Incomplete Contracts and the Product Cycle

Pol Antràs*
Harvard University, NBER and CEPR

This Draft: June 2004

Abstract

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, the incompleteness of international contracts is shown to lead to the emergence of product cycles. Because of contractual frictions, goods are initially manufactured in the North, where product development takes place. As the good matures and becomes more standardized, the manufacturing stage of production 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. The model gives rise to a new version of the product cycle 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), as well as microeconomic implications consistent with the findings of the empirical literature on the product cycle.

Keywords Product Cycle, Property-rights Theory, Multinational Firms.

JEL Classification Numbers D23, F12, F14, F21, F23, L22, L33

---

*I am grateful to Daron Acemoglu, Marios Angeletos, Gene Grossman, and Jaume Ventura for invaluable guidance, and to Richard Baldwin, Lucia Breierova, Francesco Franco, Gordon Hanson, Elhanan Helpman, Simon Johnson, Giovanni Maggi, Marc Melitz, and Roberto Rigobon for their helpful comments and suggestions. The paper was substantially improved by the thoughful comments of the Editor and two anonymous referees. I have also benefited from suggestions by seminar participants at various institutions. The first draft of this paper was written while visiting the International Economics Section at Princeton University, whose hospitality is gratefully aknowledged. I have also benefited from financial support from the Bank of Spain. All remaining errors are my own. Correspondence: Department of Economics, Harvard University, Littauer 230, Cambridge, MA 02138. Email: pantras@fas.harvard.edu.

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 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 emerging from this literature turned out to be much richer than Vernon originally envisioned. 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 transfer to the South. 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 by 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, 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 product cycles on trade flows and relative wages, Krugman (1979) specified a very simple form of technological transfer, with new goods becoming old goods at an exogenous 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

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

[2] See, for instance, Davidson and McFetridge (1984, 1985), Mansfield et al. (1979), Mansfield and Romeo (1980), and Wilson (1977). These studies will be discussed in more detail in section 4 below.

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.[3] 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.[4]

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, from that in Grossman and Helpman (1991a,b) in that 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, as well as their 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's 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

[3] See Jensen and Thursby (1987), and Segerstrom et al. (1990) for related theories of endogenous product cycles.

[4] Grossman and Helpman (1991b) claimed that purposeful imitation 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 that even in the case of the electronics industry, the spectacular increase in the market share of Korean producers might be better explained by technology transfer from foreign-based firms than by simple imitation by domestic firms in Korea.

plants in the South.[5] With no frictions to the international fragmentation of the production process, I show that the model fails to deliver a product cycle. Intuitively, 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 hold-up problems, which in turn give rise to suboptimal relationship-specific investments by the parties involved in an international 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 and overseas manufacturing, the product development manager therefore faces a trade-off between the lower costs of Southern manufacturing and the higher incomplete-contracting distortions associated with it. 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 (which affects both the manufacturing *and* the product development stages of production) 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 if the Southern wage is low enough, the good is manufactured in the South.

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 in the model, i.e., incomplete contracts, opens the door to a parallel analysis of the determinants of ownership structure, which I carry out in section 3. As in Grossman and Hart (1986), I associate

[5] There is a recent literature in international trade documenting an increasing international disintegration of the production process (cf, Feenstra, 1998, Yi, 2003). A variety of terms have been used to refer to this phenomenon: "international outsourcing", "slicing of the value chain", "vertical specialization", "global production sharing", and many others. Feenstra (1998) discusses the widely cited example of Nike, which subcontracts most parts of its production process to independent manufacturing plants in Asia.

ownership with the entitlement of some residual rights of control. When parties undertake noncontractible, relationship-specific investments, the allocation of these 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) is shown to dictate that residual rights be controlled by the party whose investment contributes most to the value of the relationship. In terms of the model, the attractiveness for a Northern product-development manager of integrating the transfer of production to the South is shown to be increasing in the output elasticity of product development, and thus decreasing in the maturity of the good at the time of the transfer.

As a result, a new version of the product cycle emerges. If the maturity at which manufacturing is shifted to the South is low enough, production will be transferred internally to a wholly-owned foreign affiliate in the South, and the Northern firm will become a multinational firm. In such case, only at a later stage in the product's life cycle will the product development manager find it optimal to give away the residual rights of control, and assign assembly to an independent subcontractor in the South, an arrangement which is analogous to the Northern firm licensing its technology (hi-tech input). 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. In section 4, I discuss several cross-sectional and time-series implications of the model and relate them to the empirical literature on the product cycle. For instance, the model is shown to be useful for understanding the evolution of the Korean electronics industry after the Korean War.

The model developed in sections 2 and 3 focuses first on the profit-maximizing choice of location by a single Northern product development manager. In section 5, I 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. I show that as long as contracts governing international transactions are incomplete, the equilibrium wage in the North necessarily exceeds that in the South, hence justifying the trade-off analyzed in sections 2 and 3. Furthermore, in spite of the 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 incompleteness of contracts. In Section 5, I also discuss some macroeconomic implications of the model that complement the work of Krugman (1979).

The rest of the paper is structured as follows. Section 2 develops a simple dynamic model that shows how the presence of incomplete contracts gives rise to product cycles. In section 3, I allow for intrafirm production transfers and describe the richer product life-cycle that emerges from it. Section 4 reviews the findings of the empirical literature on the product cycle and relates them to the predictions of the model. In section 5, I embed the framework of sections 2 and 3 in a general-equilibrium model of North-South trade and study the effects of incomplete contracting on relative wages and the speed of production transfer. Section 6 offers some concluding comments.

2 Incomplete Contracts and the Life Cycle of a Product

This section develops a simple model in which a product development manager decides 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 simple dynamic extension in which the good gets standardized over time.

2.1 Set-up

Consider a world with 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:

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

where p is the price of the good and λ is a parameter that the producer takes as given.[6]

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

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

capture the mere manufacturing or assembly of the good. 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:

$$y = \zeta_z x_h^{1-z} x_l^z, \qquad 0 \leq z \leq 1, \tag{2}$$

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 one 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. For simplicity, 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 in both countries. Production of any type of bad-quality input can be undertaken at a positive but negligible cost. All types of inputs are assumed to be freely tradable.

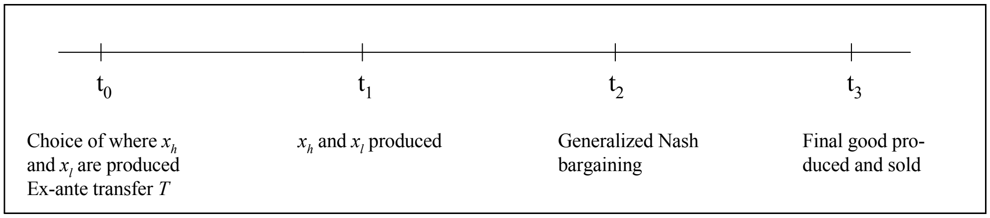
There are two types of producers: a research center and a manufacturing plant. A research center is defined as the producer of the hi-tech input and will thus always locate in the North. The research center needs to contract with an *independent* manufacturing plant for the provision of the low-tech input.[7] As discussed in the introduction, I allow for an international fragmentation of the production process. Before any investment is made, a research center decides whether to produce a hi-tech input, and if so, whether to obtain the low-tech input from an independent manufacturing plant in the North or from one in the South. Upon entry, the manufacturer makes a lump-sum transfer T to the research center. Because, ex-ante, there is a large number of identical, potential manufacturers of the good, competition among them will make T adjust so as to make the chosen manufacturer break even.[8] The research center chooses the location of manufacturing to maximize its ex-ante profits, which include the transfer.

Investments are assumed to be relationship-specific. The research center tailors the hi-tech input specifically to the manufacturing plant, while the low-tech input is customized according to the specific needs of the research center. In sum, the investments in labor

[7] In section 3.1, I allow the research center to obtain the low-tech input from an integrated plant.

[8] When y is produced by the manufacturing plant, the transfer T can be interpreted as a lump-sum licensing fee for the use of the hi-tech input. The presence of this transfer simplifies the description of the industry equilibrium in section 5. For the results in the present section, it would suffice to assume that no firm is cash-constrained, so that the equilibrium location of manufacturing maximizes the joint value of the relationship.

Figure 1: Timing of Events



needed to produce x_h and x_l are incurred upon entry and are useless outside the relationship.

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.[9] Hence, the manager of the research center and that of a Southern manufacturing 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 negligible 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.[10]

When the research center chooses to transact with a manufacturing 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 the low-tech input is produced 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 is observable to both parties and thus the costless bargaining will yield an ex-post efficient outcome. I model this ex-post

[9] This can be interpreted as a physical constraint imposed on the outside party, which might not be able to verify the quality of both inputs when these are produced in distant locations More generally, the 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.

[10] I take the fact that contracts are incomplete as given. Aghion et al. (1994), Nöldeke and Schmidt (1995) and others, have shown that allowing for specific-performance contracts may lead to efficient ex-ante relationship-specific investments. Nevertheless, Che and Hausch (1997) have identified conditions under which specific-performance contracts do not lead to first-best investment levels and may actually have no value.

bargaining as a Symmetric Nash Bargaining game in which the parties share equally the ex-post gains from trade.[11] Because 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.

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

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. The decision of where to produce the low-tech input is instead nontrivial. In his choice, the manager of the research center compares the ex-ante profits associated with two options, which I analyze in turn.

A. Manufacturing by an Independent Plant in the North

Consider first the case of a research center that decides to deal with an independent manufacturing plant in the North. In that 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 (1) and (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 the following optimal price for the final good:

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

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 in turn equal to

$$\pi^N(z) = (1-\alpha)\lambda\left(\frac{w^N}{\alpha}\right)^{-\alpha/(1-\alpha)}. \quad (3)$$

B. Manufacturing by an Independent Plant in the South

Consider next the problem faced by a research center that decides to transact with a plant in the South. As discussed above, in this case the initial contract only stipulates the transfer T. The game played by the manager of the research center and that of the manufacturing

[11] In Antràs (2003b), I extend the analysis to the case of Generalized Nash Bargaining.

plant is solved by backwards induction. 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 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 Nash bargaining leaves each manager with one-half of these quasi-rents. Rolling back in time, the research center manager sets x_h to maximize $\frac{1}{2}R - w^N x_h$, while the manufacturing plant simultaneously chooses x_l to maximize $\frac{1}{2}R - w^S x_l$.[12] Combining the first-order conditions of these two programs yields the following optimal price for the final good:

$$p^S(z) = \frac{2\left(w^N\right)^{1-z}\left(w^S\right)^z}{\alpha}.$$

If parties could write complete contracts in international transactions, the research center would instead set a price equal to $\left(w^N\right)^{1-z}\left(w^S\right)^z/\alpha$. The overinflated price reflects the distortions arising from incomplete contracting. Intuitively, because in the ex-post bargaining the parties fail to capture the full marginal return to their investments, they will tend to underinvest in x_h and x_l. As a result, output will tend to be suboptimal and the move along the demand function will also be reflected in an inefficiently high price.

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

$$\pi^S(z) = \left(1 - \frac{1}{2}\alpha\right)\lambda\left(\frac{2\left(w^N\right)^{1-z}\left(w^S\right)^z}{\alpha}\right)^{-\alpha/(1-\alpha)}. \tag{4}$$

2.3 The Equilibrium Choice

From comparison of equations (3) and (4), it follows that the low-tech input will be produced in the South only if $A(z) \leq \omega \equiv w^N/w^S$, where

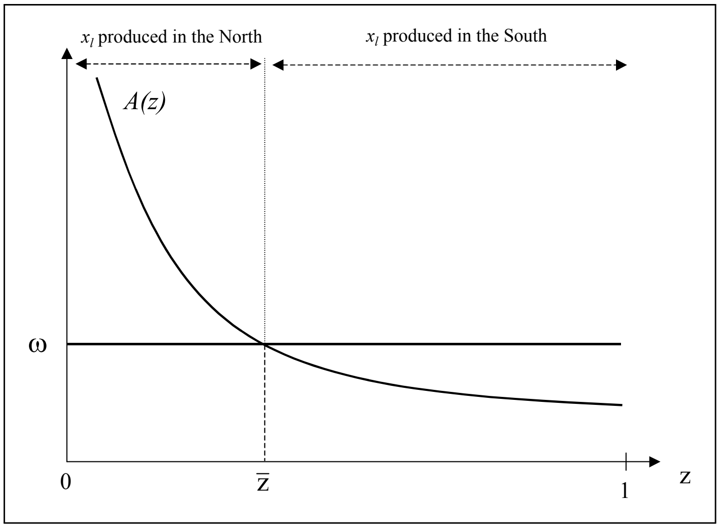
$$A(z) \equiv \left(\frac{1-\alpha}{\left(1-\frac{1}{2}\alpha\right)\left(\frac{1}{2}\right)^{\alpha/(1-\alpha)}}\right)^{(1-\alpha)/\alpha z}. \tag{5}$$

It is straightforward to show that $A(z)$ is non-increasing in z for $z \in [0,1]$, with $\lim_{z\to 0} A(z) = +\infty$ and $A(1) > 1$.[13] This implies that (i) for high enough product-development intensities of the final good, manufacturing is necessarily assigned to a manufacturing plant in the North; and (ii) unless the wage in the North is higher than that in the South, manufacturing by an

[12] It is easily checked that in equilibrium both parties receive a strictly positive ex-post payoff from producing a good-quality input. It follows that bad-quality inputs are never produced.

[13] This follows from the fact that $(1-\alpha x)\,x^{\alpha/(1-\alpha)}$ is increasing in x for $\alpha \in (0,1)$ and $x \in (0,1)$.

Figure 2: The Choice of Location



independent plant in the South will never be chosen. Intuitively, the benefits of Southern assembly are able to offset the distortions created by incomplete contracting only when the manufacturing stage is sufficiently important in production or when the wage in the South is sufficiently lower than that in the North. To make matters interesting, I assume that:

Condition 1: There exists a $z_c \in (0,1)$ such that $A(z_c) < \omega$.[14]

This ensures that $\pi^N(z_c) < \pi^S(z_c)$ for some $z_c \in (0,1)$. Figure 2 depicts the profit-maximizing choice of location as a function of z. It is apparent that:

Lemma 1 *Under 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 (5) and ω is the relative wage in the North.*

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

[14] This condition will in fact be shown to necessarily hold in the general-equilibrium model (which is why I avoid labelling it as an assumption), where the relative wage in the North will necessarily adjust to ensure positive labor demand in the South.

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.[15] Vernon believed that the unstandardized nature of new goods was crucial for understanding that they would be first produced in a high-wage country.

To capture this standardization process in a simple way, consider the following simple dynamic extension of the static model developed above. Time is continuous, indexed 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 (1). The relative wage ω is assumed to be time-invariant.[16] The output elasticity of the low-tech input is instead assumed to increase through time. In particular, this elasticity is given by

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

I therefore assume that the product-development intensity of the good is inversely related to its maturity. Following the discussion in the introduction, this is meant to capture the idea that most goods require a lot of R&D and product development in the early stages of their life cycle, while 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 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.[17] Finally, I assume that the structure of firms 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 future profit streams are sufficiently discounted.

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 1:

[15] In discussing 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 1930s, or the variegated radio designs of the 1920s with the uniform models of the 1930s." (Vernon, 1966, p.197).

[16] The latter assumption will be relaxed in the section 5, where ω will be endogenized.

[17] For instance, if such costs were increasing in dz/dt, then a discrete increase in z would be infinitely costly. A full fledged modeling of the standardization decision is left for future research.

Proposition 1 *The model displays a product cycle. When the good is relatively new or unstandardized, i.e., $t \leq h^{-1}(\bar{z})$, the manufacturing stage of production takes place in the North. When the good is relatively mature or standardized, i.e., $t > h^{-1}(\bar{z})$, manufacturing is undertaken in 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, the higher is the speed of standardization, $1/\theta$, and the lower is the threshold intensity $\bar{z}$. Furthermore, because the cut-off $\bar{z}$ is itself a decreasing function of ω, it follows that the higher is the relative wage in the North, the earlier will production transfer occur.[18]

As argued in the introduction, the fact that international contracts are not perfectly enforceable is important for product cycles to emerge. To illustrate this, consider the case in which the quality of intermediate inputs were verifiable by an outside court even in international transactions, so that the manager of the research center and that of the Southern manufacturing plant could also write enforceable contracts. It is straightforward to check that, in such case, profits for the research center would be $\pi^S(z) = (1-\alpha)\lambda \left(\left(w^N\right)^{1-z} \left(w^S\right)^z / \alpha \right)^{-\alpha/(1-\alpha)}$. Comparing this expression with equation (3), it follows that labor demand in the South would be positive if and only if $\omega \geq 1$ (this is the analog of Condition 1 above). If $\omega > 1$, profits would satisfy $\pi^N(z) \leq \pi^S(z)$ for all $z \in [0,1]$, with strict inequality for $z > 0$. The production process would therefore be broken up from time 0 and no product cycles would arise. If instead $\omega = 1$, $\pi^N(z)$ and $\pi^S(z)$ would be identical for all $z \in [0,1]$ and the location of manufacturing would be indeterminate, in which case product cycles would emerge with probability zero.

Arguably, incomplete contracting is just one of several potential frictions that would make manufacturing stay in the North for a period of time. It is important to emphasize, however,

[18] 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 ω, I have held the contracting environment constant. Recent empirical studies suggest that countries with better legal systems tend to have higher levels of per-capita income (Hall and Jones, 1999, Acemoglu et al., 2001). If I allowed for this type of correlation in the model, production might not be transferred earlier, the higher ω.

that not any type of friction would give rise to product cycles in the model. The fact that incomplete contracts distort both the manufacturing stage *and* the product development stage in production is of crucial importance. For instance, introducing a transport cost or a communication cost that created inefficiencies only in the provision of the low-tech input would not suffice to give rise to product cycles in the model. In this paper, I choose to emphasize the role of incomplete contracts because they are an important source of frictions in the real world and, also, because they are a very useful theoretical tool for understanding firm boundaries, which are the focus of the next section. The type of organizational cycles unveiled by the empirical literature on the product cycle could not easily be rationalized in theoretical frameworks in which production transfer to low-wage countries was delayed merely by transport costs or communication costs.[19] Instead, they will emerge naturally in the extension below.

3 Firm Boundaries and the Product Cycle

Consider next the same set-up as in the previous section with the following new feature. The research center is now given the option of vertically integrating the manufacturing plant and, in the case of Southern assembly, becoming 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 manufacturing plant). In particular, by integrating the production of the low-tech input, the manager of the manufacturing plant becomes an employee of the research center manager. This implies that if the manufacturing plant manager refuses to trade after the sunk costs have been incurred, the research center manager now has the option of firing the overseas manager and seizing 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.[20]

If there were no costs associated with firing the manufacturing plant manager, there

[19] To illustrate this point, consider the case in which the Northern productivity advantage in product development is bounded and the production process cannot be fragmented across borders (e.g., because of prohibitive transport costs or communication costs). Under these circumstances, the whole production process would shift from the North to the South at some point along the life-cycle of the good, but the model would deliver no predictions for the dynamic organizational structure of firms.

[20] See Antràs (2003a) and Antràs and Helpman (2004) for related set-ups.

would be no surplus to bargain over after production, and the manufacturing plant manager would ex-ante optimally set $x_l = 0$ (which of course would imply $y = 0$). In that case, integration would never be chosen. To make things more interesting, I assume that firing the manufacturing plant manager results in a negative productivity shock that leads to a loss of a fraction $1 - \delta$ of final-good production. Under this assumption, the surplus of the relationship remains positive even under integration.[21] I take the fact that δ is strictly less than one as given, but this assumption could be rationalized in a richer framework.

The rest of this section is structured as follows. I will first revisit the static, partial-equilibrium model developed in section 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.

3.1 Firm Behavior

In section 2.2, I computed ex-ante profits for the research center under two possible modes of organization: (A) manufacturing by an independent plant in the North; and (B) manufacturing by an independent plant in the South. The possibility of vertical integration introduces two additional options: manufacturing by a vertically integrated plant in the North and manufacturing by a vertically integrated plant in the South. 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 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. Manufacturing by a Vertically-Integrated Plant in the South

Consider then the problem faced by a research center and its integrated manufacturing 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 manufacturing 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 (1) will translate into sale revenues of $\delta^{\alpha} R$. The quasi-rents of the

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

relationship are therefore given by $(1-\delta^{\alpha})\,R$. Symmetric Nash bargaining leaves each party with its default option plus one-half of the quasi-rents. The research center therefore sets x_h to maximize $\frac{1}{2}(1+\delta^{\alpha})\,R - w^N x_h$, while the Southern manufacturing plant simultaneously chooses x_l to maximize $\frac{1}{2}(1-\delta^{\alpha})\,R - w^S x_l$. Relative to case (B) in section 2, integration enhances the research center's incentives to invest ($\frac{1}{2}(1+\delta^{\alpha}) > \frac{1}{2}$) but, at the same time, it reduces the manufacturing plant's incentives to invest. Combining the first-order conditions of these two programs yields the following optimal price for the final good:

$$p_M^S(z) = \frac{2\left(w^N\right)^{1-z}\left(w^S\right)^z}{\alpha\left(1+\delta^{\alpha}\right)^{1-z}\left(1-\delta^{\alpha}\right)^z}.$$

Incomplete contracting again distorts the optimal price charged for the final good. Notice, however, that in this case the distortions are higher, the higher is z. Setting T so as to make the integrated manufacturing plant break even leads to the following expression for the research center's ex-ante profits:

$$\pi_M^S(z) = \left(1-\frac{1}{2}\alpha\left(1+\delta^{\alpha}(1-2z)\right)\right)\lambda\left(\frac{2\left(w^N\right)^{1-z}\left(w^S\right)^z}{\alpha\left(1+\delta^{\alpha}\right)^{1-z}\left(1-\delta^{\alpha}\right)^z}\right)^{-\alpha/(1-\alpha)}, \tag{6}$$

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 2.3, where I showed that under Condition 1, there exists a unique $\bar{z} = A^{-1}(\omega)$ such that $\pi^N(z) > \pi^S(z)$ for $z < \bar{z}$, and $\pi^N(z) < \pi^S(z)$ for $z > \bar{z}$. Consider next the choice between Northern assembly and Southern assembly by an integrated firm. Comparing equations (3) and (6), it follows that $\pi_M^S(z) \geq \pi^N(z)$ only if $A_M(z) \leq \omega$, where

$$A_M(z) = \left(\frac{1-\alpha}{1-\frac{1}{2}\alpha\left(1+\delta^{\alpha}(1-2z)\right)}\right)^{(1-\alpha)/\alpha z}\left(\frac{2}{\left(1+\delta^{\alpha}\right)^{1-z}\left(1-\delta^{\alpha}\right)^z}\right)^{1/z}. \tag{7}$$

It is straightforward to show that $\lim_{z\to 0} A_M(z) = +\infty$ and that $A_M(z) > 1$ for all $z \in [0,1]$. As with the comparison involving arm's length production transfers, when the low-tech input is not very important in production, the cost-saving benefit of producing it in the South is outweighed by the costs of incomplete contracting, which distort the marginal cost of production of both the hi-tech and the low-tech inputs.[22] It thus follows from this discussion that, as in section 2, the low-tech input will again be produced in the North whenever z is sufficiently low, that is, whenever the good is sufficiently unstandardized.

Consider next the choice between Southern assembly by an independent firm (or outsourcing) and Southern assembly by an integrated firm (or insourcing). It is straightforward to check that insourcing will dominate outsourcing whenever $A_M(z) < A(z)$, while outsourcing will dominate insourcing whenever $A_M(z) > A(z)$.[23] Furthermore, the following result – analogous to Proposition 1 in Antràs (2003a) – is proved in Appendix A.1.

Lemma 2 *There exists a unique cutoff $\bar{z}_{MS} \in (0,1)$ such that $A_M(\bar{z}_{MS}) = A(\bar{z}_{MS})$. Furthermore, $A_M(z) < A(z)$ for $0 < z < \bar{z}_{MS}$, and $A_M(z) > A(z)$ for $\bar{z}_{MS} < z \leq 1$.*
Proof. *See Appendix A.1.* ■

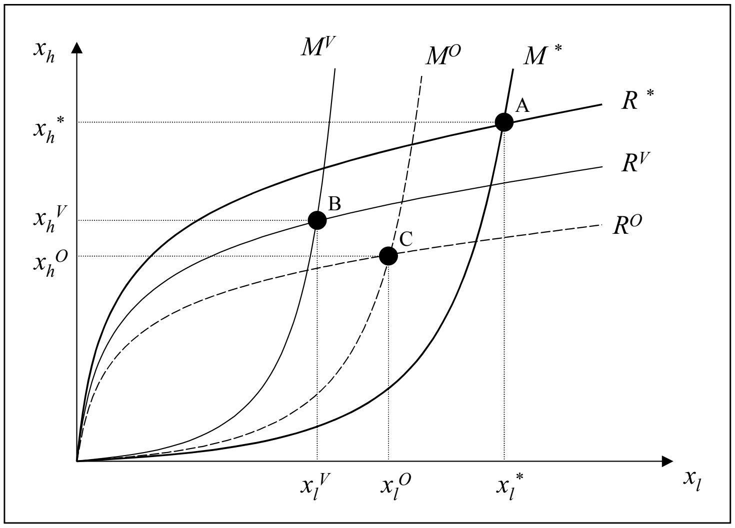
This implies that there exists a unique cutoff $\bar{z}_{MS}$ such that insourcing dominates outsourcing for all $z < \bar{z}_{MS}$, with the converse being true for $z > \bar{z}_{MS}$. The logic of this result lies at the heart of Grossman and Hart's (1986) seminal contribution. When contracts governing transactions are incomplete, ex-ante efficiency dictates that residual rights should be controlled by the party undertaking a relatively more important investment in a relationship. If production of the final good requires mostly product development (i.e., z is low), the investment made by the manufacturing 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 manufacturing plant by giving away these same residual rights.

Figure 3 illustrates this point by depicting the amounts of inputs produced under each organizational mode, as well as those prevailing under complete contracting. The curves M^* and R^* represent the reaction functions $x_h^*(x_l)$ and $x_l^*(x_h)$ under complete contracts, with the corresponding equilibrium at point A. Similarly, B and C depict the incomplete-contract equilibria corresponding to vertical integration and arm's length transacting, respectively. It is clear from the graph that incomplete contracting leads to underproduction of both x_h and

[22] 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. Integration, however, affects the allocation of power in the relationship and this explains why $A_M(z)$ is different from $A(z)$ in equation (5).

[23] This follows directly from $A(z) = \omega \cdot \left(\pi^N(z)/\pi^S(z)\right)^{(1-\alpha)/\alpha z}$ and $A_M(z) = \omega \cdot \left(\pi^N(z)/\pi_M^S(z)\right)^{(1-\alpha)/\alpha z}$.

Figure 3: Underproduction and Ownership Structure



x_l. The crucial point to notice from Figure 3, however, is that because the manufacturing plant has relatively less bargaining power under integration, the underproduction in x_l is relatively higher under integration than under outsourcing. Furthermore, 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 outsourcing will tend to be lower, the more important is the low-tech input in production (i.e., the higher z).

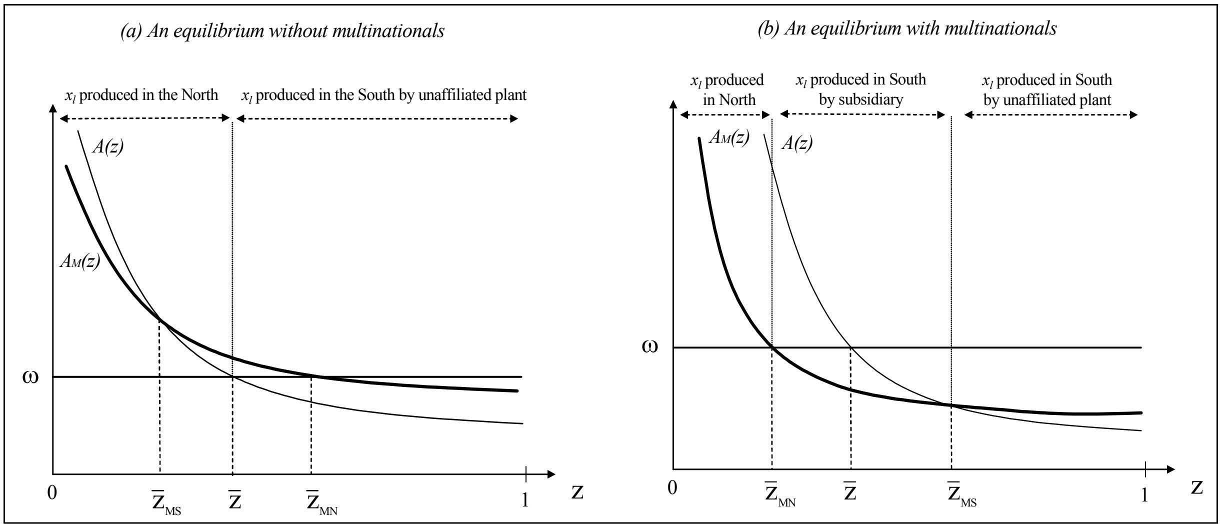
A corollary of Lemmas 1 and 2 is that, as in section 2, when z is sufficiently high (i.e., when $z > \max\{\bar{z}, \bar{z}_{MS}\}$), the low-tech input will again be produced in the South by a nonintegrated manufacturing plant. Remember also that we have established that for sufficiently low z, the low-tech input is necessarily produced in the North. It remains to analyze what happens for intermediate values of z, where multinational firms may potentially arise.

Notice first that if $A_M(z) > \omega$ for all $z \in [0,1]$, Northern assembly strictly dominates Southern insourcing for all $z \in [0,1]$ and multinational firms do not emerge. Furthermore, in such case, the choice between Northern assembly and foreign outsourcing is identical to that in section 2.[24] Let us therefore focus on the case in which $A_M(z) < \omega$ for some $z \in [0,1]$. This analysis is simplified by assuming that δ is not too high, which ensures that the function $A_M(z)$ is a decreasing function of z for *all* $z \in [0,1]$.[25] As shown in Appendix A.2, a sufficient

[24] In particular, because $A(\bar{z}_{MS}) = A_M(\bar{z}_{MS}) > \omega$ and given that $A'(z) < 0$, it must be the case that $\bar{z} > \bar{z}_{MS}$, and thus the equilibrium is as described in Lemma 1.

[25] The $A_M(z)$ curve is decreasing in z for low values of z even when δ approaches one. Assumption 1 rules out cases in which $A_M(z)$ might tilt up for high values of z. Such cases are discussed in Appendix A.2. The

Figure 4: Firm Boundaries and the Product Cycle



condition for this to be the case is:

Assumption 1: $\delta^{\alpha} \leq 1/2$.

Under Assumption 1, there exists a unique cutoff $\bar{z}_{MN} = A_M^{-1}(\omega) \in (0,1)$ such that $\pi^N(z) > \pi_M^S(z)$ for $z < \bar{z}_{MN}$, and $\pi^N(z) < \pi_M^S(z)$ for $z > \bar{z}_{MN}$. This in turn implies that the low-tech input will be produced in the North only if $z < \min\{\bar{z}, \bar{z}_{MN}\}$. Furthermore, it is easily verified that the three thresholds $\bar{z}$, $\bar{z}_{MN}$, and $\bar{z}_{MS}$ must satisfy one of the following: (i) $\bar{z}_{MS} = \bar{z} = \bar{z}_{MN}$, (ii) $\bar{z}_{MS} < \bar{z} < \bar{z}_{MN}$, or (iii) $\bar{z}_{MN} < \bar{z} < \bar{z}_{MS}$.[26]

Figure 4 is instructive in understanding this result. The figure depicts the curves $A(z)$ and $A_M(z)$, which under Assumption 1, are both decreasing in z. Lemma 2 ensure that these curves intersect just once and that $A(z) > A_M(z)$ if and only if $z < \bar{z}_{MS}$. For any relative wage ω, it is clear that either $\bar{z}_{MS} < \bar{z} < \bar{z}_{MN}$ (left panel) or $\bar{z}_{MN} < \bar{z} < \bar{z}_{MS}$ (right panel). The case $\bar{z}_{MS} = \bar{z} = \bar{z}_{MN}$ occurs with probability zero and will be ignored hereafter.

As indicated in both panels in Figure 4, 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 manufac-

results are very similar with the exception that under certain parameter values, the model may feature more complex product-cycle dynamics.

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

turing 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 $\pi_M^S(z) > \max\left\{\pi^N(z), \pi^S(z)\right\}$, 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 3 *If $\bar{z}_{MS} < \min\{\bar{z}, \bar{z}_{MN}\}$, the low-tech input in the North for $z < \bar{z}$, and in the South by an unaffiliated party for $z > \bar{z}$. If instead $\bar{z}_{MS} > \min\{\bar{z}, \bar{z}_{MN}\}$, 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 $\bar{z} > \bar{z}_{MS}$.*

3.2 Dynamics: The Product Cycle

Consider now the dynamics developed in section 2.4 and assume that δ is also time-invariant, implying that not only $\bar{z}$, but also $\bar{z}_{MN}$ and $\bar{z}_{MS}$ are constant through time. The following is a straightforward corollary of Lemma 3:

Proposition 2 *The model displays a product cycle. If $\bar{z}_{MS} < \min\{\bar{z}, \bar{z}_{MN}\}$, the product cycle is as described in Proposition 1. If instead $\bar{z}_{MS} > \min\{\bar{z}, \bar{z}_{MN}\}$, the following product cycle emerges. When the good is relatively new, i.e., $t < h^{-1}(\bar{z}_{MN})$, the manufacturing stage of production takes place in the North. For an intermediate maturity of the good, $h^{-1}(\bar{z}_{MN}) < t < h^{-1}(\bar{z}_{MS})$, manufacturing is shifted to the South but is undertaken within firm boundaries. When the good is relatively standardized, i.e., $t > h^{-1}(\bar{z}_{MS})$, production is shifted to an unaffiliated party in the South.*

This is the central result of this paper. It states that if the threshold maturity level $\min\{\bar{z}, \bar{z}_{MN}\}$ at which manufacturing is shifted to the South is high enough, the transfer of production will occur at arm's length and multinationals will not emerge in equilibrium. Conversely, if this threshold maturity level is low enough, manufacturing will be shifted to the South within the boundaries of the Northern firm by establishing a wholly-owned foreign affiliate. In that case, arm's length assembly in the South will only be observed at a later stage in the life cycle of the good. The model may thus generate both endogenous product cycles as well as endogenous organizational cycles.

4 Empirical Evidence

This section reviews some implications of this extended version of the model and contrasts them with the findings of the empirical literature on the product cycle. For simplicity, I will mostly focus on the case in which $\bar{z}_{MS} > \min\{\bar{z}, \bar{z}_{MN}\}$, so that the model features both intrafirm as well as arm's-length production transfers.

Consider first the time-series implications of the model. These are well summarized by Proposition 2. The model predicts that industries will emerge in low-wage countries only with some lag. Furthermore, the model predicts that in the initial phases of the presence of the industry in the South, foreign direct investment from rich countries should constitute an important part of the industry. Eventually, unaffiliated domestic producers should gain the bulk of the Southern market share, but importantly the model predicts that foreign licensing should still play an important role in those later phases.

The model is consistent with the evolution of the Korean electronics industry from the early 1960s to the late 1980s.[27] 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.[28] These foreign subsidiaries tended to assemble components exclusively for export using imported parts. 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.[29] The percentage approached 100% in the case of exports of computer terminals 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.[30]

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

[28] 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.

[29] 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 producer".

[30] As pointed out by a referee, a caveat in mapping Proposition 2 with the evolution of the Korean electronics industry is that, during this period, Korean wages were growing faster than U.S. wages (i.e., ω was steadily

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. To see this, imagine attempting to test the model with data on a cross-section of production transfers. The model would then predict that the probability of a particular transfer occurring within firm boundaries should be decreasing in the maturity of the product at the time of the transfer. This maturity should in turn be negatively correlated with the age of the product and positively correlated with both its R&D intensity as well as with its speed of standardization.

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. 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.[31]

In 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 indeed higher the newer and more radical was a technology and the larger was the fraction of the transferor's resources devoted to scientific R&D.

There is also some evidence that the probability of internalization might be decreasing in the speed of standardization. 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

fall). Notice, however, that because $\bar{z}_{MS}$ is independent of ω, the model would still predict the simple three-stage product cycle provided that ω does not fall at a rate faster than $A(z)$ and $A_M(z)$, as would be the case if the good standardizes at a sufficiently fast rate.

[31] In the previous case of the Korean electronics industry, 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 CD players to ten Korean producers, while keeping within firm boundaries the assembly of their deck mechanisms.

and teachable and the less complex was the technology.

The dataset used by Davidson and McFetridge (1985) also includes information on the characteristics of the country receiving the transfer. The model predicts that 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 2.3, I showed that $\bar{z}$ is a decreasing function of the relative wage ω. By way of implicit differentiation, and making use of Assumption 1, one can show that $\bar{z}_{MN}$ is also decreasing in ω. 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.[32] It thus follows that in a cross-section of production transfers, the probability of internalization should be decreasing in the labor costs of the recipient country. This prediction is consistent with the findings of Davidson and McFetridge (1985). In their sample of 1,376 transfers, they found that a higher GNP per capita of the recipient country (arguably, a proxy for ω in the model) was associated with a lower probability of internalization. Importantly, their results are robust to controlling for several institutional characteristics of the recipient country (remember the discussion in footnote 18).[33]

One further implication of the model is that relative to the case in which only arm's length transactions are permitted, the emergence of intrafirm production transfer by multinational firms accelerates the shift of production towards the South (remember that $\bar{z}_{MN} < \bar{z}$ whenever multinational firms are active in the model). This result 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 (p. 36).

[32] 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 of technology.

[33] In parallel work using aggregate industry data from the U.S. Department of Commerce, Contractor (1984) found similar results.

5 Incomplete Contracts and the Product Cycle in General Equilibrium

In this section, the partial-equilibrium model developed above is embedded in a 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.

5.1 Set-up

Consider again a world with two countries, the North and the South. The North is endowed with L^N units of labor at any time $t \in (0, \infty)$, 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. I consider an economy in which exogenous inventions continuously increase 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. Preferences of the infinitely-lived representative consumer in each country are given by:

$$U = \int_0^\infty e^{-\rho t} \int_0^{N(t)} \log \left(\int_0^{n_j(t)} y_j\left(i,t\right)^\alpha di \right)^{1/\alpha} dj dt, \tag{8}$$

where ρ 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 (8), 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).[34] As is well known, the instantaneous utility function in (8) gives rise to a constant price-elasticity of demand for any variety i in any industry j:

$$y_j\left(i,t\right) = \lambda_j(t) p_j\left(i,t\right)^{-1/(1-\alpha)}, \tag{9}$$

[34] For simplicity, equation (8) assumes an infinite intertemporal elasticity of substitution in aggregate consumption. Because of the static nature of the consumer's problem, this is an immaterial assumption and the same results would apply for any well-behaved instantaneous utility function.

where

$$\lambda_j(t) = \frac{1}{N(t)} \frac{E\,(t)}{\int_0^{n_j(t)} p_j(i', t)^{-\alpha/(1-\alpha)} di'}, \tag{10}$$

and $E\,(t)$ 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 (1) in the partial-equilibrium model above.

Production of each final-good variety is also as described in sections 2 and 3, 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 (2), with a common time-varying elasticity $z(t - t_{0j}, \theta_j)$, where t_{0j} is the date at which industry j is born and θ_j is an industry-specific parameter that captures differences in the speed of standardization across industries in the same cohort. As before, I assume that $\partial z(\cdot)/\partial\,(t - t_{0j}) > 0$, $z(0, \theta_j) = 0$, and $\lim_{t-t_{0j}\to\infty} z(t - t_{0j}, \theta_j) = 1$. That is, 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. The industry-specific parameter θ_j is assumed to be drawn at period t_{0j} from a time-invariant distribution $G\,(\theta)$. 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 $n_j\,(t)$ of varieties in each industry always adjusts so as to make all 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 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. The contracting environment is also analogous to that of the partial-equilibrium model and, in particular, the parameter δ is time-invariant and common for all varieties and industries. These assumptions, coupled with the absence of means of saving in the model, permit a period-by-period analysis of the dynamic, general-equilibrium model.

5.2 General Equilibrium without Multinational Firms

To better illustrate the workings of the general equilibrium, it is useful to first study the case in which intrafirm production transfers are ruled out. Consider then the equilibrium in any industry j at any period $t \in [0, \infty)$.[35] Facing the same technology and contracting environment, all producers 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 $z(t - t_{0j}, \theta_j) < \bar{z}(t)$, and in the South if $z(t - t_{0j}, \theta_j) > \bar{z}(t)$, with the choice remaining indeterminate for $z(t - t_{0j}, \theta_j) = \bar{z}(t)$. The equilibrium number of varieties produced in industry j at time t can be solved for by using prices to compute $\lambda_j(t)$, and then setting operating profits in (3) and (4) equal to fixed costs, as dictated by free entry. This yields

$$n_j(t) = \begin{cases} (1-\alpha) E(t) / \left[N(t) w^N(t) f \right] & \text{if } z(t - t_{0j}, \theta_j) < \bar{z}(t) \\ \left(1 - \frac{1}{2}\alpha\right) E(t) / \left[N(t) w^N(t) f \right] & \text{if } z(t - t_{0j}, \theta_j) > \bar{z}(t) \end{cases} . \tag{11}$$

Naturally, the equilibrium number of varieties in industry j depends positively on total spending in the industry and negatively on fixed costs.

Free entry ensures that profits are zero and thus all income accrues to labor. In the general equilibrium, world income equals world spending on all goods:

$$w^N(t)L^N + w^S(t)L^S = E(t), \tag{12}$$

and the labor market clears in each country. By Walras' law, we can focus on the equilibrium in the labor market in the South. Southern labor will only be demanded by those manufacturing plants belonging to an industry with $z(t - t_{0j}, \theta_j) > \bar{z}(t)$. It is straightforward to show that labor demand by each manufacturing plant in the South can be expressed as $L_l^S = \frac{1}{2}\alpha z(\cdot) E(t) / \left(w^S(t) N(t) n_j(t) \right)$. Denoting by $F_t(z)$ the fraction of industries with $z(t - t_{0j}, \theta_j) < \bar{z}(t)$ at time t and letting $f_t(z)$ be the associated probability density function, the Southern labor-market clearing condition can be expressed as:

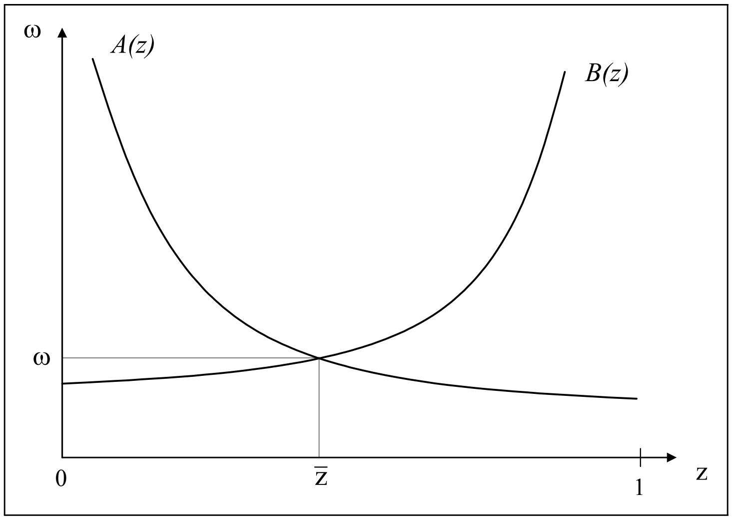
$$\int_{\bar{z}(t)}^{1} \frac{1}{2} \alpha z E(t) f_t(z) dz = w^S(t) L^S. \tag{13}$$

Defining $\xi_t(a, b) \equiv \int_a^b z f_t(z) dz$ and using (12), equation (13) can be rewritten as follows:

$$\omega(t) = B_t(\bar{z}(t)) \equiv \frac{2 - \alpha \xi_t(\bar{z}(t), 1)}{\alpha \xi_t(\bar{z}(t), 1)} \frac{L^S}{L^N}. \tag{14}$$

[35] The unit elasticity of substitution between varieties in different industries implies that we can analyze firm behavior in each industry independently. This assumption, which is made for tractability, comes at the cost of obscuring potentially interesting cross-industry interactions in the production transfer decision.

Figure 5: General Equilibrium



$B_t(\bar{z}(t))$ 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 $\omega(t)$, so an increase in $\omega(t)$ is necessary to bring the Southern labor market 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 $+\infty$.[36] Figure 5 depicts the curve $B_t(\cdot)$ in the (z, ω) space.

The other equilibrium condition that pins down $\bar{z}(t)$ and $\omega(t)$ comes from the partial equilibrium in section 2. In particular, since α is common across industries, $\bar{z}(t)$ is also common across industries and is implicitly defined by the equal profitability condition $\omega(t) = A(\bar{z}(t))$, where $A(\cdot)$ is defined in equation (5). Remember that $A(\bar{z}(t))$ is a decreasing function of z satisfying $\lim_{\bar{z}(t)\to 0} A(\bar{z}(t)) = +\infty$ and $A(1) > 1$. The function $A(\cdot)$ is depicted in Figure 5 together with the function $B_t(\cdot)$. It is apparent from Figure 5 that there exists a unique equilibrium pair $(\bar{z}(t), \omega(t))$ at each period $t \in [0, \infty)$. Furthermore, the fact that $A(1)$ is greater than 1 ensures that the equilibrium wage in the North is higher than that in the South, i.e., $\omega(t) > A(1) > 1$. This implies that Condition 1 in section 2 necessarily holds in the general equilibrium, thus granting validity to the analysis in sections 2 and 3.

It is interesting to notice that in spite of the 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 et al. (1977). Notice,

[36] Since the North always produces the hi-tech input, labor demand in the North is positive even when $\bar{z}(t)$ goes to 0, and consequently $B_t(0)$ is greater than zero.

however, that comparative advantage as represented by the curve $A(\cdot)$ is here endogenous and arises from a combination of the Northern productivity advantage in product development, the continuous standardization of goods, and the fact that contracts are incomplete.

The general equilibrium of the dynamic model is simply the sequence of period-by-period general equilibria. Moreover, as shown on Appendix A.3, the economy will converge to a stationary equilibrium in which the distribution function $F_t(z)$ is time-invariant distribution, and therefore $\bar{z}(t)$ and $\omega(t)$ are also time-invariant. In the equilibrium, all industries will necessarily follow product cycles, with varieties being manufactured first in the North and later in the South. In sum,

Proposition 3 *The economy converges to a stationary equilibrium in which the relative wage in the North is higher than one ($\omega > 1$).*

Proof. *See Appendix A.3.* ■

To illustrate the properties of the general equilibrium, consider again the particular functional form:

$$z(t - t_{0j}, \theta_j) = 1 - e^{-(t-t_{0j})/\theta_j}, \tag{15}$$

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 2.4, and given that the threshold $\bar{z}(t)$ is common across all industries, the model predicts that industries with higher rates of standardization will transfer manufacturing to the South earlier. Furthermore, in the general equilibrium, the cross-industry distribution of standardization rates will have an additional effect on the timing of production transfer, through its impact on the world distribution of product-development intensities, as given by $F_t(z)$. To see this, assume that θ_j is drawn at t_{0j} from an exponential distribution with mean θ_μ, i.e., $G(\theta_j) = 1 - e^{-\theta_j/\theta_\mu}$. Under these assumptions, Appendix A.3 shows that $F_t(z)$ converges to a time-invariant distribution function characterized by:

$$F(z) = \frac{g\theta_\mu \ln\left(\frac{1}{1-z}\right)}{1 + g\theta_\mu \ln\left(\frac{1}{1-z}\right)}. \tag{16}$$

Furthermore, it is easily verified that the steady-state relative wage in the North is increasing in the relative population size of the South L^S/L^N (as in Dornbusch et al., 1977) and in the rate of exogenous invention g, while it is decreasing in the average rate of standardization

$1/\theta_\mu$ (see Appendix A.3 for details). These comparative statics on the relative wage are similar to those obtained by Krugman (1979) in the first paper to explore the macroeconomic implications of the product cycle hypothesis.[37] By endogenizing the timing of production transfer, the present model delivers additional implications of a shift in the parameters g, θ_μ, and L^S/L^N. For instance, an increase in the rate of invention g 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.[38]

In the working paper version of this paper (Antràs, 2003b), I also analyze the implications of a shift from the stationary equilibrium with incomplete contracts described above to a stationary equilibrium with complete contracts. This improvement in the contracting environment in international transactions is shown to unambiguously decrease the relative wage in the North. Intuitively, Northern wages include a premium or rent that partly stems from the incompleteness of international contracts. In a world of complete contracts, this rent disappears and the relative wage in the North is reduced. In Antràs (2003b), I also show that this shift towards complete contracts unambiguously increases the steady-state welfare in the South, while having an ambiguous effect on steady-state Northern welfare.

5.3 General Equilibrium with Multinationals

Consider now the general equilibrium with multinational firms. In order to solve for the equilibrium relative wage, it is useful to define $\widetilde{z}(t) = \min\{\bar{z}_{MN}(t), \bar{z}(t)\}$ which, according to Lemma 3 constitutes the threshold z above which production is shifted to the South. This is the only variable apart from $\omega(t)$ that we need to pin down in the general equilibrium, given that $\bar{z}_{MS}$ does not depend on $\omega(t)$.

Computing labor demand by Southern manufacturing plants and imposing labor market clearing in the South yields the following equilibrium condition relating ω and $\widetilde{z}$, which is analogous to equation (14) (see Antràs, 2003b, for details):

$$\omega = B_{MNE}(\widetilde{z}) \equiv \begin{cases} \frac{2-\alpha(1-\delta^\alpha)\xi(\widetilde{z},\bar{z}_{MS})-\alpha\xi(\bar{z}_{MS},1)}{\alpha(1-\delta^\alpha)\xi(\widetilde{z},\bar{z}_{MS})+\alpha\xi(\bar{z}_{MS},1)}\frac{L^S}{L^N} & \text{if } \widetilde{z} < \bar{z}_{MS} \\ \frac{2-\alpha\xi(\widetilde{z},1)}{\alpha\xi(\widetilde{z},1)}\frac{L^S}{L^N} & \text{if } \widetilde{z} > \bar{z}_{MS} \end{cases}. \quad (17)$$

To save on notation, equation (17) already imposes that the distribution function $F_t(z)$ will converge to a time-invariant distribution, so that we can focus on the stationary general

[37] Krugman (1979) concluded from his comparative static results that increased technological change in the North redistributes income from the South to the North. My analysis suggests that this is not necessarily the case when increased technological change takes the form of an increase in the rate at which goods standardize.

[38] Increases in θ_μ and L^S/L^N have an analogous effect on the timing of production transfer.

equilibrium and safely drop time subscripts. It is apparent from equation (17) that if $\widetilde{z} > \bar{z}_{MS}$, the equilibrium is one without multinationals and the equilibrium condition naturally collapses back to the one in the previous section, i.e., $B_{MNE}(\widetilde{z}) = B(\widetilde{z})$. If instead $\widetilde{z} < \bar{z}_{MS}$, multinationals indeed arise in equilibrium. Furthermore, from the standard logic in Grossman and Hart (1986), an integrated manufacturing plant manager will underinvest relatively more than a non-integrated one. Hence, for a given z, Southern labor demand is relatively lower for vertically-integrated manufacturing plants, implying that $B_{MNE}(\widetilde{z}) > B(\widetilde{z})$ for $\widetilde{z} < \bar{z}_{MS}$. Finally, notice that $B_{MNE}(\widetilde{z})$ is a continuous and increasing function of $\widetilde{z}$, satisfying $B_{MNE}(0) > B(0) > 0$ and $\lim_{\widetilde{z}\to 1} B_{MNE}(\widetilde{z}) = +\infty$.

As in section 5, the other equilibrium condition that pins down $\widetilde{z}$ and ω comes from firm behavior. In particular, because $\bar{z}$ is implicitly defined by $\omega = A(\bar{z})$, and $\bar{z}_{MN}$ is implicitly defined by $\omega = A_M(\bar{z}_{MN})$, it follows that $\widetilde{z} = \min\{\bar{z}_{MN}, \bar{z}\}$ is implicitly defined by:

$$\omega = A_{MNE}(\widetilde{z}) \equiv \begin{cases} A_M(\widetilde{z}) & \text{if } \widetilde{z} < \bar{z}_{MS} \\ A(\widetilde{z}) & \text{if } \widetilde{z} > \bar{z}_{MS} \end{cases} . \tag{18}$$

Again, if $\widetilde{z} > \bar{z}_{MS}$, multinationals are not active in equilibrium and $A_{MNE}(\widetilde{z}) = A(\widetilde{z})$. On the other hand, as discussed in section 3, if $\widetilde{z} < \bar{z}_{MS}$, then it must be the case that $A_{MNE}(\widetilde{z}) = A_M(\widetilde{z}) < A(\widetilde{z})$. Overall, $A_{MNE}(\widetilde{z})$ is a continuous and (under Assumption 1) decreasing function of $\widetilde{z}$ in $[0, 1]$, satisfying $\lim_{\widetilde{z}\to 0} A_{MNE}(\widetilde{z}) = +\infty$ and $A_{MNE}(1) > 1$.

It thus follows that there again exists a unique general-equilibrium $(\widetilde{z}, \omega)$ pair. Depending on parameter values, the equilibrium is one without multinational firms (if the equilibrium $\widetilde{z}$ is lower than $\bar{z}_{MS}$) or one with multinational firms. In either case, industries necessarily follow product cycles, with varieties in those industries first being manufactured in the North and later in the South. In addition to these endogenous product cycles, an equilibrium with multinational firms also features endogenous organizational cycles, with production being shifted to the South first within firm boundaries and only later to independent Southern firms.

Furthermore, given that $A_{MNE}(\widetilde{z}) \leq A(\widetilde{z})$ and $B_{MNE}(\widetilde{z}) \geq A(\widetilde{z})$ for all $\widetilde{z}$, it follows that relative to a world with only arm's length transacting, allowing for intrafirm production 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 ω. This result complements the previous discussion on the effects of more liberal FDI policies on firm behavior, but it illustrates how these policies may also produce significant general equilibrium effects.[39]

Finally, let us briefly return to the particular example analyzed above in which $F_t(z)$ converges to equation (16). As in the equilibrium with only arm's length transacting, an increase in g, θ_μ or L^S/L^N can be shown to increase the relative wage in the North and to reduce the timing of production transfer $\widetilde{z}$. Furthermore, these parameter changes can be shown to (weakly) increase the measure of product-development intensities for which multinational firms exist. Intuitively, by creating an excess supply of Northern labor, an increase in g, θ_μ, and L^S/L^N, raise both the relative wage as well as the threshold product-development intensity $(1-\widetilde{z})$ below which manufacturing is transferred to the South. From the result in Lemma 2, it then becomes more likely that the Northern research center will decide to keep this transfer within firm boundaries.

6 Conclusions

This paper has 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 also delivers a few macroeconomic implications that complement the work of Dornbusch et al. (1977) and Krugman (1979).

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.[40] As in Antràs (2003a), I instead set forth a purely organizational, property-rights model of the multinational firm.[41] Multinational firms emerged in equilibrium whenever transaction-cost minimization dictated that certain goods would be transacted more

[39] In Antràs (2003b), I also analyze the welfare implications of a shift from a stationary equilibrium with only arm's-length transacting to a stationary equilibrium with multinational firms. It is shown that, provided that its effect on relative wages is small enough, the steady-state welfare of both countries is higher in a world with intrafirm production transfers by multinational firms.

[40] This previous literature builds on the seminal work of Helpman (1984) and Markusen (1984), and is extensively reviewed in Caves (1996) and Markusen (1995). Ethier (1986), Ethier and Markusen (1996), and Glass and Saggi (2002) study the choice between foreign direct investment and licensing, but in frameworks in which the internalization decision is unrelated to the allocation of some residual rights of control.

[41] This paper is related to an emerging literature on general-equilibrium models of ownership structure (c.f., McLaren, 2000, Grossman and Helpman, 2002, Antràs, 2003a, Antràs and Helpman, 2004). In Antràs (2003a), 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 model was extended in Antràs and Helpman (2004) to account for intraindustry heterogeneity in organizational choices.

efficiently within firm boundaries than at arm's length. Relative to a world with only arm's length transacting, I showed that foreign direct investment might help alleviate contractual frictions in global production sharing, thereby anticipating the transfer of certain stages of the production to low-wage countries.

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.

A Appendix

A.1 Proof of Lemma 2

Let $\Theta(z) = (A_M(z)/A(z))^z$, which using equations (5) and (7) simplifies to

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

Straightforward algebra delivers that $\Theta'(z) > 0$ if and only if

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

The right hand side is decreasing in z and is therefore no larger than $(1-\alpha)\delta^\alpha / \left(1 - \frac{1}{2}\alpha(1+\delta^\alpha)\right)$, which in turn can be shown to be no larger than $\ln[(1+\delta^\alpha)/(1-\delta^\alpha)]$. To see this last statement, simply define the function $\vartheta(\delta) = \ln[(1+\delta^\alpha)/(1-\delta^\alpha)] - (1-\alpha)\delta^\alpha / \left(1 - \frac{1}{2}\alpha(1+\delta^\alpha)\right)$ and notice it is increasing in δ and satisfies $\vartheta(0) = 0$.

Next, the fact that $(1-\alpha x)x^{\alpha/(1-\alpha)}$ is increasing in x for $\alpha \in (0,1)$ and $x \in (0,1)$ implies that $\Theta(0) < 1$ and $\Theta(1) > 1$. Hence, there exists a unique $\bar{z}_{MS}$ such that $\Theta(z) < 1$ for $0 < z < \bar{z}_{MS}$, $\Theta(z) > 1$ for $\bar{z}_{MS} < z < 1$, and $\Theta(z) = 1$ for $z = \bar{z}_{MS}$. Notice that Assumption 1 is not necessary for this result. To complete the proof, note that for $z \in (0,1)$, $A_M(z) < A(z)$ if and only if $\Theta(z) < 1$; $A_M(z) > A(z)$ if and only if $\Theta(z) > 1$; and $A_M(z) = A(z)$ if and only if $\Theta(z) = 1$. Finally, for $z = 1$, straightforward manipulation yields $A_M(1) > A(1)$.

A.2 Discussion of Assumption 1

As argued in the main text, the proof of Lemma 3 is straightforward when $A_M(z)$ is non-increasing in z for all $z \in [0,1]$. Simple differentiation of equation (7) shows that this is the case whenever

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

It is easy to show that $\partial r(\cdot)/\partial z \geq 0$ for all $z \in [0,1]$ (with strict inequality for $z > 0$), which implies that $A_M(z)$ will be decreasing for all z if the slope at $z = 1$ is negative. Differentiating $r(1,\delta,\alpha)$ with respect to δ, one can show that the slope at 1 is increasing in δ and is negative when evaluated at $\delta = (1/2)^{1/\alpha}$. It follows that if $\delta^\alpha < 1/2$, $A_M(z)$ is non-increasing in z for all $z \in [0,1]$.

When $\delta^\alpha > 1/2$, the $A_M(z)$ curve will still be decreasing in z for sufficiently low z, but it may tilt up when z is sufficiently close to one. In those cases, firm behavior will still be exactly as described in Lemma 3 of section 4 provided that $A_M(z)$ equals ω for only one $z \leq \bar{z} \equiv A^{-1}(\omega)$. In other words, the $A_M(z)$ curve may intersect ω for high values of z, but this is immaterial because arm's length transacting dominates insourcing at those values of z.

Conversely, when $\delta^\alpha > 1/2$ and $A_M(z)$ equals ω for two values of z less than $\bar{z}$, firm behavior is a bit more complex than as described in Lemma 3.[42] Denote these two thresholds by $\bar{z}_{MN}$ and $\bar{z}'_{MN}$. As in Lemma 3, firms will produce the low-tech input in the North for $z < \bar{z}_{MN}$, and they will contract with an arm's-length Southern producer when $z > \bar{z}_{MS}$. The only difference is that multinational firms will emerge only in the interval $(\bar{z}_{MN}, \bar{z}'_{MN}) \subset (\bar{z}_{MN}, \bar{z}_{MS})$, while Northern assembly will be the preferred option not only for $z < \bar{z}_{MN}$ but also for $z \in (\bar{z}'_{MN}, \bar{z}_{MS})$. In the dynamic extension of the model, this implies that the model may predict that assembly returns to the North for some intermediate levels of standardization.

The economics behind these results are as follows. Incomplete contracting has two effects on profits – compare equation (3) with equations (4) or (6). On the one hand, contractual frictions lead

[42] The convexity of $A_M(z)$ in $[0,1]$ ensures that $A_M(z) = \omega$ for at most two values of z.

to underproduction of both x_h and x_l, which translates into lower sale revenues and profits. On the other hand, these frictions create rents, thereby increasing the fraction of revenues that producers are able to capture as operating profits. This second effect is second order in the sense that, holding relative wages constant, operating profits are always higher under complete contracts than under incomplete contracts. This is a desirable property of the model, because the existence of the second effect seems much less robust to alternative modelling strategies.[43] What matters for Lemma 3, however, is how the overall distortions vary with z. This derivative tends to be dominated by the underproduction effect which dictates that $A_M(z)$ be a non-increasing function of z for all $z \in [0,1]$. For sufficiently high δ, however, the second effect may actually dominate thus complicating the analysis.

A.3 Proof of Proposition 3

The fact that $\omega > 1$ follows immediately from $A(z) > 1$ for all $z \in [0,1]$. Furthermore, because $A_M(z) > 1$ for all $z \in [0,1]$, this also holds in the general equilibrium with multinationals.

Let us now concentrate on the stationary properties of the model. Remember that it is sufficient to show that $F_t(z)$ converges to a time invariant distribution. Assuming that the function $z(\cdot)$ is invertible, the fraction of industries with $z(t-t_{0j},\theta_j) < z$ can be expressed as the following probability:

$$F_t(z) = \Pr\left(z(t-t_{0j},\theta_j) < z\right) = \Pr\left(t - t_{0j} < \psi(\theta_j, z)\right).$$

Using the fact that the number of industries grows at constant rate g, the distribution of $t - t_{0j}$ conditional on a particular realization of θ_j is simply

$$F_t(z|\theta_j) = \Pr\left(t - t_{0j} < \psi(\theta_j,z)\,|\theta_j\right) = \begin{cases} 1 - e^{-g\cdot\psi(\theta_j,z)} & \text{if } \psi(\theta_j,z) \le t \\ \varphi_0(\theta_j,z,t) & \text{if } \psi(\theta_j,z) > t \end{cases},$$

where $\varphi_0(\theta_j,z,t)$ depends on the characteristics of the initial N_0 industries. As $t \to \infty$, however, $F_t(z|\theta_j)$ converges to $F(z|\theta_j) = 1 - e^{-g\cdot\psi(\theta_j,z)}$.

Next notice, that by the Glivenko-Cantelli theorem, the distribution of θ_j within each cohort converges uniformly to $G(\theta)$. Because this distribution is time invariant, it follows that the overall distribution for $\psi(\theta_j,z) \le t$ will also be characterized by $G(\theta)$. This in turn implies that as $t \to \infty$, the fraction of industries with $z(t-t_{0j},\theta_j) < z$ will converge to the time-invariant distribution:

$$F(z) = \int_{-\infty}^{\infty} F(z|\theta_j)\, dG(\theta_j).$$

Consider now the example in section 5 in which $z(t-t_{0j},\theta_j) = 1 - e^{-(t-t_{0j})/\theta_j}$ and $G(\theta_j) = 1 - e^{-\theta_j/\theta_\mu}$. In this case, $\psi(\theta_j,z) = \theta_j \ln(1-z)$ and

$$F(z) = \int_0^{\infty} \left(1 - (1-z)^{g\theta_j}\right) \cdot \frac{e^{-\theta_j/\theta_\mu}}{\theta_\mu} \cdot d\theta_j = \frac{g\theta_\mu\left(\ln\left(\frac{1}{1-z}\right)\right)}{1 + g\theta_\mu\left(\ln\left(\frac{1}{1-z}\right)\right)}$$

as claimed in equation (16) above.

Letting $f(z)$ be the corresponding probability density function and integrating by parts yields:

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

To proof the claims in section 5.2 regarding the effects of g and θ_μ, 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 (A.1). Similarly, to proof the claims in section 5.3, notice that $\xi(\widetilde{z},\bar{z}_{MS}) + \xi(\bar{z}_{MS},1) = \xi(\widetilde{z},1)$, and therefore $\alpha(1-\delta^\alpha)\xi(\widetilde{z},\bar{z}_{MS}) + \alpha\xi(\bar{z}_{MS},1) = \alpha(1-\delta^\alpha)\xi(\widetilde{z},1) + \alpha\delta^\alpha\xi(\bar{z}_{MS},1)$. The results then follow again from $\partial\xi(z,1)/\partial\theta_\mu < 0$ and $\partial\xi(z,1)/\partial g < 0$.

[43] In particular, by introducing a rent-absorbing fixed factor in production, this effect would disappear. This extension would, however, complicate the general equilibrium.

References

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.

Aghion, Philippe, Mathias Dewatripont and Patrick Rey (1994), "Renegotiation Design with Unverifiable Information," *Econometrica*, 62:2, pp. 257-282.

Antràs, Pol (2003a), "Firms, Contracts, and Trade Structure," *Quarterly Journal of Economics*, 118:4, pp. 1375-1418.

Antràs, Pol (2003b), "Incomplete Contracts and the Product Cycle," Working Paper no. 9945. Cambridge: Mass,: NBER.

Antràs, Pol and Elhanan Helpman (2004), "Global Sourcing," *Journal of Political Economy*, 112:3, pp. 552-580.

Bloom, Martin (1992), *Technological Change in the Korean Electronics Industry*, Development Center Studies, OECD, Paris.

Caves, Richard E. (1996), *Multinational Enterprise and Economic Analysis*, Second Edition, Cambridge University Press.

Che, Yeon-Koo and Donald B. Hausch (1999), "Cooperative Investments and the Value of Contracting," *American Economic Review*, 89:1, pp. 125-147.

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.

Cyhn, J. W. (2002), *Technology Transfer and International Production: the Development of the Electronics Industry in Korea*, Edward Elgar, Cheltenham.

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.

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.

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.

Ethier, Wilfred J. (1986), "The Multinational Firm," *Quarterly Journal of Economics*, 101:4, pp. 805-833.

Ethier, Wilfred J. and James R. Markusen (1996), "Multinational Firms, Technology Diffusion and Trade," *Journal of International Economics*, 41:1, pp. 1-28.

Feenstra, Robert C. (1998), "Integration of Trade and Disintegration of Production in the Global Economy," *Journal of Economic Perspectives*, 12:4, 31-50.

Glass, Amy J., and Kamal Saggi (2002), "Licensing versus Direct Investment: Implications for Economic Growth", *Journal of International Economics*, 56:1, pp. 131-153.

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.

Grossman Gene M., and Elhanan Helpman (1991a), "Endogenous Product Cycles," *Economic Journal*, 101:4, pp. 1214-1229.

Grossman, Gene M., and Elhanan Helpman (1991b), "Quality Ladders and Product Cycles," *Quarterly Journal of Economics*, 106:2, pp. 557-586.

Grossman, Gene M., and Elhanan Helpman (2002), "Integration vs. Outsourcing in Industry Equilibrium," *Quarterly Journal of Economics*, 117:1, pp. 85-120.

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.

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.

Hart, Oliver and John Moore (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98:6, pp. 1119-1158.

Helpman, Elhanan (1984), "A Simple Theory of International Trade with Multinational Corporations", *Journal of Political Economy*, 92:3, pp. 451-471.

Hirsch, Seev (1967), *Location of Industry and International Competitiveness*, Oxford: Clarendon Press.

Jensen, Richard and Marie Thursby (1987), "A Decision Theoretic Model of Innovation, Technology Transfer, and Trade," *Review of Economic Studies*, 54:4, pp. 631-647.

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.

Krugman, Paul (1979), "A Model of Innovation, Technology Transfer, and the World Distribution of Income," *Journal of Political Economy*, 87:2, pp. 253-266.

McLaren, John (2000), "Globalization and Vertical Structure," *American Economic Review*, 90:5, pp. 1239-1254.

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.

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.

Markusen, James R. (1984), "Multinationals, Multi-Plant Economies, and the Gains from Trade" *Journal of International Economics*, 16, pp. 205-226.

Markusen, James R. (1995), "The Boundaries of Multinational Enterprises and the Theory of International Trade," *Journal of Economic Perspectives*, 9:2, pp. 169-189.

Moran, Theodore H. (2001), *Parental Supervision: The New Paradigm for Foreign Direct Investment and Development*, Washington DC: Institute for International Economics.

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.

Parry, Thomas G. (1975), "The Product Cycle and International Production: U.K. Pharmaceuticals," *Journal of Industrial Economics*, 24:1, pp. 21-28.

Segerstrom, Paul S., T. C. A. Anant, and Elias Dinopoulos (1990), "A Schumpeterian Model of the Product Life Cycle," *American Economic Review*, 80:5, pp. 1077-1091.

UNCTAD (1995), *World Investment Report: Transnational Corporations and Competitiveness*, New York: United Nations.

Vernon, Raymond (1966), "International Investment and International Trade in the Product Cycle," *Quarterly Journal of Economics*, 80:2, pp. 190-207.

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.

Williamson, Oliver E. (1985), *The Economic Institutions of Capitalism*, Free Press.

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.

Yi (2003), "Can Vertical Specialization Explain the Growth of World Trade?" *Journal of Political Economy*, 111:1, pp. 52-102.