\delta^\alpha}{\phi(1 - \delta^\alpha)}\right) - \frac{(2 - \alpha)(1 - \alpha)(1 - \phi)\,\delta^\alpha}{(1 - \alpha\phi)(1 - \alpha\,(\phi + (1 - \phi)\delta^\alpha)}$$

From simple differentiation of this expression, it follows that $\vartheta'(\delta) > 0$ if and only if $(1 - \alpha\rho)^2 - (2 - \alpha)(1 - \alpha)(1 - \rho)\,\rho > 0$ for some $\rho \in (0, 1)$. But it is simple to check that this is in fact true all $\alpha, \rho \in (0, 1)$, and therefore $\vartheta(\delta) > \vartheta(0) = 0$.

Finally, notice that $\Theta(0) = \frac{1 - \alpha\bar{\phi}}{1 - \alpha\phi}\left(\frac{\bar{\phi}}{\phi}\right)^{\frac{\alpha}{1 - \alpha}} > 1$ and $\Theta(1) = \frac{1 - \alpha(1 - \bar{\phi})}{1 - \alpha(1 - \phi)}\left(\frac{1 - \bar{\phi}}{1 - \phi}\right)^{\frac{\alpha}{1 - \alpha}} < 1$, where the inequalities follow from $\bar{\phi} > \phi$ and the fact that $(1 - \alpha x)x^{\frac{\alpha}{1 - \alpha}}$ is an increasing function of $x$ for $\alpha \in (0, 1)$ and $x \in (0, 1)$. QED.

# Chapter 4

# Global Sourcing[1]

## 4.1   Introduction

A firm that chooses to keep the production of an intermediate input within its boundaries can produce it at home or in a foreign country. When it keeps it at home, it engages in standard vertical integration. And when it makes it abroad, it engages in foreign direct investment (FDI) and intra-firm trade. Alternatively, a firm may choose to outsource an input in the home country or in a foreign country. When it buys the input at home, it engages in domestic outsourcing. And when it buys it abroad, it engages in arm's-length trade. Intel Corporation provides an example of the FDI strategy; it assembles most of its microchips in wholly-owned subsidiaries in China, Costa Rica, Malaysia, and the Philippines. On the other hand, Nike provides an example of the arm's-length import strategy; it subcontracts most of its manufacturing to independent producers in Thailand, Indonesia, Cambodia, and Vietnam.

Growth of international specialization has been a dominant feature of the international economy. Amongst the many examples that illustrate this trend, two are particularly telling. Citing Tempest (1996), Feenstra (1998) illustrates Mattel's global sourcing strategy in the production of its star product, the Barbie doll. "Of the $2 export value for the dolls when they leave Hong Kong for the United States," he writes, "about 35 cents

---

[1]**This chapter is joint work with Professor Helpman from Harvard University.**

covers Chinese labor, 65 cents covers the cost of materials," — which are imported from Taiwan, Japan, and the United States — "and the remainder covers transportation and overheads, including profits earned in Hong Kong" (pp.35-36). The World Trade Organization provides another example in its 1998 annual report. In the production of an "American" car, 30 percent of the car's value originates in Korea, 17.5 percent in Japan, 7.5 percent in Germany, 4 percent in Taiwan and Singapore, 2.5 percent in the U.K., and 1.5 percent in Ireland and Barbados. That is, "...only 37 percent of the production value... is generated in the United States" (p.36).

Importantly, the increasing international disintegration of production is large enough to be noticed in aggregate statistics. Feenstra and Hanson (1996) use U.S. input–output tables to infer U.S. imports of intermediate inputs. They find that the share of imported intermediates increased from 5.3% of total U.S. intermediate purchases in 1972 to 11.6% in 1990. Campa and Goldberg (1997) find similar evidence for Canada and the U.K. (but not for Japan). Hummels, Ishii and Yi (2001), who use a narrower concept of international specialization, i.e., the fraction of imported inputs embodied in the production of goods destined for export, find that in 9 OECD countries and 4 emerging market economies this fraction increased — by almost 30% on average — between 1970 and 1990 (again, not in Japan).[2]

But how important is intra-firm relative to arm's-length trade in intermediate inputs? A firm-level data analysis is needed to answer this question, and no such analysis is available at this point in time. And despite the fact that the business press has stressed the spectacular growth of foreign outsourcing, Hanson, Mataloni and Slaughter (2002) document an equally impressive growth of trade within multinational firms. Nevertheless, the fact that imports from foreign affiliates of U.S.-based firms has fallen from 23.9% of total U.S. imports in 1977 to 16.3% in 1999, suggests that foreign outsourcing might have outpaced foreign intra-firm sourcing by U.S. firms.[3]

Other studies have documented a rise in the prevalence of *domestic* outsourcing by

---

[2]In related work, Yeats (2001) constructs a direct measure of trade in components, taking advantage of a recent revision of the Standard International Trade Classification (SITC) system. His data for machinery and transport equipment suggest that, between 1978 and 1995, international trade in components has grown at a faster rate than international trade in final-stage products.

[3]

U.S. firms. *The Economist* (1991), Bamford (1994) and Abraham and Taylor (1996), all report rising subcontracting in particular industries or activities. A systematic analysis of this trend is not available. Nevertheless, Figure 4-1 provides indirect evidence of a decline in vertical integration. The average number of four-digit SIC segments in which a U.S. publicly-traded manufacturing company operates, declined from 2.72 in 1979 to 1.81 in 1997.[4] The figure suggests that U.S. manufacturing firms have become increasingly specialized, which indicates a trend towards less vertical integration.

To address issues that arise from the choice of outsourcing versus integration and home versus foreign production, we need a theoretical framework in which companies make endogenous organizational choices. We propose such a framework in this paper by integrating two recent strands of the literature.

Melitz (2002) and Helpman, Melitz and Yeaple (2003) have studied the effects of within sectoral heterogeneity on the decisions of firms to serve foreign markets. By allowing productivity to differ across firms, they show that low-productivity firms serve the domestic market but not foreign markets, while high-productivity firms also serve

Table 4.1. Foreign insourcing

| Year | Imports from foreign affiliates by U.S. parents | Total U.S. imports | Share of foreign insourcing in U.S. imports |
|------|------|------|------|
| 1977 | 36,266 | 151,534 | 23.9 |
| 1982 | 39,288 | 243,952 | 16.1 |
| 1989 | 74,738 | 473,211 | 15.8 |
| 1994 | 114,881 | 663,256 | 17.3 |
| 1999 | 166,990 | 1,024,618 | 16.3 |

Sources: BEA Direct Investment data set and U.S. Census

Table 4.1 reports data for the five years in which the BEA conducted comprehensive surveys on the universe of U.S. firms engaging in foreign direct investment. As is evident from the table, the large drop in the share of insourcing occurred sometime between the late 1970s and the early 1980s, and it remained relatively constant during the last 20 years. This share is only a rough measure of the relative importance of foreign insourcing, however, because both the numerator and the denominator include trade in finished goods, and the denominator may also incorporate insourcing by U.S. affiliates of foreign-based multinationals.

[4]The data for this figure are taken from Fan and Lang (2000), who constructed it from the Compustat data set. One might worry that the trend in Figure 4-1 is driven by a composition effect, i.e., by a relative increase in the number of firms in relatively specialized manufacturing sectors. To examine this possibility, we regressed the four-digit SIC segments per firm on a time trend and firm fixed effects. The coefficient on the time trend was negative, with a T-statistic of $-66.93$. We interpret this to imply that firms specialized more over time.

Figure 4-1: The average number of four-digit SIC segments in which firms operate.
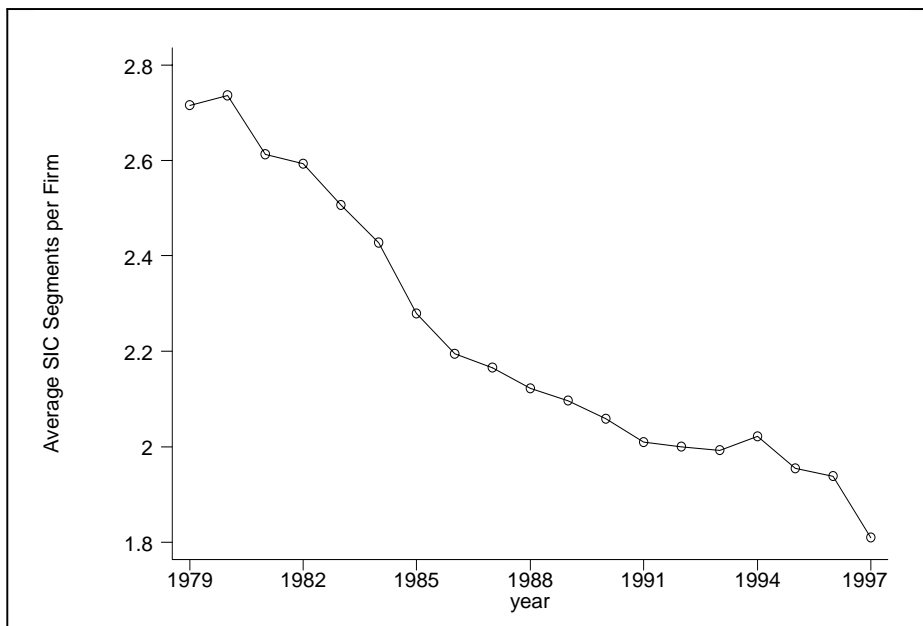


foreign markets. Allowing for horizontal foreign direct investment, Helpman, Melitz and Yeaple also showed that, amongst the firms that serve foreign markets, the more productive ones engage in foreign direct investment while the less productive firms export. Importantly, affiliate sales relative to exports are larger in sectors with more productivity dispersion. Their approach emphasized variations across firms within industries, without addressing the organizational choices of firms that need to acquire intermediate inputs.

Grossman and Helpman (2002a) addressed the choice between outsourcing and integration in a one-input general equilibrium framework, assuming that all firms of a given type are equally productive. Their firms face the friction of incomplete contracts in arm's-length relationships, which they weigh against the less-efficient production of inputs in integrated companies. As a result, some sectors have only vertically integrated firms while others have only disintegrated firms. Grossman and Helpman identify sectoral characteristics that lead to one or the other equilibrium structure. This approach has been extended by Antràs (2002a) to a trading environment, by introducing two important new features. First, the friction of incomplete contracts also exists within integrated firms, and — as in Grossman and Hart (1986) — integration provides well defined prop-

erty rights. However, these property rights may or may not give integration an advantage over outsourcing. Second, there are two inputs, one controlled by the final-good producer, the other by another supplier, inside or outside the firm. The relative intensity of these inputs turns out to be an important determinant of the choice between integration and outsourcing.

By embodying this structure in a Helpman and Krugman (1985) style two-sector general equilibrium model of trading countries, Antràs shows that the sector that is relatively intensive in the input controlled by the final-good producers integrates, while the sector that is relatively intensive in the other input outsources. As a result, in the former sector there is intra-firm trade in inputs, while in the latter sector there is arm's-length trade.

Building on this literature, we develop a theoretical model that combines the within-sectoral heterogeneity of Melitz (2002) with the structure of firms in Antràs (2002a). The final-good producer controls the supply of headquarter services while an operator of the production facility of intermediate inputs controls their production. This allows us to explore the joint variations within and across sectors, i.e., productivity within sectors and technological and organizational features across sectors, in shaping organizational forms, trade and foreign direct investment. In particular, we show that in a world of two countries, North and South, in which final-good producers are based in the North, final-good producers who operate in the same sector but differ by productivity sort into integrated companies that produce inputs in the North (do not engage in foreign trade), integrated companies that produce inputs in the South (engage in FDI and intra-firm trade), disintegrated companies that outsource in the North (do not engage in foreign trade), and disintegrated companies that outsource in the South (import inputs at arm's length).

We show that in low-headquarter-intensive sectors firms do not integrate; low-productivity firms outsource in the North while high-productivity firms outsource in the South. In sectors with high headquarter intensity all four organizational forms emerge in equilibrium. Importantly, high-productivity firms import inputs while low-productivity firms acquire them in the North. However, amongst the firms that acquire inputs in the same country, the low-productivity firms outsource while the high-productivity firms integrate. This

implies that the least-productive firms outsource in the North while the most productive firms integrate in the South via foreign direct investment.

We use the model to study the relative prevalence of different organizational forms. We show how prevalence depends on the wage gap between the North and the South, the trading costs of intermediate inputs, the degree of productivity dispersion within a sector, the distribution of bargaining power, the size of the ownership advantage (which may be different in the two countries), and the intensity of headquarter services. Our model predicts that relatively more final-good producers rely on imported intermediates in sectors with higher productivity dispersion or lower headquarter intensity. And in sectors with integration and outsourcing, which are the sectors with high headquarter intensity, industries with higher productivity dispersion have relatively more final-good producers who integrate. This is true for a comparison of integration versus outsourcing in each of the countries. As a result, such sectors have more intra-firm trade relative to arm's-length trade. These results illustrate the types of issues that can be addressed with our model.

Our model is developed in the next section. In section 3 we characterize an industry's equilibrium. Then, in section 4, we describe the equilibrium sorting of firms into different organizational forms, and we study in section 5 the prevalence of each mode of organization. This is also the section that examines the effects of variations within and across sectors on the relative prevalence of organizational forms. Section 6 offers a short summary with concluding comments.

## 4.2   The Model

Consider a world with two countries, the North and the South, and a unique factor of production, labor. The world is populated by a unit measure of consumers with identical preferences represented by:

$$U = x_0 + \frac{1}{\mu} \sum_{j=1}^{J} X_j^{\mu}, \ \ 0 < \mu < 1,$$

where $x_0$ is consumption of a homogeneous good, $X_j$ is an index of aggregate consumption in sector $j$, and $\mu$ is a parameter. Aggregate consumption in sector $j$ is a CES function

$$X_j = \left[ \int x_j(i)^\alpha di \right]^{1/\alpha}, \quad 0 < \alpha < 1,$$

of the consumption of different varieties $x_j(i)$, where the range of $i$ will be endogenously determined. The elasticity of substitution between any two varieties in a given sector is $1/(1-\alpha)$. We assume that $\alpha > \mu$, so that varieties within a sector are more substitutable for each other than they are substitutable for $x_0$ or for varieties from a different sector. This leads to the inverse demand function for each variety $i$ in sector $j$:

$$p_j(i) = X_j^{\mu-\alpha} x_j(i)^{\alpha-1}. \tag{4.1}$$

In every country the differentiated product sectors are assumed to be small relative to the size of the local labor market. As a result, producers in these sectors face a perfectly elastic supply of labor.[5] We denote by $w^N$ the wage rate in the North and by $w^S$ the wage rate in the South. These wage rates are fixed and $w^N > w^S$.[6]

The demand parameters $\mu$ and $\alpha$ are the same in every industry, which helps to focus attention on cross-sectoral differences in technology and cross-country differences in organizational costs. Our aim is to explore how differences in technology interact with organizational choices in shaping industrial structure, trade flows and FDI.

Only the North has the know-how to produce final-good varieties in the non-homogeneous sectors. To start producing a variety in sector $j$ a firm needs to bear a fixed cost of entry consisting of $f_E$ units of Northern labor. Upon paying this fixed cost, the unique producer of variety $i$ in sector $j$ draws a productivity level $\theta$ from a known distribution $G(\theta)$.[7] After observing this productivity level, the final-good producer decides whether

---

[5]A simple way to ensure this property is to assume that there is a continuum of sectors $j$ rather than a finite number.

[6]The assumption of fixed wage rates and a higher wage rate in the North can be justified in general equilibrium by assuming that $w_\ell$ is the productivity of labor in producing $x_0$ in country $\ell$, $\ell = N, S$, and that labor supply is large enough in every country so that both countries produce $x_0$.

[7]We can accommodate sectoral differences in the fixed cost $f_E$ and the distribution $G(\theta)$, but chose them to be identical for simplicity.

to exit the market or start producing; in the latter case an additional fixed cost of organizing production needs to be incurred. As discussed below, this additional fixed cost is a function of the structure of ownership and the location of production.

Production of any final-good variety requires a combination of two variety-specific and freely tradable intermediate inputs, $h_j(i)$ and $m_j(i)$, which we associate with headquarter services and manufactured components, respectively. Output of every variety is a sector-specific Cobb-Douglas function of the inputs,

$$x_j(i) = \theta \left( \frac{h_j(i)}{\eta_j} \right)^{\eta_j} \left( \frac{m_j(i)}{1 - \eta_j} \right)^{1-\eta_j}, \quad 0 < \eta_j < 1. \tag{4.2}$$

Notice that, up to the productivity parameter $\theta$, the technology is identical for all varieties in a given sector, but sectors differ in the relative intensity of headquarter services, as represented by $\eta_j$. The larger is $\eta_j$ the more intensive is the sector in headquarter services.

The unit cost function for producing intermediate inputs is identical in all sectors but varies by country. Production of one unit of headquarter services $h_j(i)$ in the North requires one unit of Northern labor, while the South is much less efficient at producing headquarter services. We assume that the productivity advantage of the North is so large that headquarter services are only produced in the North. On the other hand, production of one unit of $m_j(i)$ requires one unit of labor in the North and in the South.

There are two types of producers: final-good producers and operators of manufacturing plants for components. Only final-good producers have the know-how to produce headquarter services. On the other hand, every final-good producer needs to contract with a manufacturing-plant operator for the provision of components. We allow international fragmentation of the production process, so that the final-good producer can choose to transact with a manufacturing plant in the North or in the South.

It follows from our assumptions that final-good producers locate in the North. Upon paying the fixed cost of entry $w^N f_E$ and observing the productivity level $\theta$, the unique final-good producer of variety $i$ in sector $j$ decides whether to match with an operator of a manufacturing plant in the North or with one in the South. Simultaneously, the final-good producer chooses whether to vertically-integrate the manufacturing plant or engage instead in an arm's-length transaction.

We specify a very simple matching technology. After paying the fixed cost of search in a given market, a final-good producer finds a match with probability one. We assume that final-good producers in the North need to incur a higher fixed cost to search in the unfamiliar South than in the familiar North. We also assume that the *status quo* is for the firms to remain non integrated. In addition to the search costs, a final good producer incurs management and negotiation costs that depend on the organizational form. All these costs, the sum of which we term *fixed organizational costs*, are in terms of Northern labor. We denote them by $w^N f_k^\ell$, where $k$ is an index of the ownership structure and $\ell$ is an index of the country in which the manufacturing of components takes place.

The ownership structure takes one of two forms: vertical integration $V$ or outsourcing $O$. The location of the manufacturing of components is in one of two sites: in the North $N$ or in the South $S$. Therefore $k \in \{V, O\}$ and $\ell \in \{N, S\}$. An *organizational form* consists of an ownership structure and a location for the production of components. Because the status quo is for the final-good producer to be non-integrated, we assume that the fixed organizational costs are higher for a vertically integrated firm, no matter in which country it owns the manufacturing plant for components. Namely, $f_V^\ell > f_O^\ell$ for $\ell = N, S$.[8] Note that when a final-good producer owns a manufacturing plant of components in the North this represents a standard situation of vertical integration. On the other hand, when a final-good producer owns the manufacturing plant of components in the South, this represents vertical foreign direct investment (FDI).

We finally assume that the fixed organizational costs are higher in the South, because the fixed costs of search, monitoring, and communication are higher in the foreign country. Combined with the assumption that integration entails higher costs of organization than outsourcing, the fixed organizational costs are ranked as follows:

$$f_V^S > f_O^S > f_V^N > f_O^N. \tag{4.3}$$

The location of the manufacturing of components and the mode of ownership are

---

[8]One can imagine situations in which this may not be the case. Nevertheless, we believe that this assumption is appropriate in many instances, and we therefore maintain it in the main analysis. It is not difficult, however, to see how various results change when outsourcing requires higher fixed costs. We shall point out how some of the results differ in this case.

chosen ex-ante by the final-good producer to maximize the joint value of the relationship, as measured by the sum of the operating profits of the final-good producer and the manufacturing plant net of all fixed costs of production. This can be justified by assuming that the final-good producer sets a fee for participation in the relationship that has to be paid by the operator of the manufacturing plant. This fee can be positive or negative, i.e., the operator can make a payment to the final good producer or vice versa. The purpose of the fee is to secure the participation of the operator in the relationship at minimum cost to the final-good producer. When the supply of operators of manufacturing plants is infinitely elastic, the operator's profits from the relationship net of the participation fee are equal in equilibrium to his outside option.[9] For simplicity, we set the operators' outside option equal to zero. It is, however, easy to extend the analysis to cases in which these outside options are positive and different in the North and in the South.

The setting is one of incomplete contracts. Final-good producers and manufacturing-plant operators cannot sign ex-ante enforceable contracts specifying the purchase of specialized intermediate inputs for a certain price. In addition, the parties cannot write enforceable contracts contingent on the amount of labor hired or on the volume of sales revenues obtained when the final good is sold. One can use arguments of the type developed by Hart and Moore (1999) and Segal (1999) to justify this specification. Namely, that the parties cannot commit not to renegotiate an initial contract and that the precise nature of the required input is revealed only ex-post, and it is not verifiable by a third party. To simplify the analysis, we just impose these constraints on the contracting environment.

Because no enforceable contract can be signed ex-ante, final-good producers and manufacturing-plant operators bargain over the surplus from the relationship after the inputs have been produced. We model this ex-post bargaining as a Generalized Nash Bargaining game in which the final-good producer obtains a fraction $\beta \in (0, 1)$ of the ex-post gains from the relationship.[10]

Following the property-rights approach to the theory of the firm, we assume that ex-post bargaining takes place both under outsourcing and under integration. The dis-

---

[9]See Antràs (2002a,b) for more details.

[10]This specification is similar to Grossman and Helpman (2002a) and Antràs (2002a,b).

tribution of surplus is sensitive, however, to the mode of organization. More specifically, the outside option for a final-good producer is assumed to be different when it owns the manufacturing plant than when it does not. In the latter case, a failure to reach an agreement on the distribution of the surplus leaves both parties with no income, because the inputs are tailored specifically to the other party in the transaction. However, by vertically integrating the production of components, the final-good producer is effectively buying the right to fire the operator of the manufacturing plant and seize the inputs $m_j(i)$. If there were no costs associated with firing the operator of the manufacturing plant, the final-good producer would always have an incentive to seize the inputs $m_j(i)$ ex-post, and the manufacturing-plant operator would have an incentive to choose $m_j(i) = 0$ ex-ante (which of course would imply $x_j(i) = 0$). In this case integration would never be chosen. We therefore assume that firing the manufacturing-plant operator results in a loss of a fraction $1 - \delta^\ell$ of final-good production.[11] We also assume that $\delta^N > \delta^S$. This captures the notion that a contractual breach is likely to be more costly for the final-good producer when the manufacturing plant is located in the South.

## 4.3  Equilibrium

Consider the payoffs in the bargaining game for a pair of firms producing in sector $j$. Since from now on we discuss a particular sector, we drop for simplicity the index $j$ from all the variables. If the parties agree in the bargaining, the potential revenue from the sale of the final goods is $R(i) = p(i)x(i)$, which, using (4.1) and (4.2), can be written as

$$R(i) = X^{\mu-\alpha}\theta^\alpha \left(\frac{h(i)}{\eta}\right)^{\alpha\eta} \left(\frac{m(i)}{1-\eta}\right)^{\alpha(1-\eta)}. \tag{4.4}$$

If they fail to agree, however, the outside option for the manufacturing-plant operator is always 0 while that for the final-good producer varies with the ownership structure and the location of components manufacturing.

---

[11]The fact that the fraction of final-good production lost is independent of $\eta_j$ greatly simplifies the analysis, but it is not necessary for the qualitative results discussed below.

When the final-good producer outsources components, its outside option is also 0 regardless of the location of the manufacturing plant. In this event the final-good producer gets $\beta R(i)$ while the manufacturing-plant operator gets $(1 - \beta) R(i)$.

The final-good producer has more leverage under vertical integration. When the parties fail to reach an agreement, the final-good producer can sell an amount $\delta^\ell x(i)$ of output when its manufacturing plant is in country $\ell$, which yields the revenue $(\delta^\ell)^\alpha R(i)$. The ex-post gains from trade are in this case $\left[1 - (\delta^\ell)^\alpha\right] R(i)$. In the bargaining, the final-good producer receives its default option plus a fraction $\beta$ of the quasi-rents, i.e. $(\delta^\ell)^\alpha R(i) + \beta \left[1 - (\delta^\ell)^\alpha\right] R(i)$, while the operator of the manufacturing plant obtains $(1 - \beta) \left[1 - (\delta^\ell)^\alpha\right] R(i)$.

Notice that the payoffs in the bargaining game are proportional to the revenue. Denoting by $\beta_k^\ell R(i)$ the payoff of the final-good producer under ownership structure $k$ and the location of components manufacturing in country $\ell$, the assumption $\delta^N > \delta^S$ implies that

$$\beta_V^N = \left(\delta^N\right)^\alpha + \beta \left[1 - \left(\delta^N\right)^\alpha\right] > \beta_V^S = \left(\delta^S\right)^\alpha + \beta \left[1 - \left(\delta^S\right)^\alpha\right] > \beta_O^N = \beta_O^S = \beta. \quad (4.5)$$

That is, final-good producers are able to appropriate higher fractions of revenue under integration than under outsourcing, with this fraction being higher when integration takes place in the North.

Since the delivery of the inputs $h(i)$ and $m(i)$ is not contractible ex-ante, the parties choose their quantities noncooperatively; every supplier maximizes its own payoff. In particular, the final-good producer provides an amount of headquarter services that maximizes $\beta_k^\ell R(i) - w^N h(i)$ while the manufacturing-plant operator provides an amount of components that maximizes $\left(1 - \beta_k^\ell\right) R(i) - w^\ell m(i)$. Combining the first-order conditions of these two programs, using (4.4), the total value of the relationship, as measured by operating profits, can be expressed as:

$$\pi_k^\ell (\theta, X, \eta) = X^{(\mu-\alpha)/(1-\alpha)} \theta^{\alpha/(1-\alpha)} \psi_k^\ell (\eta) - w^N f_k^\ell \quad (4.6)$$

where

$$\psi_k^\ell(\eta) = \frac{1 - \alpha\left[\beta_k^\ell \eta + \left(1 - \beta_k^\ell\right)(1 - \eta)\right]}{\left[\frac{1}{\alpha}\left(\frac{w^N}{\beta_k^\ell}\right)^\eta \left(\frac{w^\ell}{1-\beta_k^\ell}\right)^{1-\eta}\right]^{\alpha/(1-\alpha)}}. \tag{4.7}$$

Note that among the arguments of the profit function $\pi_k^\ell(\theta, X, \eta)$, the first one is firm-specific while the others are industry-specific. Moreover, while $\eta$ is a parameter, the consumption index $X$ is endogenous to the industry but exogenous to the producer of a specific variety of the final good.

Upon observing its productivity level $\theta$, a final-good producer either pays the fixed organizational cost $w^N f_k^\ell$ and chooses the ownership structure and the location of manufacturing that maximizes (4.6), or exits the industry and forfeits the fixed cost of entry $w^N f_E$. It is clear from (4.6) that the latter occurs whenever $\theta$ is below a threshold $\theta$, denoted by $\underline{\theta} \in (0, \infty)$, at which the operating profits

$$\pi(\theta, X, \eta) = \max_{k \in \{V,O\}, \ell \in \{N,S\}} \pi_k^\ell(\theta, X, \eta) \tag{4.8}$$

equal zero. Namely, $\underline{\theta}$ is implicitly defined by

$$\pi(\underline{\theta}, X, \eta) = 0. \tag{4.9}$$

This threshold productivity level depends on the sector's aggregate consumption index $X$, i.e., $\underline{\theta}(X)$.

In solving the problem on the right-hand-side of (4.8), a final-good producer effectively chooses the triplet $\left(\beta_k^\ell, w^\ell, f_k^\ell\right)$ that maximizes (4.6). It is straightforward to see that $\pi_k^\ell(\theta, X, \eta)$ is decreasing in both $w^\ell$ and $f_k^\ell$. For this reason final-good producers prefer to organize production so as to minimize both variable and fixed costs. On account of variable costs, Southern manufacturing is preferred to Northern manufacturing regardless of the ownership structure (because $w^N > w^S$). On account of fixed costs, however, the ranking of profit levels is the reverse of the ranking of fixed cost levels in (4.3).

Next note that if the final-good producer could freely choose its fraction of revenue

Figure 4-2: The Profit-Maximizing Distribution of Revenue
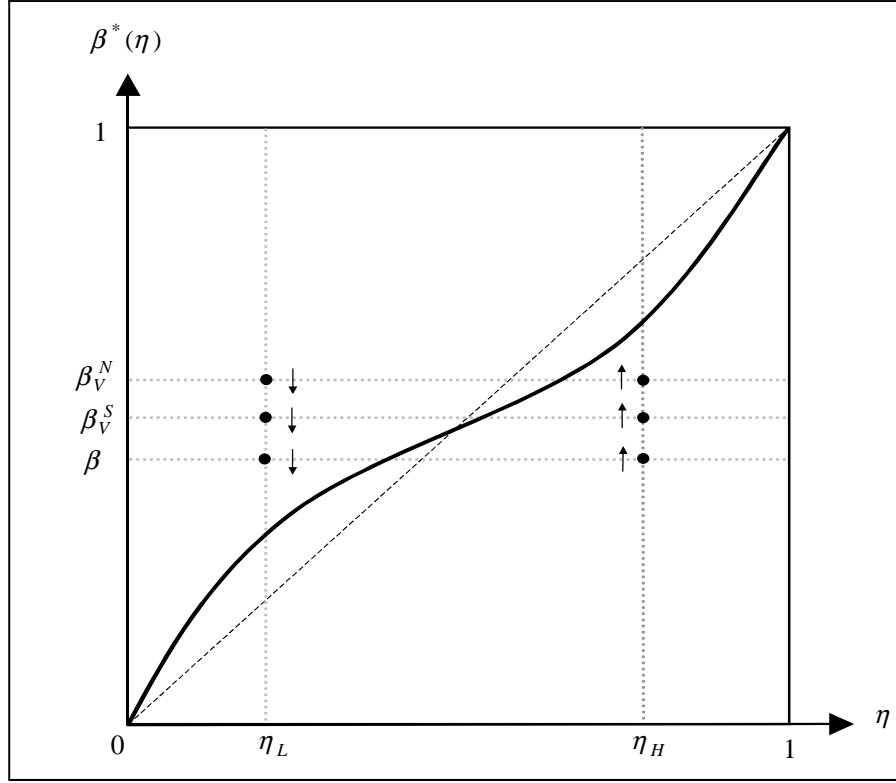


$\beta_k^\ell$, it would choose $\beta^* \in [0,1]$ that maximizes $\psi_k^\ell(\eta)$. This fraction is

$$\beta^*(\eta) = \frac{\eta(\alpha\eta + 1 - \alpha) - \sqrt{\eta(1-\eta)(1-\alpha\eta)(\alpha\eta + 1 - \alpha)}}{2\eta - 1}. \qquad (4.10)$$

Although a higher $\beta_k^\ell$ gives the final-good producer a larger fraction of the revenue, it also induces the supplier of components to produce fewer components. As a result, the final-good producer trades the choice of a larger fraction of the revenue for a smaller revenue level.

The function $\beta^*(\eta)$ is depicted by the solid curve in Figure 4-2. It rises in $\eta$; $\beta^*(0) = 0$ and $\beta^*(1) = 1$.[12] To understand these properties, notice that in the ex-post bargaining neither the final-good producer nor the manufacturing-plant operator appropriate the full

---

[12]Notice also that it does not depend on factor prices and that it is less nonlinear the higher is $\alpha$.

marginal return to their investments in the supply of headquarter services and components, respectively. This leads them to underinvest in the provision of these inputs. Each party's severity of underinvestment is inversely related to the fraction of the surplus that it appropriates. Ex-ante efficiency then requires giving a larger share of the revenue to the party undertaking the relatively more important investment. As a result, the higher the intensity of headquarter services (the larger is $\eta$), the higher is the profit-maximizing fraction of the surplus accruing to the final-good producer (the higher is $\beta^*$).

Following Grossman and Hart (1986), we do not allow a free ex-ante choice of the division rule of the surplus. The choice of ownership structure and the location of the manufacturing of components are the only instruments for affecting the division rule, in the sense that the final-good producer is constrained to choose a $\beta_k^\ell$ in the set $\{\beta_V^N, \beta_O^N, \beta_V^S, \beta_O^S\}$. When $\eta$ is close to 1, higher values of $\beta_k^\ell$ yield higher profits. Given the ordering in (4.5), this implies that the final-good producer would have chosen domestic integration if there were no other differences in the costs and benefits of the competing organizational forms. Conversely, when $\eta$ is close to 0, lower values of $\beta_k^\ell$ yield higher profits, and the final-good producer would have chosen outsourcing in the absence of other differences in the costs and benefits of the organizational forms.

Naturally, there are other differences in the costs and benefits of various organizational forms. As a result, the profit-maximizing choice of an ownership structure and the location of the manufacturing of components depends on a firm's productivity level. When $\theta$ is small, changes in $\beta_k^\ell$ and $w^\ell$ have small effects on profits, because changes in $\psi_k^\ell(\eta)$ have small effects on profits (see (4.6) and (4.7)). Under the circumstances differences in fixed costs dominate the choice of an organizational form, which gives domestic outsourcing a particular advantage. On the other hand, when $\theta$ is large, fixed costs are less important, and the combinations of $\beta_k^\ell$ and $w^\ell$ that raise $\psi_k^\ell(\eta)$ as much as possible are particularly advantageous. We shall see in the next section how these tradoffs play out.

Free entry ensures that, in equilibrium, the expected operating profits of a potential entrant equal the fixed cost of entry. From the discussion above, a firm in sector $j$ that draws a productivity level below $\underline{\theta}(X)$ chooses to exit, because its operating profits are negative. On the other hand, firms with $\theta \geq \underline{\theta}(X)$ stay in the industry, and they choose

organizational forms that maximize their profits. Under the circumstances the free-entry condition can be expressed as

$$\int_{\underline{\theta}(X)}^{\infty} \pi(\theta, X, \eta) \, dG(\theta) = w^N f_E. \tag{4.11}$$

This condition provides an implicit solution to the sector's real consumption index $X$. Using the sector's consumption index, it is then possible to calculate all other variables of interest, such as the threshold productivity level of surviving entrants, the organizational forms of final-good producers with different productivity levels, and the number of entrants.

In order to gain insights into the prevalence of alternative organizational forms and into how they differ across sectors, we focus in what follows on two types of sectors: those with relatively high headquarter intensity and those with relatively low headquarter intensity. Our aim is to characterize the differences in organizational forms between these sector types. Intermediate factor intensities can be similarly analyzed, but they are more complex and provide no new insights.

## 4.4  Organizational Forms

We first consider a sector with low headquarter intensity $\eta$, such that $\beta^*(\eta) < \beta_O^N = \beta_O^S = \beta$. For concreteness, we refer to it as a low-tech sector. This case is depicted in Figure 4-2 by $\eta = \eta_L$, where the arrows indicate the direction in which profits rise with changes in $\beta_k^{\ell}$.[13] Components are particularly important in the production process of a low-tech sector and the profits of firms in this type of sector are decreasing in the fraction of revenue that accrues to the final-good producer. On this account, profits are highest under outsourcing (both domestic and foreign) and lowest under domestic integration. In addition, the fixed costs of outsourcing are lower than the fixed costs of integration. Therefore a final-good producer in a low-tech sector never integrates into manufacturing of components. In choosing between domestic outsourcing and foreign

---

[13]Note that the following analysis applies to every industry in which $\beta^*(\eta) < \beta$; this is necessarily the case when $\eta$ is low enough.

Figure 4-3: Equilibrium in the Low-Tech Sector
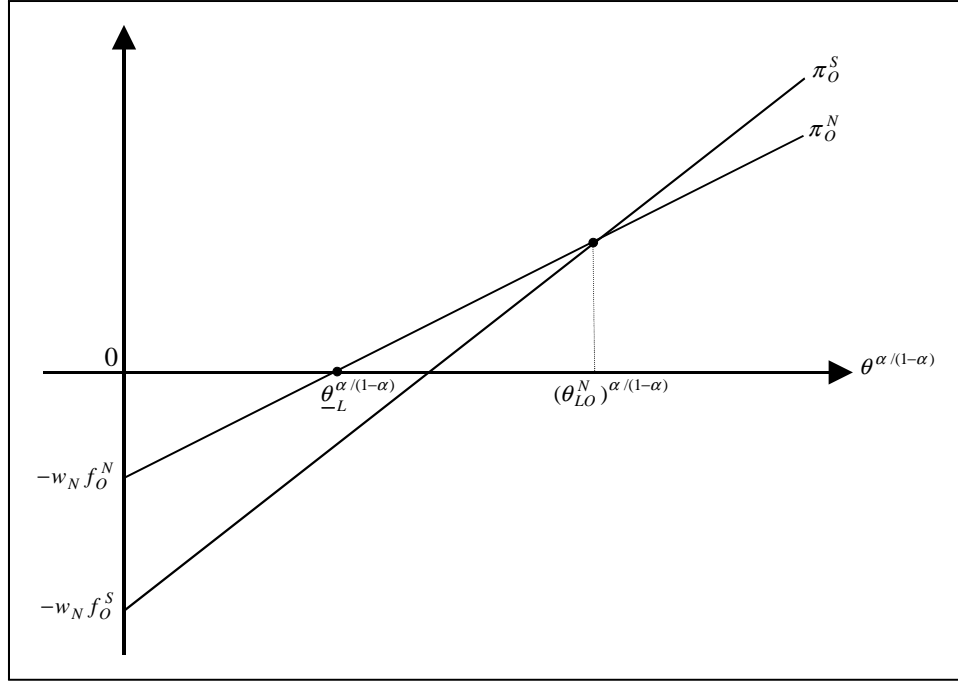


outsourcing, however, such a producer trades-off the lower variable costs of Southern manufacturing against the lower fixed organizational costs in the North.

Two types of equilibria exist in this case. Which type applies depends on whether the cross-country difference in the wage rate is large or small relative to the cross-country difference in the fixed organizational costs of outsourcing. Figure 4-3 depicts the case in which the wage differential is small relative to the fixed-cost differential.[14] The transformed measure of productivity $\theta^{\alpha/(1-\alpha)}$ is measured along the horizontal axis while operating profits are measured along the vertical axis. It is evident from (4.6) that the operating profits $\pi_k^\ell$ are linear in $\theta^{\alpha/(1-\alpha)}$, with the slope being proportional to $\psi_k^\ell(\eta)$ and the intercept being equal to $-w^N f_k^\ell$. The figure depicts profits from outsourcing only, because in a low-tech sector profits from outsourcing in country $\ell$ are higher than profits from integration in country $\ell$. Note also that profits from outsourcing in the South have a steeper slope than profits from outsourcing in the North, because wages are lower

---

[14]That is, $w^N/w^S < \left(f_O^S/f_O^N\right)^{(1-\alpha)/\alpha(1-\eta)}$.

143

in the South and therefore $\psi_O^S(\eta) > \psi_O^N(\eta)$.[15]

Firms with productivity below $\underline{\theta}_L$ expect negative profits under all organizational forms. Therefore they exit the industry. Firms with productivity between $\underline{\theta}_L$ and $\theta_{LO}^N$ attain the highest profits by outsourcing in the North while firms with productivity above $\theta_{LO}^N$ attain the highest profits by outsourcing in the South.[16] The cutoffs $\underline{\theta}_L$ and $\theta_{LO}^N$ are given by

$$\left.\begin{array}{l} \underline{\theta}_L = X^{(\alpha-\mu)/\alpha} \left[\frac{w^N f_O^N}{\psi_O^N(\eta)}\right]^{(1-\alpha)/\alpha}, \\[2ex] \theta_{LO}^N = X^{(\alpha-\mu)/\alpha} \left[\frac{w^N\left(f_O^S - f_O^N\right)}{\psi_O^S(\eta) - \psi_O^N(\eta)}\right]^{(1-\alpha)/\alpha}. \end{array}\right\} \tag{4.12}$$

It is also clear from Figure 4-3 that the intersection point of the two profit lines takes place at a negative profit level when the fixed organizational costs of outsourcing in the South are close to the fixed organizational costs of outsourcing in the North.[17] In this case the threshold productivity level $\underline{\theta}_L$ is defined by the point of intersection of the profit line $\pi_O^S$ with the horizontal axis. As a result, all firms with productivity below this threshold exit while all firms with higher productivity levels outsource in the South. Evidently, in this equilibrium no firm outsources in the North.

We shall treat the case described in Figure 4-3 as the generic case of a low-tech sector. In this event the free entry condition (4.11), together with (4.6) and (4.8), imply

$$w^N X^{(\alpha-\mu)/(1-\alpha)} = \frac{\psi_O^N(\eta)\left[V\left(\theta_{LO}^N\right) - V\left(\underline{\theta}_L\right)\right] + \psi_O^S(\eta)\left[1 - V\left(\theta_{LO}^N\right)\right]}{f_E + f_O^N\left[G\left(\theta_{LO}^N\right) - G\left(\underline{\theta}_L\right)\right] + f_O^S\left[1 - G\left(\theta_{LO}^N\right)\right]}, \tag{4.13}$$

where

$$V(\theta) = \int_0^\theta y^{\alpha/(1-\alpha)} dG(y).$$

Equations (4.12) and (4.13) provide implicit solutions for the cutoffs $\underline{\theta}_L$ and $\theta_{LO}$ and for the aggregate consumption index $X$.

We next consider a sector with high headquarter intensity $\eta$, such that $\beta^*(\eta) > \beta_V^N$. We refer to it as a high-tech sector. A sector of this type is represented by $\eta = \eta_H$ in

---

[15] Equation (4.7) implies that $\psi_O^S(\eta)/\psi_O^N(\eta) = \left(w^N/w^S\right)^{\alpha(1-\eta)/(1-\alpha)} > 1$.

[16] The upper envelope of the profit lines in Figure 4-3 represent the profit function $\pi(\theta, X, \eta)$ (from (4.8)) for low-tech sectors.

[17] This happens when $w^N/w^S > \left(f_O^S/f_O^N\right)^{(1-\alpha)/\alpha(1-\eta)}$.

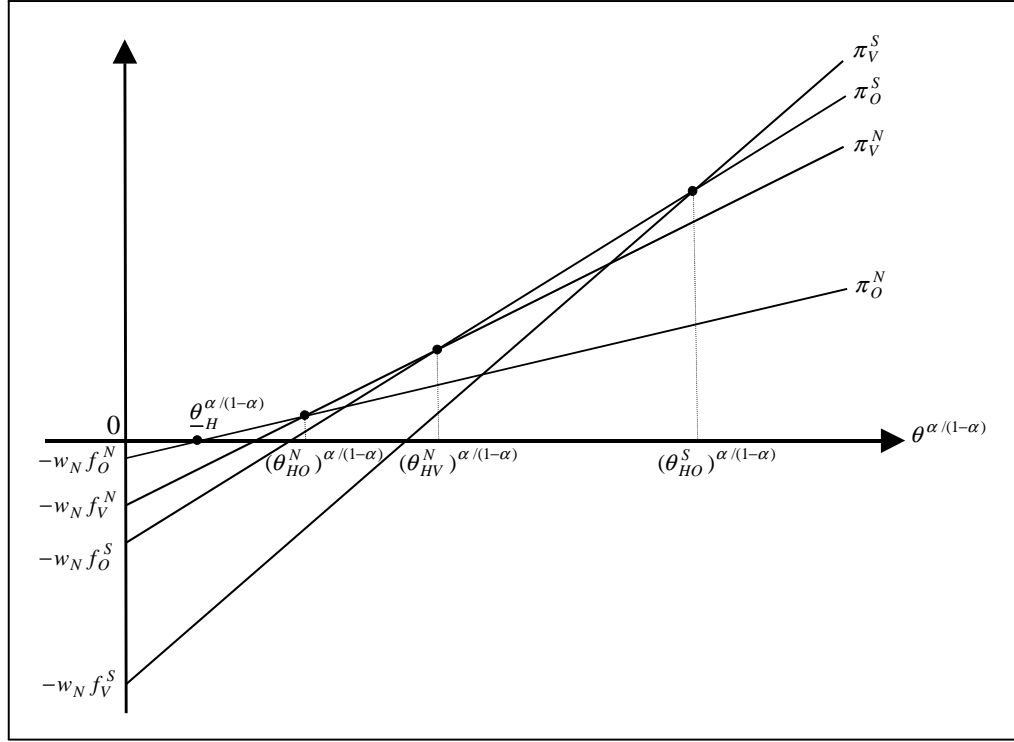Figure 4-4: Equilibrium in the Hi-Tech Sector



Figure 4-2. In this sector profits are increasing in the fraction of revenue $\beta_k^\ell$, as shown by the arrows in the figure. In a high-tech sector the marginal product of headquarter services is high, making underinvestment by final-good producers especially costly. As a result, ex-ante maximization of value favors integration over outsourcing as long as there are no other differences in the benefits and costs of alternative organizational forms. In particular, equation (4.7) implies that in a high-tech sector $\psi_V^\ell(\eta) > \psi_O^\ell(\eta)$ for $\ell = N, S$. Namely, in every country the slope of the profit function is steeper when the firm is vertically integrated than when it outsources components.

Now compare the slope of the profit function of an integrated firm that produces components in the North with the slope of the profit function of a firm that outsources its components in the South, where the slope is measured relative to the $\theta^{\alpha/(1-\alpha)}$ axis (see Figure 4-4). The integrated firm has the advantage of being able to save a positive fraction $\delta^N$ of output when it severs its ties with the operator of the production facility

of components, while the outsourcing firm saves no output at all when it severs its ties with the arm's-length supplier of components. On the other hand, the integrated firm faces higher production costs, because the wage rate is higher in the North. For these reasons the profit function of the firm outsourcing in the South can be steeper or flatter than the profit function of the integrated firm in the North, depending on whether $\delta^N$ is small or large, respectively, relative to the wage differential. That is, $\psi_O^S(\eta)$ can be larger or smaller than $\psi_V^N(\eta)$.

First consider the case in which the wage differential is large relative to $\delta^N$, so that $\psi_O^S(\eta) > \psi_V^N(\eta)$.[18] Under these circumstances

$$\psi_V^S(\eta) > \psi_O^S(\eta) > \psi_V^N(\eta) > \psi_O^N(\eta). \tag{4.14}$$

Given the ordering in (4.3) and (4.14), the order of the intercepts and the slopes of the profit functions are as depicted in Figure 4-4. The intersection point of $\pi_O^N$ with the horizontal axis is to the left of the intersection point of this profit line with $\pi_V^N$, the latter intersection point is to the left of the intersection point of $\pi_V^N$ with $\pi_O^S$, and this last intersection point is to the left of the intersection point of $\pi_O^S$ with $\pi_V^S$. We take this situation to be the generic case of a high-tech sector. In this case all firms with productivity below $\underline{\theta}_H$ exit the industry, those with productivity between $\underline{\theta}_H$ and $\theta_{HO}^N$ outsource in the North, those with productivity between $\theta_{HO}^N$ and $\theta_{HV}^N$ integrate in the North, those with productivity between $\theta_{HV}^N$ and $\theta_{HO}^S$ outsource in the South, and those with productivity above $\theta_{HO}^S$ integrate in the South (engage in vertical FDI). It is easy to see that either one of the first three organizational forms may not exist in equilibrium, but that the last one always exists. That is, there always exist high-productivity final-good producers who choose to manufacture components in the South. The cutoffs depicted in

---

[18]This condition is satisfied if and only if

$$\left(\frac{w^N}{w^S}\right)^{1-\eta} > \frac{\left\{1 - \alpha\left[\beta_V^N \eta + \left(1 - \beta_V^N\right)(1-\eta)\right]\right\}^{(1-\alpha)/\alpha}\left(\beta_V^N\right)^\eta\left(1 - \beta_V^N\right)^{1-\eta}}{\left\{1 - \alpha\left[\beta\eta + (1-\beta)(1-\eta)\right]\right\}^{(1-\alpha)/\alpha}\beta^\eta(1-\beta)^{1-\eta}}$$

(see (4.7)). This inequality always holds in low-tech sectors (in which the right-hand-side is smaller than one), but may not hold in high-tech sectors (in which the right-hand-side is larger than one).

---

Figure 4-4 are given by

$$
\left.
\begin{aligned}
\underline{\theta}_H &= X^{(\alpha-\mu)/\alpha} \left[ \frac{w^N f_O^N}{\psi_O^N(\eta)} \right]^{(1-\alpha)/\alpha}, \\
\theta_{HO}^N &= X^{(\alpha-\mu)/\alpha} \left[ \frac{w^N\left(f_V^N - f_O^N\right)}{\psi_V^N(\eta) - \psi_O^N(\eta)} \right]^{(1-\alpha)/\alpha}, \\
\theta_{HV}^N &= X^{(\alpha-\mu)/\alpha} \left[ \frac{w^N\left(f_O^S - f_V^N\right)}{\psi_O^S(\eta) - \psi_V^N(\eta)} \right]^{(1-\alpha)/\alpha} \\
\theta_{HO}^S &= X^{(\alpha-\mu)/\alpha} \left[ \frac{w^N\left(f_V^S - f_O^S\right)}{\psi_V^S(\eta) - \psi_O^S(\eta)} \right]^{(1-\alpha)/\alpha}.
\end{aligned}
\right\}
\tag{4.15}
$$

We can also use the free entry condition (4.11) to derive an equation that is analogous to (4.13). This equation together with (4.15) can then be used to solve for the cutoffs and the consumption index $X$.[19]

Next consider cases in which $\psi_O^S(\eta) < \psi_V^N(\eta)$, i.e., the profit function from integration in the North has a larger slope than the profit function from outsourcing in the South. This happens when the wage differential is small relative to $\delta^N$.[20]  In this event the ordering in (4.14) is not preserved. There are two possibilities: either $\psi_V^S(\eta) > \psi_V^N(\eta) > \psi_O^S(\eta) > \psi_O^N(\eta)$ or $\psi_V^N(\eta) > \psi_V^S(\eta) > \psi_O^S(\eta) > \psi_O^N(\eta)$.[21]  The former case arises when the wage differential is small relative to $\delta^N$, but not so small relative to the difference between $\delta^N$ and $\delta^S$. On the other hand, the latter case arises when the wage differential

---

[19]Suppose instead that the fixed costs of outsourcing are higher than the fixed costs of integration in each one of the countries, but that the fixed costs of integration in the South are higher than the fixed costs of outsourcing in the North. Then, in a high-tech sector integration dominates outsourcing in each one of the countries, because the fixed costs of integration are lower than the fixed costs of outsourcing and the profit function of an integrated firm is steeper than the profit function of an outsourcing firm. As a result, no firm outsources. Amongst the final-good producers who stay in the industry, low-productivity firms integrate in the North while high-productivity firms integrate in the South. The reversal of the ordering of the fixed costs also affects the sorting of firms by organizational form in low-tech sectors. Now, in a low-tech sector integration may dominate outsourcing at certain productivity levels. In particular, the least productive final-good producers who stay in the industry may integrate in the North, some more productive firms may outsource in the North, still higher-productivity firms may integrate in the South, and the most productive firms outsource in the South.

[20]Namely,

$$
\left( \frac{w^N}{w^S} \right)^{1-\eta} < \frac{\left\{ 1 - \alpha \left[ \beta_V^N \eta + \left(1 - \beta_V^N\right)(1-\eta) \right] \right\}^{(1-\alpha)/\alpha} \left(\beta_V^N\right)^\eta \left(1 - \beta_V^N\right)^{1-\eta}}{\left\{ 1 - \alpha \left[ \beta\eta + (1-\beta)(1-\eta) \right] \right\}^{(1-\alpha)/\alpha} \beta^\eta (1-\beta)^{1-\eta}}.
$$

[21]Note that $\psi_O^S(\eta) > \psi_O^N(\eta)$ as long as the wage rate is lower in the South.

is small even relative to the difference between $\delta^N$ and $\delta^S$. If, for example, the wage rates are almost identical, then the fact that an integrated final-good producer in the North can save a larger fraction of output than an integrated final-good producer in the South can when both sever their ties with the components manufacturers, makes the former's profits more sensitive to productivity changes than the latter's. As a result, $\psi_V^N(\eta) > \psi_V^S(\eta)$.

In the first case, when $\psi_V^S(\eta) > \psi_V^N(\eta) > \psi_O^S(\eta) > \psi_O^N(\eta)$, integration in the North dominates outsourcing in the South, because it has lower fixed costs of organization and higher profits per unit productivity $\theta$. Namely, the profit line $\pi_V^N$ in Figure 4-4 has a higher intercept and a larger slope than $\pi_O^S$. In this event no firm chooses to outsource in the South, and the model predicts that — amongst the firms that do not exit the industry — low-productivity firms outsource in the North, high-productivity firms integrate in the South, and firms with intermediate productivity levels integrate in the North.

In the second case, when $\psi_V^N(\eta) > \psi_V^S(\eta) > \psi_O^S(\eta) > \psi_O^N(\eta)$, integration in the North dominates both outsourcing in the South and integration in the South. As a result, at most two organizational forms survive in equilibrium: low-productivity firms that outsource in the North and high-productivity firms that integrate in the North.

We have characterized the organizational forms in low-tech and high-tech sectors. The choice of organizational forms by firms with different productivity levels is depicted in Figure 4-5. This figure describes the generic cases. First, low-tech firms do not integrate into the production of components; low-productivity firms outsource them in the North while high-productivity firms outsource them in the South. The least productive firms exit. On the other hand, integration always takes place in high-tech sectors. The most productive firms integrate in the South via foreign direct investment while somewhat less productive firms outsource in the South. Firms with even lower productivity acquire components in the North. Amongst those, the more productive firms integrate while the less productive outsource. The least productivity firms exit. Note that surviving firms with the lowest productivity outsource in the North in both low-tech and high-tech sectors. And more generally, less productive firms acquire components in the North while more productive firms acquire them in the South.

This sorting pattern differs from the sorting pattern derived by Grossman and Help-

Figure 4-5: Organizational Forms
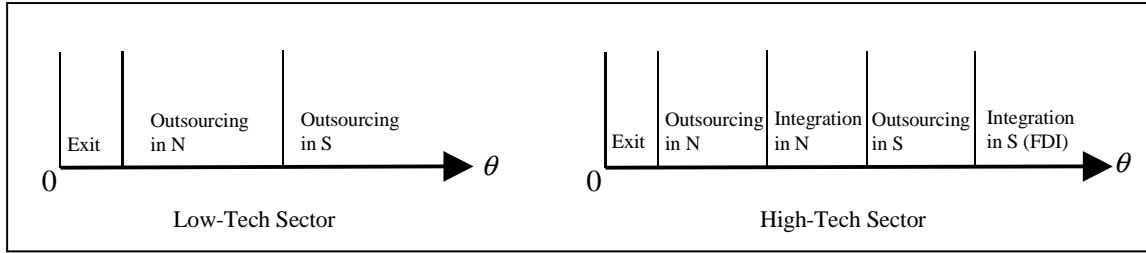


man (2002b) for organizational structures that use managerial incentives à la Holmstrom and Milgrom (1994).[22] Contrary to our results, in their model surviving low-productivity firms acquire components in the South. Within this group less-productive firms outsource while more-productivity firms integrate via FDI. While no one outsources inputs in the North, there exist modestly-high productive firms that integrate in the North. However, the most-productive firms, like the least-productive firms, outsource in the South.

Evidently, the two models predict very different sorting patterns. It would be interesting and useful to gauge which pattern better fits reality. There exists, however, no evidence that bears directly on this question. And it is hard to see how to test these predictions with the available data.

It also is important to bear in mind that the most suitable theory of the firm can differ across sectors. Namely, the Grossman-Hart property-rights approach may be most suitable for some industries while the Holmstrom-Milgrom managerial-incentives approach may be most suitable for others.[23] This possibility would complicate every empirical analysis that tries to explain the cross-sectional variation in organizational forms. An appreciation of this possibility also raises an interesting theoretical question, the answer to which may help to design an empirical strategy: How do companies in a particular industry choose between the property-rights approach and the managerial-incentives ap-

---

[22]They did not distinguish between low- and high-tech sectors, however, although one can interpret their production technology as having $\eta = 0$, i.e., a zero output elasticity with respect to headquarter services. For this reason a comparison of the cross-section variation of organizational forms that is based on the low-tech high-tech distinction cannot be made with their work.

[23]The empowerment of workers may also be an important determinant of the structure of firms. Puga and Trefler (2002) and Marin and Verdier (2003) have developed general equilibrium frameworks in which every firm chooses endogenously the structure of authority within the organization.

proach in the organization of their activities? Or, more broadly, how do they choose endogenously the structure of ownership and incentives? To sort out the determinants of the organizational forms of industries together with an endogenous choice of incentives schemes and the structure of ownership is a major challenge for future research.

## 4.5 Prevalence of Organizational Forms

Our analysis has so far focused on the sorting patterns of firms into different organizational forms: outsourcing in the North, integration in the North, outsourcing in the South, and integration in the South. How prevalent are these organizational forms? And what determines their relative prevalence in different industries? To answer these questions we need a measure of prevalence. We choose the relative fractions of the varieties of final goods that are produced under these organizational forms as our measures of relative prevalence. We show in the appendix, however, that our results do not depend on this choice, in the sense that other measures — such as market shares of final goods — deliver similar results.[24]

Following Melitz (2002) and Helpman, Melitz and Yeaple (2002), we choose $G(\theta)$ to be a Pareto distribution with shape $k$.[25] That is,

$$G(\theta) = 1 - \left(\frac{b}{\theta}\right)^k \quad \text{for } \theta \geq b > 0. \tag{4.16}$$

Under this assumption, the distribution of sales is also Pareto, a feature consistent with the evidence.[26] We use this distribution in the following analysis. As in the previous section, we focus on low- and high-tech sectors.

### 4.5.1 Low-tech sector

Recall that in a low-tech sector no firm integrates into the production of intermediates, because the outsourcing of components delivers higher operating profits. In the generic

---

[24]See also Grossman and Helpman (2002b) for this point.

[25]$k$ has to be large enough to ensures a finite variance of the size distribution of firms.

[26]See Axtell (2001) and Helpman, Melitz and Yeaple (2002).

case depicted in Figure 4-3, final-good producers with productivity below $\underline{\theta}_L$ exit the industry, those with productivity between $\underline{\theta}_L$ and $\theta_{LO}^N$ outsource in the North, and firms with higher productivity levels outsource in the South. Denote by $\sigma_{LO}^\ell$ the fraction of *active* final-good producers who outsource in country $\ell$. Then in the generic low-tech sector $\sigma_{LO}^S = \left[1 - G\left(\theta_{LO}^N\right)\right] / \left[1 - G\left(\underline{\theta}_L\right)\right]$ and $\sigma_{LO}^N = 1 - \sigma_{LO}^S$. The Pareto distribution (4.16) then implies that $\sigma_{LO}^S = \left(\underline{\theta}_L/\theta_{LO}^N\right)^k$. Substituting (4.12) into this expression, we obtain

$$\sigma_{LO}^S = \left[\frac{\psi_O^S(\eta) - \psi_O^N(\eta)}{\psi_O^N(\eta)} \frac{f_O^N}{f_O^S - f_O^N}\right]^{k(1-\alpha)/\alpha}. \tag{4.17}$$

First consider the effect of the Southern wage rate. A lower wage in the South raises the profitability of outsourcing in the South, by increasing $\psi_O^S(\eta)$. In this event, (4.17) implies a rise in the share of final-good producers that outsource components in the South. It can also be shown that the threshold productivity level $\underline{\theta}_L$ is higher the lower the wage rate in the South. The lower wage raises profits from outsourcing components in the South, therefore shifting upwards the profit line $\pi_O^S$ in Figure 4-3. But this raises the expected profits of entrants into the industry, attracting new producers of final goods. As a result the real consumption index $X$ rises, shifting down both profit lines. The final outcome is a higher threshold $\underline{\theta}_L$, which implies that a larger fraction of the final-good producers who enter the industry exit upon learning their productivity level. Evidently, the lower wage in the South induces a reorganization among the final-good producers in the North that leads them to rely more on arm's-length imports of components.

The model can easily be extended to incorporate transport costs for intermediate inputs. If the shipment of components is subjected to melting-iceberg-type transport costs, then a fall in transport costs is very similar to a decline in the Southern wage rate. It follows that lower transport costs lead to exit of a larger fraction of entrants (as in Melitz, 2002) and to a larger fraction of final-good producers who outsource components in the South.[27]

Second, we have assumed for simplicity that an outsourcing final-good producer ap-

---

[27]In the U.S. manufacturing sector, the sum of tariff duties and freight costs has steadily fallen from 11.3% of the Customs value of imports in 1974 to 5.1% in 2001. We computed these figures from data available on Robert Feenstra's website.

propriates a fraction $\beta$ of the surplus from its relationship with the supplier of parts, irrespective of whether the supplier is in the North or in the South. Imagine, however, a situation in which this fraction can differ between the countries, and that the final-good producer now gets a smaller fraction of the surplus from outsourcing in the South (but still higher than $\beta^*(\eta)$, so that the condition of a low-tech sector remains valid). Such a decline in the bargaining power in the South raises the profitability of outsourcing in the South via an increase in $\psi_O^S(\eta)$. As a result, the fraction of final-good producers who outsource in the South rises.

Third, consider an increase in the dispersion of productivity, which is represented by a decline of $k$. Since the expression in the brackets on the right hand side of (4.17) represents the ratio of the cutoffs $\underline{\theta}_L/\theta_{LO}^N$ and this ratio is smaller than one, it follows that a rise in dispersion raises the fraction of final-good producers who outsource in the South.[28]

Finally, note that the degree of a sector's headquarter intensity affects its relative prevalence of outsourcing in the two countries. Since $\psi_O^S(\eta)/\psi_O^N(\eta) = \left(w^N/w^S\right)^{(1-\eta)\alpha/(1-\alpha)}$ (see (4.7)), it follows that among the low-tech sectors those whose technology is more intensive in headquarter services outsource relatively less in the South. Intuitively, the less important are components in the production of the final good, the less important are the cost savings from outsourcing components in the South as compared to the higher fixed organizational costs of this activity.

### 4.5.2 High-tech sector

In the generic case of the high-tech sector, there are four organizational forms, ordered from low- to high-productivity firms: outsourcing in the North, integration in the North, outsourcing in the South and integration in the South (see Figures 4-4 and 4-5). We denote by $\sigma_{Hk}^\ell$ the share of products that are supplied with the organizational form $(k, \ell)$, where $k$ is the ownership structure and $\ell$ is the location of production of components.

---

[28]This is similar, in terms of the mechanism at work, to the finding in Melitz (2002) that more dispersion raises the share of exporting firms in domestic output, and the finding in Helpman, Melitz and Yeaple (2002) that more dispersion raises horizontal FDI relative to exports.

Using the Pareto distribution (4.16) and the cutoffs (4.15), these shares are

$$
\left.
\begin{aligned}
\sigma_{HO}^{N} &= 1 - \left[ \frac{\psi_V^N(\eta) - \psi_O^N(\eta)}{\psi_O^N(\eta)} \frac{f_O^N}{f_V^N - f_O^N} \right]^{k(1-\alpha)/\alpha}, \\
\sigma_{HV}^{N} &= \left[ \frac{\psi_V^N(\eta) - \psi_O^N(\eta)}{\psi_O^N(\eta)} \frac{f_O^N}{f_V^N - f_O^N} \right]^{k(1-\alpha)/\alpha} - \left[ \frac{\psi_O^S(\eta) - \psi_V^N(\eta)}{\psi_O^N(\eta)} \frac{f_O^N}{f_O^S - f_V^N} \right]^{k(1-\alpha)/\alpha}, \\
\sigma_{HO}^{S} &= \left[ \frac{\psi_O^S(\eta) - \psi_V^N(\eta)}{\psi_O^N(\eta)} \frac{f_O^N}{f_O^S - f_V^N} \right]^{k(1-\alpha)/\alpha} - \left[ \frac{\psi_V^S(\eta) - \psi_O^S(\eta)}{\psi_O^N(\eta)} \frac{f_O^N}{f_V^S - f_O^S} \right]^{k(1-\alpha)/\alpha}, \\
\sigma_{HV}^{S} &= \left[ \frac{\psi_V^S(\eta) - \psi_O^S(\eta)}{\psi_O^N(\eta)} \frac{f_O^N}{f_V^S - f_O^S} \right]^{k(1-\alpha)/\alpha}.
\end{aligned}
\right\}
\tag{4.18}
$$

We again first consider a lowering of the wage rate in the South. Lower wages in the South raise the profitability of foreign integration and foreign outsourcing. In particular, (4.7) implies that $\psi_V^S(\eta)$ and $\psi_O^S(\eta)$ increase while $\psi_V^N(\eta)$ and $\psi_O^N(\eta)$ do not change. It then follows from (4.18) that $\sigma_{HO}^N$ does not change. Namely, the share of products that are supplied by final-good producers who outsource in the North remains the same. On the other hand, the share of products supplied by vertically integrated producers in the North, $\sigma_{HV}^N$, declines. The reason is that low-productivity firms that outsource in the North are too far from productivity levels that make the acquisition of inputs in the South profitable. As a result, small changes in the profitability of importing inputs, be it through arm's-length transactions or via FDI, does not make the purchase of inputs in the South attractive to these firms. On the other hand, amongst the integrated producers in the North the most productive are indifferent between integration in the North and outsourcing in the South. Therefore, for these firms a decline in the South's wage rate tilts the balance in favor of outsourcing in the South. For this reason the share of final-good producers who outsource in the South, $\sigma_{HO}^S$, rises.[29] Finally, the share of firms that integrate in the South, $\sigma_{HV}^S$, also rises. Evidently, lower labor costs in the South induce a reorganization that favors the acquisition of components in the South, but it has a disproportionately large effect on outsourcing as compared to FDI. At the same time the unfavorable effect on the acquisition of inputs in the North falls disproportionately on integration. It follows that outsourcing rises overall relative to integration.

A fall in transport costs of intermediate inputs has the same effects as a fall in $w^S$.

---

[29] This is easy to see from (4.18) by noting that the ratio $\psi_V^S(\eta)/\psi_O^S(\eta)$ is independent of the wage rate $w^S$.

It is interesting to note that the recent trends described in the introduction are in line with the model's predictions about falling costs of doing business in the South. Feenstra and Hanson (1996) point out that transport costs have declined and foreign assembly has increased both in-house and at arm's length. Furthermore, Table 1 suggested that the growth of foreign outsourcing might have outpaced that of foreign direct investment. Finally, as predicted by the model, U.S. domestic outsourcing seems to have increased relative to U.S. domestic integration at a time of falling trade barriers (see Figure 4-1).[30]

Second, consider the effect of $\delta^\ell$, the share of output that a final-good producer who is integrated in country $\ell$ retains in case it severs its ties with the operator of the production unit of components. We start with an increase in this share in the South. This improves the outside option of an integrated producer in the South in its bargaining with the operator of the production unit of components. The better outside option translates into higher effective bargaining power, as measured by $\beta_V^S$. As is evident from (4.7), a higher $\beta_V^S$ raises $\psi_V^S(\eta)$ without affecting the slopes of other profit functions. Equations (4.18) then imply that the shares of products that are supplied by final-good producers who acquire components in the North, either via outsourcing or integration, do not change. In this event, the fraction of final goods that use imported components does not change too, except that amongst those who use imported components the share of outsourcing firms declines while the share of integrated firms rises.

An increase in $\delta^N$ raises the effective bargaining power of an integrated producer in the North. As a result, integration in the North becomes more profitable and the slope $\psi_V^N(\eta)$ rises; the other slopes do not change. It then follows from (4.18) that the share of final-good producers who outsource in the North declines, the share of integrated firms in the North rises, the share of outsourcing firms in the South declines, and the share of integrated firms in the South does not change. The interesting implication is that the improvement in the attractiveness of domestic integration changes the relative prevalence of foreign outsourcing relative to FDI in favor of the latter.

Third, consider an increase in the primitive bargaining power $\beta$. It can be shown

---

[30]As in the a low-tech sector, lower labor costs in the South or lower transport costs of intermediates increase the cutoff productivity level below which final-good producers exit the industry in a high-tech sector. This implies a higher proportion of exiting firms.

that it reduces the ratios $\psi_V^S(\eta)/\psi_O^\ell(\eta)$ and $\psi_V^N(\eta)/\psi_O^\ell(\eta)$ for $\ell = N, S$ as well as $\psi_V^N(\eta)/\psi_V^S(\eta)$. The reason is that an increase in $\beta$ shifts the bargaining power in favor of the final-good producer, regardless of ownership structure. As a result, the final-good producer finds it relatively more profitable to outsource. In this event the share of final-good producers who outsource components rises in the North as well as in the South. On the other hand, the share of final-good producers who integrate declines in the North as well as in the South. Moreover, the fraction of firms that import components rises. That is, the rise in the share of outsourcing firms in the South is larger than the fall in the share of firms that engage in foreign direct investment. It follows that an increase in the final-good producer's bargaining power biases the acquisition of inputs towards imports on the one hand and towards outsourcing as opposed to integration on the other.

Fourth, we examine an increase in the degree of dispersion of the distribution of productivity, as represented by a fall in the shape parameter $k$. It is evident from (4.18) that a decline in $k$ reduces the share of firms that outsource in the North and increases the share of firms that integrate in the South. The effect on the share of firms that integrate in the North or outsource in the South is ambiguous, however. Yet the share of final-good producers who import components from the South rises, and so does the prevalence of FDI relative to outsourcing in the South (i.e., the ratio $\sigma_{HV}^S/\sigma_{HO}^S$) and the prevalence of integration relative to outsourcing in the North (i.e., the ratio $\sigma_{HV}^N/\sigma_{HO}^N$).

Finally, we consider variations in headquarter intensity. In sectors with higher headquarter intensity domestic outsourcing is favored relative to foreign outsourcing and integration is favored relative to outsourcing. That is, $\psi_O^N(\eta)/\psi_O^S(\eta)$ and $\psi_V^\ell(\eta)/\psi_O^\ell(\eta)$ for $\ell = N, S$ are higher in sectors with larger values of $\eta$.[31] Equations (4.18) then imply that the share of final-good producers who outsource in the North falls with $\eta$ while the share of final-good producers who integrate in the North rises. Moreover, the sum of these two shares goes up, implying that a larger $\eta$ reduces the fraction of firms that import components. As for the composition of imported components, we cannot sign the effects of $\eta$ on the share of firms that import from integrated subsidiaries. Nevertheless, (4.18) implies that the ratio $\sigma_{HV}^S/\sigma_{HO}^S$ rises and, hence, that $\sigma_{HO}^S$ falls. Namely, FDI becomes

---

[31]See Antràs, (2002a).

more prevalent relative to arm's-length imports. It follows that in a cross-section of high-tech sectors the relative prevalence of integration rises and the relative prevalence of outsourcing falls with headquarter intensity. This prediction is in line with the findings of Antràs (2002a), who shows that in a panel of 23 manufacturing industries and four years of data, the share of intra-firm imports in total U.S. imports is significantly higher, the higher the R&D intensity of the industry.[32]

## 4.6 Concluding Comments

We have developed a theoretical framework for studying global sourcing strategies. In our model, heterogeneous final-good producers choose organizational forms. That is, they choose ownership structures and locations for the production of intermediate inputs. Headquarter services are always produced in the home country (the North). Intermediate inputs can be produced at home or in the low-wage South, and the production of intermediates can be owned by the final-good producer or by an independent supplier. When a final-good producer owns the production unit of components and this unit is located in the North, the organizational form is one of standard vertical integration. When, on the other hand, the production unit of the intermediate inputs is located in the South, the organizational form is one of integration with vertical foreign direct investment. This type of FDI generates intra-firm international trade. A final-good producer who does not integrate into the production of components outsources them to independent suppliers. Such a final-good producer can outsource in the home country or in the South. In the latter case outsourcing generates international trade at arm's length.

Final-good producers and operators of components production units make relationship-specific investments which are governed by imperfect contracts. In choosing between a domestic and a foreign supplier of parts, a final-good producer trades off the benefits of lower variable costs in the South against the benefits of lower fixed costs in the North.

---

[32]Controlling for several industry characteristics, Antràs (2002a) finds that a 1% increase in the ratio of industry R&D expenditures over industry sales leads to a 0.42% increase in the fraction of that industry's U.S. imports that are transacted within firm boundaries. The effect is significant at the 1% significance level.

On the other hand, in choosing between vertical integration and outsourcing in one of the countries, the final-good producer trades off the benefits of ownership from vertical integration against the benefits of better incentives for the supplier of parts under outsourcing. These tradeoffs induce firms with different productivity levels to sort by organizational form. We show that the equilibrium sorting patterns depend on the wage differential between the North and the South, on the ownership advantage in each one of the countries (as measured by the fraction of output that a final-good producer can recoup in the event of a breakup of his relationship with an integrated supplier of components), on the distribution of the bargaining power between final-good producers and operators of components production facilities, and on the headquarter intensity of the production process.

A key result is that high-productivity firms acquire intermediate inputs in the South while low-productivity firms acquire them in the North. Amongst the final-good producers who acquire inputs in the same country the low-productivity firms outsource while the high-productivity firms integrate. In sectors with a very low intensity of headquarter services no firm integrates; low-productivity firms outsource at home while high-productivity firms outsource abroad.

We construct industry equilibria and use them to characterize the relative prevalence of alternative organizational forms. By relative prevalence we mean the fraction of final-good producers who choose a particular organizational form. Relative prevalence depends on all the features of an industry that affect the sorting pattern of its firms into various organizational forms. In addition, it depends on the degree of dispersion of productivity across the industry's firms. Using these relationships, we describe how differences in industry characteristics affect the relative prevalence of various organizational forms.

Two results stand out. First, sectors with more dispersion of productivity rely more on imported inputs. And moreover, amongst the headquarter-intensive final-good producers who acquire inputs in a particular country, the number of integrated firms is higher relative to the number of outsourcing firms the more dispersed is productivity within the sector. Second, the higher a sector's headquarter intensity, the less it relies on imported intermediate inputs. And amongst the headquarter-intensive final-good producers who acquire inputs in a particular country, the number of integrated firms is higher relative

to the number of outsourcing firms the more headquarter intensive is the sector.

Our model has also interesting implications for the effects of a widening of the wage gap between the North and the South, or a reduction of the trading costs of intermediate inputs (both changes produce similar outcomes). As one would expect, reducing the relative cost of foreign sourcing raises the prevalence of organizational forms that rely on imported inputs. Importantly, however, such shifts in costs also affect the relative attractiveness of outsourcing versus integration. Not only does it raise the relative prevalence of outsourcing overall, it also raises the number of outsourcing final-good producers relative to the number of integrated final-good producers in every country. This means in particular that it raises arm's-length trade relative to intra-firm trade.

As is evident from these results, our model provides rich predictions about issues of central concern to the quest for a better understanding of the changing landscape of foreign trade and investment. Since we laid out the empirical motivation for this study in the introduction, it suffices to point out in these concluding comments that our approach helps to better appreciate the complexity of trade and investment in a world in which firms choose endogenously their organizational forms. It also should help in designing empirical studies of the ever evolving world trading system.

## 4.7 Appendix

In the main text, we measured the relative prevalence of different organizational forms with the fraction of final-good varieties produced under each type of organization. In this appendix we show that the use of other measures yields similar results.

First consider the case of market shares, i.e., the fraction of industry sales captured by each organizational form. It is straightforward to show that firm revenues can be expressed as:

$$R_k^\ell (\theta, X, \eta) = \frac{X^{(\mu-\alpha)/(1-\alpha)} \theta^{\alpha/(1-\alpha)} \psi_k^\ell (\eta)}{1 - \alpha \left[ \beta_k^\ell \eta + \left( 1 - \beta_k^\ell \right) (1 - \eta) \right]}.$$

Therefore in the generic low-tech sector, the market share of foreign outsourcing is

$$\xi_{LO}^S = \frac{\left[ 1 - V \left( \theta_{LO}^N \right) \right] \rho_O^S (\eta)}{\left[ 1 - V \left( \theta_{LO}^N \right) \right] \rho_O^S (\eta) + \left[ V \left( \theta_{LO}^N \right) - V \left( \underline{\theta}_L \right) \right] \rho_O^N (\eta)}. \tag{4.19}$$

where

$$\rho_k^\ell (\eta) = \frac{\psi_k^\ell (\eta)}{1 - \alpha \left[ \beta_k^\ell \eta + \left( 1 - \beta_k^\ell \right) (1 - \eta) \right]} = \left[ \frac{1}{\alpha} \left( \frac{w^N}{\beta_k^\ell} \right)^\eta \left( \frac{w^\ell}{1 - \beta_k^\ell} \right)^{1-\eta} \right]^{-\alpha/(1-\alpha)}. \tag{4.20}$$

When the productivity index $\theta$ is drawn from a Pareto distribution with the shape parameter $k$, the distribution of firm sales is also Pareto with the shape parameter $k - \alpha/(1-\alpha)$. Making use of the properties of the Pareto distribution, (4.19) can be expressed as:

$$\xi_{LO}^S = \frac{1}{1 + \left[ \left( \frac{\psi_O^N(\eta) \left( f_O^S - f_O^N \right)}{\left[ \psi_O^S(\eta) - \psi_O^N(\eta) \right] f_O^N} \right)^{k(1-\alpha)/\alpha - 1} - 1 \right] \frac{\rho_O^N(\eta)}{\rho_O^S(\eta)}}.$$

Because $\beta_O^N = \beta_O^S = \beta$, it follows that $\rho_O^N (\eta) / \rho_O^S (\eta) = \psi_O^N (\eta) / \psi_O^S (\eta)$, and $\xi_{LO}^S$ is increasing in $\psi_O^S (\eta) / \psi_O^N (\eta)$. This implies that, as in the main text, the prevalence of Southern outsourcing is decreasing in the Southern wage rate, in transport costs, and in the importance of headquarter services as measured by $\eta$. Furthermore, because $\theta_{LO}^N > \underline{\theta}_L$, it is straightforward to show that an increase in dispersion (a fall in $k$) raises

the market share of final-good producers outsourcing in the South. Finally, a fall in the South's bargaining power increases $\psi_O^S(\eta)$ and $\rho_O^S(\eta)$ when $\eta < \beta$, a condition that may or may not be more restrictive than the condition that defines the low-tech sector (i.e., $\beta^*(\eta) < \beta$).[33] When $\eta < \beta$, a fall in the bargaining power in the South raises the market share of Southern outsourcing. When, instead, $\eta > \beta$, the effect is ambiguous.

In the generic high-tech sector, total sale revenues are $X^{(\mu-\alpha)/(1-\alpha)}\theta^{\alpha/(1-\alpha)}\widehat{R}(\eta)$, where $\widehat{R}(\eta)$ is given by:

$$
\begin{aligned}
\widehat{R}(\eta) &= \left[V\left(\theta_{HO}^N\right) - V\left(\underline{\theta}_H\right)\right]\rho_O^N(\eta) + \left[V\left(\theta_{HV}^N\right) - V\left(\theta_{HO}^N\right)\right]\rho_V^N(\eta) + \\
&\quad + \left[V\left(\theta_{HO}^S\right) - V\left(\theta_{HV}^N\right)\right]\rho_O^S(\eta) + \left[1 - V\left(\theta_{HO}^S\right)\right]\rho_V^S(\eta), \quad\quad (4.21)
\end{aligned}
$$

and $\rho_k^\ell(\eta)$ is defined in (4.20). The market share of each type of organizational form is then:

$$
\left.
\begin{aligned}
\xi_{HO}^N &= \left[V\left(\theta_{HO}^N\right) - V\left(\underline{\theta}_H\right)\right]\rho_O^N(\eta)\,/\widehat{R}(\eta), \\
\xi_{HV}^N &= \left[V\left(\theta_{HV}^N\right) - V\left(\theta_{HO}^N\right)\right]\rho_V^N(\eta)\,/\widehat{R}(\eta), \\
\xi_{HO}^S &= \left[V\left(\theta_{HO}^S\right) - V\left(\theta_{HV}^N\right)\right]\rho_O^S(\eta)\,/\widehat{R}(\eta) \\
\xi_{HV}^S &= \left[1 - V\left(\theta_{HO}^S\right)\right]\rho_V^S(\eta)\,/\widehat{R}(\eta).
\end{aligned}
\right\} \quad\quad (4.22)
$$

As is clear from equations (4.21) and (4.22), each market share is now a function of all four cutoffs $\underline{\theta}_H$, $\theta_{HO}^N$, $\theta_{HV}^N$, and $\theta_{HO}^S$. This complicates the analysis relative to the main text, but the results are similar.

First, a fall in the Southern wage or in transport costs increases $\psi_O^S(\eta)$, $\psi_V^S(\eta)$, $\rho_O^S(\eta)$, and $\rho_V^S(\eta)$, while leaving the ratios $\psi_O^S(\eta)\,/\psi_V^S(\eta)$ and $\rho_O^S(\eta)\,/\rho_V^S(\eta)$ unaffected. It is straightforward to check that, as in the main text, the ratios $\xi_{HO}^S/\xi_{HV}^S$, $\xi_{HV}^S/\xi_{HO}^N$, and $\xi_{HO}^N/\xi_{HV}^N$ all increase. It follows that global production sharing, as measured by the sum $\xi_{HO}^S + \xi_{HV}^S$, increases, as does outsourcing relative to integration in each one of the countries. This implies that $\xi_{HO}^S$ rises and $\xi_{HV}^N$ falls. The overall effect on $\xi_{HO}^N$ and $\xi_{HV}^S$ depends on whether $\widehat{R}(\eta)$ increases or decreases. If $\eta > \beta$ and $w^N/w^S$ is high enough, it can be shown that not only $\psi_V^S(\eta) > \psi_O^S(\eta) > \psi_V^N(\eta) > \psi_O^N(\eta)$, but also

---

[33]The inequality $\eta < \beta$ holds true in the low-tech sector when $\beta < 1/2$. This follows from $\beta^*(\eta) > \eta$ if and only if $\beta^*(\eta) < 1/2$ (see equation (4.10)).

$\rho_V^S(\eta) > \rho_O^S(\eta) > \rho_V^N(\eta) > \rho_O^N(\eta)$.[34] In this case $\widehat{R}(\eta)$ rises when Southern wages or transport costs fall. As a result, $\xi_{HO}^N$ falls, whereas the effect on $\xi_{HV}^S$ remains ambiguous. If instead $\eta$, $\beta$, and $w^N/w^S$ are such that $\widehat{R}(\eta)$ falls, then both $\xi_{HO}^N$ and $\xi_{HV}^S$ rise when Southern wages or transport costs decline.

Second, consider the effect of $\delta^\ell$. An increase in $\delta^S$ raises $\psi_V^S(\eta)$ without affecting the slopes of the other profits functions. Furthermore, if $\eta$ is high enough, namely $\eta > \beta_V^S$, this also increases $\rho_V^S(\eta)$ relative to $\rho_O^S(\eta)$, $\rho_V^N(\eta)$, and $\rho_O^N(\eta)$. In this case $\xi_{HV}^S$ increases and $\xi_{HO}^S$ declines, while the ratio $\xi_{HO}^N/\xi_{HV}^N$ does not change. The only difference with the main text is that the market share of final-good producers who use imported components now depends on $\delta^S$. The effect depends again on whether $\widehat{R}(\eta)$ increases or decreases with $\delta^S$. As before, if $\eta > \beta$ and $w^N/w^S$ is high enough, then $\rho_V^S(\eta) > \rho_O^S(\eta) > \rho_V^N(\eta) > \rho_O^N(\eta)$, and $\widehat{R}(\eta)$ is raised by an increase in $\delta^S$. In this case the market share of importers is increasing in $\delta^S$.

An increase in $\delta^N$ affects prevalence similarly to the the main text when $\eta > \beta_V^N$. In this case domestic integration gains market share relative to both domestic outsourcing and foreign outsourcing. As a result, the prevalence of vertical integration relative to outsourcing rises in both countries. As in the main text, $\xi_{HV}^N$ is increasing in $\delta^N$, whereas the effect on the other market shares depends on whether $\widehat{R}(\eta)$ is increasing or decreases in $\delta^N$.

Third, as in the main text, an increase in the primitive bargaining power $\beta$ reduces the ratios $\psi_V^S(\eta)/\psi_O^\ell(\eta)$, $\psi_V^N(\eta)/\psi_O^\ell(\eta)$, and $\psi_V^N(\eta)/\psi_V^S(\eta)$ for $\ell = N, S$. Moreover, it also reduces the ratios $\rho_V^S(\eta)/\rho_O^\ell(\eta)$, $\rho_V^N(\eta)/\rho_O^\ell(\eta)$, and $\rho_V^N(\eta)/\rho_V^S(\eta)$ for $\ell = N, S$. As a result, the market share of firms outsourcing in each country increases relative to the market share of firms integrating in the same country, just as in the main text. The effect on the market share of firms that import components ($\xi_{HO}^S + \xi_{HV}^S$) is, however, ambiguous.

---

[34]In particular, $\eta > \beta$ ensures that $\rho_V^S(\eta) > \rho_O^S(\eta)$ and $\rho_V^N(\eta) > \rho_O^N(\eta)$, while $\rho_O^S(\eta) > \rho_V^N(\eta)$ holds true as long as:

$$\left(\frac{w^N}{w^S}\right)^{1-\eta} > \left(\frac{\beta_V^N}{\beta}\right)^\eta \left(\frac{1-\beta_V^N}{1-\beta}\right)^{1-\eta}.$$

Fourth, it is straightforward to show that an increase in the degree of dispersion reduces the market share of firms outsourcing in the North and increases the market share of firms integrating in the South. Furthermore, as in the main text, $\xi_{HO}^S + \xi_{HV}^S$, $\xi_{HV}^S/\xi_{HO}^S$, and $\xi_{HV}^N/\xi_{HO}^N$ are decreasing in $k$.

Finally, an increase in the output elasticity of headquarter services, $\eta$, increases $\psi_O^N(\eta)/\psi_O^S(\eta)$ and $\psi_V^\ell(\eta)/\psi_O^\ell(\eta)$ for $\ell = N, S$, as well as $\rho_O^N(\eta)/\rho_O^S(\eta)$ and $\rho_V^\ell(\eta)/\rho_O^\ell(\eta)$ for $\ell = N, S$. As in the main text, the relative prevalence of domestic integration increases, both in absolute terms and relative to domestic outsourcing, while the relative prevalence of foreign outsourcing falls, both in absolute terms and relative to foreign integration. Furthermore, under mild assumptions, the market share of firms that import components falls.

Using output of each organizational form as a measure of relative prevalence also yields similar results. In particular, it can be shown that equations (4.19)-(4.22) apply to this case, with $\widetilde{\rho}_k^\ell(\eta) = \left[\rho_k^\ell(\eta)\right]^{1/\alpha}$ replacing $\rho_k^\ell(\eta)$. The comparative statics are therefore similar to those for market shares.

# Bibliography

[1] Abraham, Katharine G., and Susan K. Taylor (1996), "Firms' Use of Outside Contractors: Theory and Evidence," *Journal of Labor Economics*, 14:3, pp. 394-424.

[2] Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, pp. 1369-1401.

[3] Aghion, Philippe, Mathias Dewatripont and Patrick Rey (1994), "Renegotiation Design with Unverifiable Information," *Econometrica*, 62:2, pp. 257-282.

[4] Aoki, Masahiko (1990), "Toward an Economic Model of the Japanese Firm," *Journal of Economic Literature*, 28:1, pp. 1-27.

[5] Antràs, Pol (2002a), "Firms, Contracts, and Trade Structure," mimeo Massachusetts Institute of Technology.

[6] Antràs, Pol (2002b), "Incomplete Contracts and the Product Cycle," mimeo Massachusetts Institute of Technology.

[7] Axtell, Robert L. (2001), "Zipf Distribution of US Firm Sizes", *Science*, 293, pp. 1818-1820.

[8] Bamford, James (1994), "Driving America to Tiers," *Financial World*, November 8, pp. 24-27.

[9] Baker, P. George and Thomas N. Hubbard (2002), "Make or Buy in Trucking: Asset Ownership, Job Design, and Information," mimeo Chicago Business School.

[10] Bloom, Martin (1992), *Technological Change in the Korean Electronics Industry*, Development Center Studies, OECD, Paris.

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

[12] Campa, Jose and Linda S. Goldberg (1997), "The Evolving External Orientation of Manufacturing Industries: Evidence from Four Countries," *Federal Reserve Bank of New York Economic Policy Review*, 4, pp. 79-99.

[13] Casson, Mark (1979), *Alternatives to the Multinational Enterprise*, London: Macmillan.

[14] Caves, Richard E. (1996), *Multinational Enterprise and Economic Analysis*, Second Edition, Cambridge University Press.

[15] Che, Yeon-Koo and Donald B. Hausch (1999), "Cooperative Investments and the Value of Contracting," *American Economic Review*, 89:1, pp. 125-147.

[16] Coase, Ronald H. (1937), "The Nature of the Firm," *Economica*, 4:16, pp. 386-405.

[17] Cohen, Wesley M. and Steven Klepper (1992), "The Anatomy of Industry R&D Intensity Distributions," *American Economic Review*, 82:4, pp. 773-799.

[18] Contractor, Farok J. (1984), "Choosing between Direct Investment and Licensing: Theoretical Considerations and Empirical Tests," *Journal of International Business Studies*, 15:3, pp. 167-188.

[19] Cyhn, J. W. (2002), *Technology Transfer and International Production: the Development of the Electronics Industry in Korea*, Edward Elgar, Cheltenham.

[20] Davidson, W. H. and Donald G. McFetridge (1984), "International Technology Transactions and the Theory of the Firm," *Journal of Industrial Economics*, 32:3, pp. 253-264.

[21] Davidson, W. H. and Donald G. McFetridge (1985), "Key Characteristics in the Choice of International Technology Transfer Mode," *Journal of International Business Studies*, 16:2, pp. 5-21.

[22] Dornbusch, Rudiger, Stanley Fischer, and Paul A. Samuelson (1977), "Comparative Advantage, Trade and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, 67:5, pp. 823-839.

[23] Dunning, John H. (1993), *Multinational Enterprises and the Global Economy*, Addison Wesley Longman, Inc.

[24] *The Economist* (1991), "The Ins and Outs of Outing", August 31, pp. 54-56.

[25] Ethier, Wilfred J. (1986), "The Multinational Firm," *Quarterly Journal of Economics*, 101:4, pp. 805-833.

[26] Ethier, Wilfred J. and James R. Markusen (1996), "Multinational Firms, Technology Diffusion and Trade," *Journal of International Economics*, 41:1, pp. 1-28.

[27] Fan, Joseph P. H., and Larry H. P. Lang (2000), "The Measurement of Relatedness: An Application to Corporate Diversification," *Journal of Business*, 73:4, pp. 629-660.

[28] Feenstra, Robert C. (1998), "Integration of Trade and Disintegration of Production in the Global Economy," *Journal of Economic Perspectives*, 12:4, 31-50

[29] Feenstra, Robert C. (1999), "Facts and Fallacies about Foreign Direct Investment," in Martin Feldstein, ed. *International Capital Flows*, University of Chicago Press and NBER, pp. 331-350.

[30] Feenstra, Robert C. and Gordon H. Hanson (1996), "Globalization, Outsourcing, and Wage Inequality," *American Economic Review*, 86:2, pp. 240-245.

[31] Grossman, Sanford J. and Oliver D. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94:4, pp. 691-719.

[32] Grossman Gene M., and Elhanan Helpman (1991a), "Endogenous Product Cycles," *Economic Journal*, 101:4, pp. 1214-1229.

[33] Grossman, Gene M., and Elhanan Helpman (1991b), "Quality Ladders and Product Cycles," *Quarterly Journal of Economics*, 106:2, pp. 557-586.

[34] Grossman, Gene M., and Elhanan Helpman (2002a), "Integration vs. Outsourcing in Industry Equilibrium," *Quarterly Journal of Economics*, 117:1, pp. 85-120.

[35] Grossman, Gene M., and Elhanan Helpman (2002b), "Managerial Incentives and the International Organization of Production," NBER Working Paper No. 9403.

[36] Gruber William, Dileep Mehta, and Raymond Vernon (1967), "The R & D Factor in International Trade and International Investment of United States Industries," *Journal of Political Economy*, 75:1, pp. 20-37.

[37] Gwartney, James D., with Chris Edwards, Robert Lawson, Walter Park, and Veronique De Rugy (2002), *Economic Freedom of the World: 2002 Annual Report*, Fraser Institute.

[38] Hall, Robert E. and Charles I. Jones (1999), "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114:1, pp. 83-116.

[39] Hanson, Gordon H., Raymond J. Mataloni, Jr., and Matthew J. Slaughter (2002), "Vertical Specialization in Multinational Firms," manuscript, University of California, San Diego.

[40] Hart, Oliver (1995), *Firms, Contracts, and Financial Structure*, Oxford: Clarendon Press.

[41] Hart, Oliver and John Moore (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98:6, pp. 1119-1158.

[42] Hart, Oliver D. and John Moore (1999), "Foundations of Incomplete Contracts," *Review of Economic Studies*, 66:1, Special Issue: Contracts, pp. 115-138.

[43] Helpman, Elhanan (1984), "A Simple Theory of International Trade with Multinational Corporations", *Journal of Political Economy*, 92:3, pp. 451-471.

[44] Helpman, Elhanan (1993), "Innovation, Imitation, and Intellectual Property Rights," *Econometrica*, 61:6, pp. 1247-1280.

[45] Helpman, Elhanan and Paul R. Krugman (1985), *Market Structure and Foreign Trade*, Cambridge: MIT Press.

[46] Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple (2003), "Exports versus FDI," NBER Working Paper No. 9439.

[47] Hummels, David, Jun Ishii, and Kei-Mu Yi (2001), "The Nature and Growth of Vertical Specialization in World Trade," *Journal of International Economics*, 54:1, pp. 75-96.

[48] Hirsch, Seev (1967), *Location of Industry and International Competitiveness*, Oxford: Clarendon Press.

[49] Holmström Bengt and Paul Milgrom (1994), "The Firm as an Incentive System,", *American Economic Review*, 84:4, pp. 972-991.

[50] Kogut, Bruce and Udo Zander (1993), "Knowledge of the Firm and the Evolutionary Theory of the Multinational Corporation," *Journal of International Business Studies*, 24:4, pp. 625-645.

[51] Knack, S., and P. Keefer (1995), "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures," *Economics and Politics* 7:3, pp. 207-27.

[52] Krugman, Paul (1979), "A Model of Innovation, Technology Transfer, and the World Distribution of Income," *Journal of Political Economy*, 87:2, pp. 253-266.

[53] Lai, Edwin L. -C. (1998), "International Intellectual Property Rights Protection and Rate of Product Innovation," *Journal of Development Economics*, 55:1, pp. 133-153.

[54] McLaren, John (2000), "Globalization and Vertical Structure," *American Economic Review*, 90:5, pp. 1239-1254.

[55] Mansfield, Edwin, and Anthony Romeo (1980), "Technology Transfer to Overseas Subsidiaries by U.S.-Based Firms," *Quarterly Journal of Economics*, 95:4, pp. 737-750.

[56] Mansfield, Edwin, Anthony Romeo, and Samuel Wagner (1979), "Foreign Trade and U. S. Research and Development," *Review of Economics and Statistics*, 61:1, pp. 49-57.

[57] Marin, Dalia and Thierry Verdier (2001), "Power Inside the Firm and the Market: A General Equilibrium Approach," mimeo University of Munich.

[58] Marin, Dalia and Thierry Verdier (2003), "Globalization and the Empowerment of Talent," manuscript, University of Munich.

[59] Markusen, James R. (1984), "Multinationals, Multi-Plant Economies, and the Gains from Trade" *Journal of International Economics*, 16, pp. 205-226.

[60] Markusen, James R. (1995), "The Boundaries of Multinational Enterprises and the Theory of International Trade," *Journal of Economic Perspectives*, 9:2, pp. 169-189.

[61] Markusen, James R. and Keith E. Maskus (2001), "General-Equilibrium Approaches to the Multinational Firm: A Review of Theory and Evidence," forthcoming in Harrigan, James, ed. *Handbook of International Trade*, Basil Blackwell.

[62] Mauro, Paolo (1995), "Corruption and Growth," *Quarterly Journal of Economics*, 110:3, pp. 681-712.

[63] Melitz, Marc J. (2002), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," NBER Working Paper No. 8881, forthcoming in *Econometrica*.

[64] Milgrom, Paul, and John Roberts (1993), "Johnson Controls, Inc., Automotive Systems Group: The Georgetown, Kentucky Plant," Stanford Graduate School of Business Case S-BE-9.

[65] Moran, Theodore H. (2001), *Parental Supervision: The New Paradigm for Foreign Direct Investment and Development*, Washington DC: Institute for International Economics.

[66] Mundell, Robert A. (1957), "International Trade and Factor Mobility," *American Economic Review*, 47:3, pp. 321-335.

[67] Nöldeke, Georg and Klaus M. Schmidt (1995), "Option Contracts and Renegotiation: A Solution to the Hold-up Problem," *RAND Journal of Economics*, 26:2, pp. 163-179.

[68] Parry, Thomas G. (1975), "The Product Cycle and International Production: U.K. Pharmaceuticals," *Journal of Industrial Economics*, 24:1, pp. 21-28.

[69] Puga, Diego, and Daniel Trefler (2002), "Knowledge Creation and Control in Organizations," NBER Working Paper No. 9121.

[70] Romalis, John (2002), "Factor Proportions and the Structure of Commodity Trade," mimeo Chicago Graduate School of Business.

[71] Rugman, Alan M. (1981), *Inside the Multinationals: The Economics of Internal Markets*, New York: Columbia University Press.

[72] Segal, Ilya (1999), "Complexity and Renegotiation: A Foundation for Incomplete Contracts," *Review of Economic Studies*, 66:1, Special Issue: Contracts, pp. 57-82.

[73] Tempest, Rone (1996), "Barbie and the World Economy," *Los Angeles Times*, September 22, A1 and A12.

[74] Tirole, Jean (1988), *The Theory of Industrial Organization*, Cambridge: MIT Press.

[75] UNCTAD (1995), *World Investment Report: Transnational Corporations and Competitiveness*, New York: United Nations.

[76] UNCTAD (2001), *World Investment Report: Promoting Linkages*, New York: United Nations.

[77] Vernon, Raymond (1966), "International Investment and International Trade in the Product Cycle," *Quarterly Journal of Economics*, 80:2, pp. 190-207.

[78] Vernon, Raymond (1979), "The Product Cycle Hypothesis in a New International Environment," *Oxford Bulletin of Economics and Statistics*, 41, pp. 255-267.

[79] Vernon, Raymond and W. H. Davidson (1979), "Foreign Production of Technology-Intensive Products by U.S.-Based Multinational Enterprises," Working Paper No. 79-5, Harvard Graduate School of Business Administration.

[80] Wells, Louis T. Jr. (1969), "Test of a Product Cycle Model of International Trade: U.S. Exports of Consumer Durables," *Quarterly Journal of Economics*, 83:1, pp. 152-162.

[81] Whinston, Michael D. (2002), "On the Transaction Cost Determinants of Vertical Integration," mimeo Northwestern University.

[82] Williamson, Oliver E. (1985), *The Economic Institutions of Capitalism*, Free Press.

[83] Wilson, Robert W. (1977), "The Effect of Technological Environment and Product Rivalry on R&D Effort and Licensing of Inventions," *Review of Economics and Statistics*, 59:2, pp. 171-178.

[84] *World Competitiveness Report* (1992), World Economic Forum and International Institute for Management Development.

[85] World Trade Organization (1998), *Annual Report 1998* (Geneva: World Trade Organization).

[86] World Trade Organization (2001), *International Trade Statistics*, Geneva.

[87] Yeats, Alexander J. (2001), "Just How Big Is Global Production Sharing?" in Arndt, Sven W. and Henryk Kierzkowski, eds., *Fragmentation: New Production Patterns in the World Economy*, Oxford University Press.

[88] Young S., N. Hood and J. Hamill (1985), "Decision-Making in Foreign Owned Multi-national Subsidiaries in the U.K., *ILO Working Paper* No. 35. Geneva: International Labor Office.

[89] Zeile, William J. (1997), "U.S Intrafirm Trade in Goods," *Survey of Current Business*, 77:2, pp 23-38.

[90] Zhang, Kevin H. and James R. Markusen (2001), "Vertical Multinationals and Host-Country Characteristics," *Journal of Development Economics*, 59:2, pp. 233-252.