Part III

Internalization

Chapter 6

The Transaction-Cost Approach

In the last two chapters, we have studied how imperfect contract enforcement across countries shapes the global sourcing decisions of firms. The models in those chapters have illustrated how, other things equal, a particularly weak institutional environment will act as a deterrent for foreign firms seeking offshoring opportunities in a particular country. Furthermore, we have studied which type of product and industry characteristics tend to be associated with a disproportionately deleterious effect of weak contracting on the profitability of offshoring.

As argued in Chapter 1, choosing the locations from which a firm sources its inputs is just one of the many organizational decisions that firms need to make when designing their global production strategies. In this chapter and the next, I will turn attention to a second key organizational decision of firms, namely the extent of control that firms choose to exert over the production of the different parts and components in their value chain. In many circumstances, ownership of the input producer's physical assets is the key method to enhance such control. For this reason, this decision is often dubbed *internalization*, since equity ownership amounts to bringing the production of intermediate inputs inside firm boundaries. When such internalization occurs across borders, the investing entity becomes a multinational firm and any flow of physical goods between related parties constitutes intrafirm international trade.[1]

[1]The internalization decision is key in John Dunning's (1981) celebrated *OLI* or 'ecletic' theory of the multinational firm, where OLI is an acronym for Ownership, Location, and Internalization. Put succinctly, the emergence of the multinational firm is explained by an **O**wnership advantage stemming from firm-specific assets that allow firms to compete in

At the risk of oversimplifying the global organizational decisions of firms, it may be useful to represent these decisions in terms of choosing a cell in the following two-by-two matrix, in which the rows denote different location decisions (domestic or foreign sourcing), while the columns are associated with distinct internalization choices:

	Internal Procurement	**External Procurement**
Domestic Sourcing	Domestic Integration	Domestic Outsourcing
Foreign Sourcing	Foreign Integration	Foreign Outsourcing

Note that when the location decision is associated with foreign sourcing, internalization will be associated with foreign direct investment and intrafirm international trade in physical goods, while in the case of foreign outsourcing, any exchange of goods will also be recorded in international trade statistics, but it will constitute non-related-party trade.

Why do firms find it optimal to carry out certain production stages within firm boundaries while conducting others at arm's-length? Beginning with the seminal work of Coase (1937), the leading approach to answering this question posits that activities take place within or across firm boundaries depending on which of these organizational modes leads to the minimization of transaction costs. Furthermore, it is widely accepted that transaction costs stem, to a large degree, from contractual incompleteness. Indeed, firm boundaries would be indeterminate and irrelevant in a world in which transactions were governed by comprehensive contracts that specify (in an enforceable way) the course of action to be taken in *any* possible contingency that the contracting parties may encounter. In such a case, formal contracting would leave no room for residual rights of control to matter for economic decisions.

The main unifying theme of the theoretical literature on (multinational) firm boundaries is thus the departure from the classical assumption of complete or perfect contracting. Nevertheless, different theories of internalization emphasize different types of contractual frictions and they also adopt different approaches with regards to how the internalization of transactions affects these frictions. In this Part III of the book, I will restrict the analysis to what I consider to be the two leading approaches to the analysis of the internalization decision. On the one hand, the so-called *Transaction-Cost Theory*

unfamiliar environments, a **L**ocation advantage that makes it efficient to exploit the firm assets in production facilities in multiple countries, and an **I**nternalization advantage that makes the within-firm exploitation of assets dominate exploitation at arm's-length.

focuses on describing the type of contractual frictions that naturally emerge in arm's-length transactions when contracts are incomplete, but tends to be much less precise about the source of transaction costs in internalized transaction. On the other hand, the *Property-Rights Theory* assumes that the contractual sources of transaction costs are not too distinct in internal versus external transactions, and instead places at the center stage the study of how the allocation of ownership rights over physical assets shapes the size of transaction costs under different organizational modes.

The transaction cost theory, which draws inspiration in Coase (1937) but is mostly associated with the work of Williamson (1971, 1975, 1985), has arguably been the leading paradigm in the analysis of the internalization decision in international environments. In line with the theory, it is typically perceived that vertical (or lateral) integration is an effective way for firms to deal with situations of contractual incompleteness in international transactions, in which it may be hard to provide incentives to subcontracted producers.

To give a particular example, consider Boeing's organizational decisions in recent years as it was struggling to complete the production phase of the new 787 Dreamliner. According to Boeing's website, the 787 'Development Team' encompasses 50 suppliers located in 10 countries and the involvement of foreign suppliers is not anecdotal as it accounts for close to 70 percent of the aircraft's parts (Newhouse, 2007, p.29).[2] The repeated delays experienced during the production phase, which ran more than three years behind schedule, have been ascribed in part to the fact that multiple suppliers did not stand by their contractual obligations.[3] Boeing responded to these delays by partially reorganizing their sourcing model and bringing some of the problematic upstream production stages within their firm boundaries. For example, during 2008 and 2009, Boeing successively acquired operations from Vought Aircraft Industries, a company that was producing the rear sections of the Dreamliner's fuselage but that had been identified as a problematic supplier. One of those acquisitions entailed forming a 50-50 joint venture with a subsidiary of Italy's Alenia Aeronautica, another key supplier for Boeing, from which it procures the horizontal stabilizer and the center fuselage for the Dreamliner, and with which it has also wrestled in past years.

The experience of Boeing illustrates that firms often design the ownership structure along their value chain in a way that attempts to minimize frictions

[2]The full list of suppliers, as of June of 2014, can be found at http://www.boeing.com/commercial/787family/dev_team.html.

[3]See http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aF6uWvMb9C08 for a more complete account.

resulting from suppliers not honoring their contractual commitments. As argued by the property-rights theory, however, vertical integration is not always a panacea to deal with those situations; the problem of getting procurement incentives right does not simply disappear when firms source inputs internally. In fact, it is not uncommon for firms to fail to get their own divisions to produce what they want at low cost, thereby leading to externalization of certain production processes and foreign direct *divestment* in international environments. One example of such an externalization is provided by Sony Corporation, which in 2010 decided to sell a 90 percent stake in its LCD TV assembly plant in Nitra, Slovakia, to Taiwanese electronics parts maker Hon Hai Precision Industry (which operates under the trade name Foxconn). As admitted by Sony, the decision was motivated by a desire to cut fixed costs and turn around their loss-making TV operations, and had been preceded a year earlier by a similar divestment in its production plant in Tijuana, Mexico.[4]

In the next two chapters, I will overview in more detail the theoretical underpinnings of these two mainstream theories of the firm, and will then show how to apply them to the study of the global ownership decisions of multinational firms. In the process, I will derive a series of testable implications that distinguish these two approaches and that will guide the empirical analysis carried out in Chapter 8.

The Transaction-Cost Approach: A Non-Technical Overview

The late Ronald Coase was the father of transaction-cost economics and his 1937 article in *Economica* marks the birth of this field of study in Economics. His paper can be interpreted as a reaction to the neoclassical approach to the optimal size of firms, which was being developed at that time. To put things in historical perspective, Viner's 1932 celebrated paper on cost and supply curves had appeared only five years earlier. The neoclassical approach was technological in nature and treated optimal scale of operation as being determined by a firm's profit-maximization problem given a cost function obtained by choosing inputs in a cost-minimizing fashion.

Coase (1937) instead argued that there are substantial transaction costs associated with running the economic system, and that firms emerge precisely when certain transactions can be undertaken with less transaction costs inside the firm than through the market mechanism. Or in his own words:

[4]For more details, see http://www.reuters.com/article/2010/03/31/sony-honhai-idUSTOE62U08020100331 and http://www.pcworld.com/article/171181/article.html.

> "The main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism" (1937, p. 390).

In his article, Coase mentions a few transaction-cost disadvantages of market transactions, including the costs of negotiating and concluding a separate contract for each exchange transaction and the costs of specifying all possible contingencies in a long-term contract. Coase also proposed certain factors that might limit the size of the firm, such as decreasing returns to the entrepreneur function or an increasing supply price of some factors, but his treatment of the costs of integration was closest in spirit to the one used by neoclassical approach to rationalize upward sloping marginal cost curves.

Coase's view of the firm did not instantly become part of mainstream economics. It was criticized for its vagueness and was dubbed tautological. What kind of empirical evidence would be supportive of Coase's theory? How could the theory be refuted? Between 1940 and 1970, the literature thus focused instead on exploring technological theories of the firm.

Oliver Williamson brought transaction-cost considerations back into the spotlight by developing much more explicit theories of the inefficiencies of market transactions, thus making the transaction-cost approach operational. His theory, as laid out in his classic 1985 book *The Economic Institutions of Capitalism*, is based on three concepts: (1) bounded rationality, (2) opportunism and (3) asset specificity.

Williamson appeals to *bounded rationality* to provide a foundation for the incompleteness of contracts. In particular, in a complex and unpredictable world, boundedly rational agents will be unable to plan ahead for all the contingencies that may arise. Furthermore, even when contingencies are foreseen, it may be hard for contracting parties to negotiate about these plans because of limited capability of describing these possible states. Finally, even when parties can plan and negotiate these contingencies, it may be hard for a third party to verify them and enforce the contract. As a result, formal contracts will tend to be incomplete and will tend to be renewed or renegotiated as the future unfolds.

By *opportunism*, Williamson means that economic actors are "self-interest seeking with guile" (1985, p. 47). The fact that agents are opportunistic is a necessary condition for the incompleteness of contracts to lead to inefficiencies. If agents could credibly pledge at the outset to execute the contract efficiently, then although the contract would have gaps, renegotiation would always occur in a joint profit maximizing manner.

Finally, Williamson points out that certain assets or investments are

relationship-specific, in the sense that the value of these assets or investments is higher inside a particular relationship than outside of it. This is important because it implies that, at the renegotiation stage, parties cannot costlessly switch to alternative trading partners and are partially locked in a bilateral relationship. This is what Williamson calls the "*fundamental transformation*" from an ex-ante competitive situation to one of bilateral monopoly.

Williamson then shows how the combination of contract incompleteness, opportunism and specificities gives rise to inefficiencies which can be interpreted as equilibrium transaction costs. In his own work, Williamson primarily emphasized ex-post inefficiencies related to haggling (i.e., parties spending resources in order to improve their bargaining position), which could lead to inefficient terminations or executions of contracts. Subsequent work, beginning with the important contribution of Klein, Crawford and Alchian (1978), has largely focused on ex-ante or hold-up inefficiencies associated with the suboptimal provision of relationship-specific investments.

By now, the reader should have a sneaking suspicion that Williamson's notion of transaction cost inefficiencies, particularly in their ex-ante form, appears to be closely related to the type of contractual inefficiencies described in Chapter 4 of this book. I will indeed confirm such a connection very shortly.

A key limitation of the transaction-cost approach, even in its most refined versions developed by Williamson, is that it does not have an awful lot to say about the nature of transaction costs associated with intrafirm transactions. In Williamson's view, contract incompleteness, opportunism and relationship-specificity are irrelevant within firm boundaries because decisions within firms are taken by fiat, so agents would not bother to spend time and resources haggling over profits. Instead, Williamson appealed to a broad notion of *governance costs* to justify the limits to firm size.

Despite the vagueness in its description of the costs of internalization, Williamson's careful description of the sources of transaction costs in market transactions provided an operational theory that could be mapped to particular observable economic variables, such as measures of contractual complexity of relationship-specificity. It is thus not surprising that a rich empirical literature on the determinants of the internalization emerged in the 1980s, shortly after Williamson's seminal work. The key papers in this literature are surveyed in Lafontaine and Slade (2007). Some of these empirical contributions relate to internalization decisions by multinational firms in open-economy environments, so I will return to them in Chapter 8.

A Transaction-Cost Model of Multinational Firm Boundaries

As many readers will have recognized, there is a clear connection between the informal description of the transaction-cost approach to the theory of the firm above and the model of global sourcing with contractual frictions developed in Chapter 4. In particular, in that chapter I illustrated how the combination of imperfect contracting, bargaining by self-interested parties, and the existence of relationship-specific investments led to production inefficiencies that would, other things equal, reduce the profitability of global sourcing. An important distinctive feature of the model in Chapter 4 is that it emphasized weak contract enforcement – rather than bounded rationality – as the cause for the lack of complete contracting. As we shall explore in Chapter 8, this will be of some relevance when choosing the variables used to proxy for the key objects in the model. A second, perhaps less crucial distinguishing aspect of the models in Chapter 4 is that they focused attention on ex-ante, hold-up inefficiencies (as in Klein, Crawford and Alchian, 1978), rather than on ex-post inefficiencies which featured prominently in Williamson's work.

There is, however, one fundamental dimension in which in Chapters 4 and 5 I departed from the transaction-cost approach. More specifically, both when discussing the models and when taking them to the data, I made no distinction between within-firm and across-firm sourcing decisions. The model introduced two agents - a final-good producer F and a manufacturing plant manager M, both with a zero reservation utility, but I was not explicit about whether or not the manufacturing plant was vertically integrated. In other words, I did not specify whether M was an employee of F or an independent subcontractor. In the empirical analysis, I studied the implications of the model for the share of imported inputs coming from different markets, regardless of whether those imports were transacted within or across firm boundaries.

The transaction-cost theory argues, however, that the production inefficiencies identified in Chapter 4 should apply only to market transactions, and not to those conducted within firm boundaries. In choosing whether or not to own upstream producers, firms trade off these contractual inefficiencies against the higher *governance costs* associated with intrafirm procurement.

Let us now turn to a more formal exposition of a transaction-cost model of multinational firm boundaries. In line with the two-by-two matrix introduced at the beginning of this chapter, I will first consider a simple two-country model along the lines of the benchmark global-sourcing model in Chapter 4. A continuum of heterogeneous final-good producing firms based in a rich

North combine locally-produced headquarter services with manufacturing components that can be produced domestically or in a foreign location, which is referred to as South. Offshoring lowers the marginal cost of production of Northern firms but entails disproportionately high fixed costs, and thus some final-good producers optimally opt out of procuring inputs abroad. I continue to index the location choice by a subscript $\ell = D$ or $\ell = O$, for **D**omestic sourcing and **O**ffshoring, respectively. Final-good production and headquarter services provision are overviewed by agent F in the North, while manufacturing production is controlled by agent M in the location where manufacturing takes place.

The main novelty in the framework is that I will now incorporate a decision as to whether M is an employee of F or the manager of an unaffiliated subcontractor. I will assume that the models developed in Chapter 4 capture accurately the type of contractual difficulties encountered in market transactions, and as a result the profit functions associated with domestic outsourcing and foreign outsourcing correspond to those derived in that chapter (more on this below). Conversely, when F internalizes its provision of manufacturing components, it can make all relevant decisions by fiat and thus can choose all investment levels in a profit-maximizing manner regardless of the contracting environment. I capture the notion of governance costs by assuming that the marginal cost of input provision is multiplied by a factor $\lambda > 1$ when M is integrated by F. Furthermore, I shall assume that the fixed costs of production are also higher under an integrated structure than under a non-integrated one.

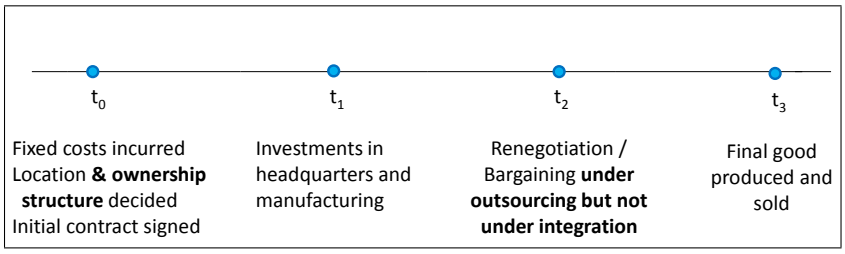

Figure 6.1: Timing of Events

Figure 6.1 depicts the timing of events in the model. This is analogous to Figure 4.1 in Chapter 4, but it includes (in bold) the new actions or assumptions associated with the modeling of the internalization decision. Notice that F is still allowed to demand an ex-ante transfer from potential M agents for their participation in production and that the key contractual difference between integration and outsourcing is that renegotiation and bargaining only

occur under the latter, while under the former, F's authority dictates M's abidance by whatever was stipulated in the initial contract.

I will index the internationalization decision by a second subscript $k = V$ or $k = O$, to refer to **V**ertical integration and **O**utsourcing, respectively. In sum, agent F will choose from one of four potential organizational modes, $(\ell, k) \in \{DO, DV, OO, OV\}$, associated with domestic outsourcing, domestic vertical integration, offshore outsourcing, and offshore vertical integration (or vertical FDI). Assuming that the fixed costs of offshoring are higher than those associated with domestic sourcing regardless of the ownership structure, we have that the ranking of fixed costs satisfies

$$f_{OV} > f_{OO} > f_{DV} > f_{DO}. \tag{6.1}$$

This is the same ranking of fixed costs adopted by Antràs and Helpman (2004, 2008), but I will discuss below the robustness of the theoretical predictions to alternative rankings of fixed costs.

Let us now discuss the equilibrium profitability of the different organizational forms available to the Northern final-good producer. In the case of vertical integration, F's authority and the existence of unrestricted lump-sum transfers ensures that these profit flows will be given by

$$\pi_{DV}(\varphi) = (w_N)^{1-\sigma} B \varphi^{\sigma-1} \lambda^{1-\sigma} - w_N f_{DV} \tag{6.2}$$

under domestic integration, and by

$$\pi_{OV}(\varphi) = \left((w_N)^{\eta} (\tau w_S)^{1-\eta}\right)^{1-\sigma} B \lambda^{1-\sigma} \varphi^{\sigma-1} - w_N f_{OV} \tag{6.3}$$

under offshore integration, where remember that the market demand term B is defined as

$$B = \frac{1}{\sigma} \left(\frac{\sigma}{(\sigma - 1) P} \right)^{1-\sigma} \beta \left(w_N L_N + w_S L_S \right).$$

Importantly, these profit flows apply to all the variants of the global sourcing model studied in Chapter 4 which considered the inclusion of generalized bargaining power, constraints on ex-ante transfers, variable degrees of relationship-specificity, partial contractibility, multiple suppliers, and a sequential production process.[5]

[5] In Chapter 4, I rationalized the existence of limitations on ex-ante transfers to potential misbehavior on the part of suppliers. Under the assumptions of the transaction-cost theory, such misbehavior can be avoided through the use of authority, and thus it is natural to focus on the case in which they have no impact on the profitability of sourcing.

The profitability of domestic outsourcing and offshore outsourcing will instead vary with changes in the contracting and economic environments, but as in Chapter 4 we can write these profit levels succinctly as

$$\pi_{DO}(\varphi) = (w_N)^{1-\sigma} B\Gamma_{DO}\varphi^{\sigma-1} - w_N f_{DO} \tag{6.4}$$

and

$$\pi_{OO}(\varphi) = \left((w_N)^{\eta}(\tau w_S)^{1-\eta}\right)^{1-\sigma} B\Gamma_{OO}\varphi^{\sigma-1} - w_N f_{OO}, \tag{6.5}$$

where $\Gamma_{DO} \leq 1$ and $\Gamma_{OO} \leq 1$ denote the contractual efficiency associated with domestic and offshore outsourcing, respectively. The determination of these terms was developed in detail in Chapter 4, so I will not repeat it here. Below, however, I will review some of the key insights that were obtained in Chapter 4 when studying the dependence of Γ_{DO} and Γ_{OO} on the deep parameters of the model.

Equilibrium Sorting

With the profit functions (6.2) through (6.5) in hand, we can now study the optimal organization of production, as summarized by the choice of location $\ell \in \{D, O\}$ and of ownership $k \in \{V, O\}$. As a first step, I will first describe how, depending on their core productivity parameter φ, firms self-select into different organizational modes. In order to build intuition and focus on the key objects that will be taken to the data in Chapter 8, I will begin by considering a benchmark case in which contracting is complete when sourcing domestically, which implies that $\Gamma_{DO} = 1$ in equation (6.4). An immediate implication of this assumption is that $\pi_{DO}(\varphi) > \pi_{VO}(\varphi)$ for all φ, and thus domestic integration is a dominated strategy in this scenario and can be safely ignored for the time being. We will later introduce versions of the model with a well-defined choice between domestic integration and outsourcing. This simplifying assumption is, however, not too unrealistic when focusing on the U.S. case, for which Atalay, Hortacsu and Syverson (2013) have shown that within-firm shipments of physical goods account for a very small share of overall shipments of U.S. establishments.

The remaining three profit functions are all linearly increasing in $\varphi^{\sigma-1}$ with a slope that is inversely related to governance costs λ in the foreign integration case and positively related to the index of contractual efficiency Γ_{OO} in the offshore outsourcing case. Figure 6.2 depicts these profit functions whenever governance costs are disproportionately high, so that foreign direct investment is a strictly dominated strategy since it entails high fixed and variable costs. In the figure, it is also implicitly assumed that wage differences across countries are high relative to the contractual inefficiencies of

offshore outsourcing, so that the most productive firms in an industry find the latter strategy to be the profit-maximizing one. The sorting pattern is thus analogous to that discussed in Chapter 4, but with offshoring taking the specific form of subcontracting.

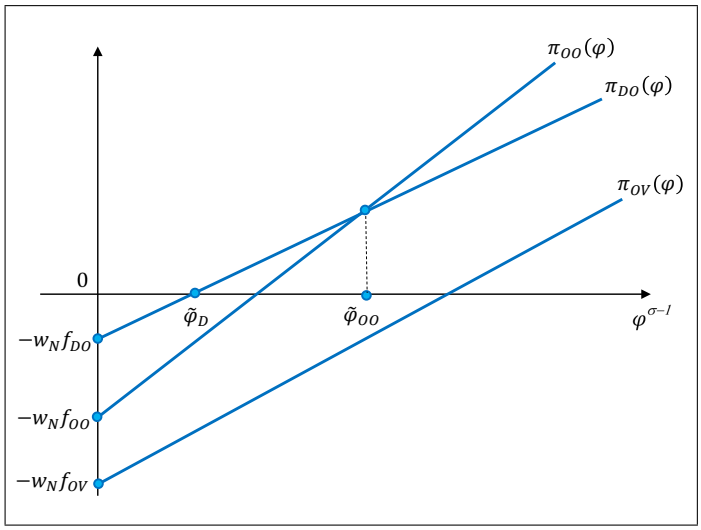

Figure 6.2: Equilibrium Sorting with High Governance Costs

In Figure 6.3, I depict instead the case in which, for reasons that I will review shortly, the contractual efficiency of foreign outsourcing is very low, and no firm finds it optimal to use this strategy when sourcing inputs abroad. As shown in the figure, this does not rule out the possibility that some relatively productive firms will still find it optimal to offshore in the South, but those transactions will unavoidably happen within firm boundaries. Naturally, for this to be the case, it is necessary for governance costs λ to be sufficiently low relative to wage differences across countries.[6]

Finally, in Figure 6.4, I illustrate the most interesting case in which wage differences are large relative to both arm's-length transaction costs and governance costs. In such a case, one can easily construct equilibria of the type depicted in the figure, with a positive measure of active firms adopting each of the three candidate organizational forms. The least productive active firms make use of domestic outsourcing, the most productive firms engage

[6]To avoid describing a taxonomy of cases, I ignore situations in which wage differences are low relative to both governance costs and contractual inefficiencies associated with market transactions. In those cases, all active firms would simply resort to domestic sourcing.

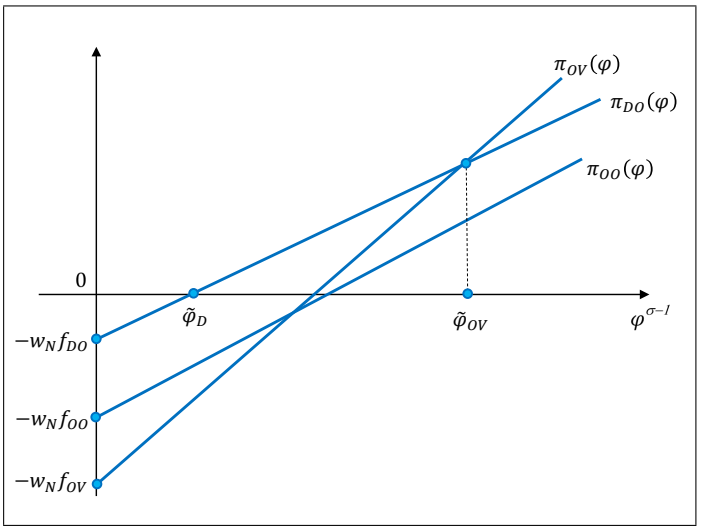

Figure 6.3: Equilibrium Sorting with High Oustourcing Contractual Inefficiencies

in vertical foreign direct investment (offshore integration), while firms with intermediate levels of productivity opt for foreign outsourcing.

It may seem that the model generates a plethora of possible types of equilibria, with small changes in the parameters leading to large changes in the nature of the equilibrium. Fortunately, however, this is not the case. First, it is important to emphasize a key robust prediction emerging from the sorting patterns depicted in Figures 6.2, 6.3 and 6.4; namely, when foreign outsourcing and foreign integration coexist within an industry (so the share of intrafirm trade is strictly between 0 and 1), it is necessarily the case that firms integrating foreign suppliers are more productive than those choosing subcontracting. Thus, the model predicts a similar type of performance premia for multinationals (relative to firms resorting to foreign outsourcing) as the one observed in the data and documented in Chapter 2 for exporters (relative to non-exporters) and for offshoring firms (relative to firms sourcing domestically).

A second reason not to be dismayed by the different cases illustrated in Figures 6.2, 6.3 and 6.4 is that although small changes in parameters might lead to the disappearance of a certain type of organizational form from the industry equilibrium, it is necessarily the case that this particular organizational form accounted for a very small fraction of the sourcing activity in the industry. Put differently, and as I am about to show formally, the relative prevalence of the different organizational forms within an industry varies

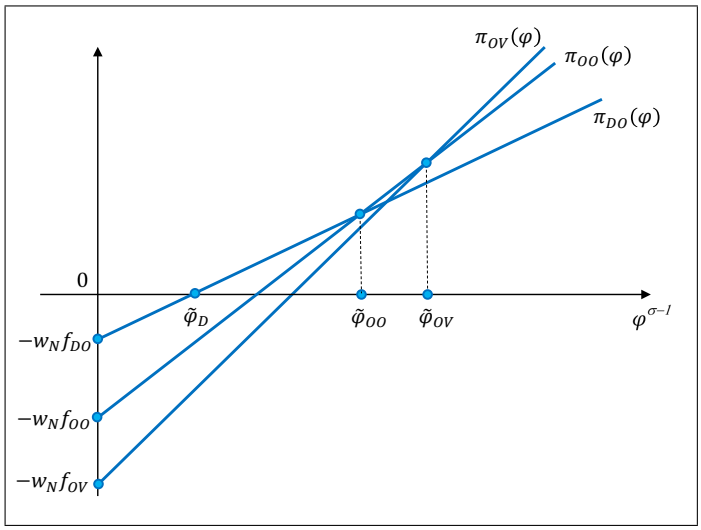

Figure 6.4: Equilibrium Sorting with Low Contractual Inefficiencies and Governance Costs

smoothly with the parameters of the model.

Relative Prevalence and Intrafirm Trade Shares

We are now ready to aggregate the decisions of the various firms in a sector and study the determinants of the relative importance of different organizational forms – domestic sourcing, foreign outsourcing, foreign integration – in a particular industry. As in Chapters 2, 4 and 5, I measure the relative prevalence of an organizational form by the relative magnitude of intermediate input purchases under that organizational form. In principle, one could aggregate arm's-length and intrafirm foreign input purchases and use the framework above to revisit the determinants of the share of *imported* manufacturing inputs over total manufacturing input purchases in a particular industry. I will not do so here because the results of such an exercise deliver results essentially identical to those derived in Chapter 4 and tested in Chapter 5. Instead, I will focus on describing the determinants of the relative prevalence of intrafirm versus arm's-length foreign sourcing decisions, as measured by the share of intrafirm intermediate input imports over total imported inputs. As explained in Chapter 1 and described in more detail in Chapter 8, this is an object we can attempt to measure with available data, so it seems natural to focus on it.

In order to trace the implications of our transaction-cost model for input

purchases, one needs again to take a stance on how these inputs are priced in market transactions. As in Chapter 4, I follow the agnostic approach of assuming that foreign inputs are priced such that these input expenditures constitute the same multiple $(1-\eta)(\sigma-1)$ of operating profits as in the case of foreign intrafirm input purchases. Under this assumption, the share Sh_{i-f} of intrafirm imported inputs over the total imported input purchases is given by

$$Sh_{i-f} = \frac{\lambda^{1-\sigma}\int_{\tilde{\varphi}_{OV}}^{\infty}\varphi^{\sigma-1}dG(\varphi)}{\Gamma_{OO}\int_{\tilde{\varphi}_{OO}}^{\tilde{\varphi}_{OV}}\varphi^{\sigma-1}dG(\varphi)+\lambda^{1-\sigma}\int_{\tilde{\varphi}_{OV}}^{\infty}\varphi^{\sigma-1}dG(\varphi)}, \tag{6.6}$$

where $\tilde{\varphi}_{OO}$ is such that

$$\frac{(w_N)^{\sigma}(f_{OO}-f_{DO})}{B\left[\left(\frac{w_N}{\tau w_S}\right)^{(1-\eta)(\sigma-1)}\Gamma_{OO}-1\right]} = (\tilde{\varphi}_{OO})^{\sigma-1}, \tag{6.7}$$

and $\tilde{\varphi}_{OV}$ is such that

$$\frac{(w_N)^{\sigma}(f_{OV}-f_{OO})}{\left(\frac{w_N}{\tau w_S}\right)^{(1-\eta)(\sigma-1)}B\left[\lambda^{1-\sigma}-\Gamma_{OO}\right]} = (\tilde{\varphi}_{OV})^{\sigma-1}. \tag{6.8}$$

For some firms to potentially find foreign outsourcing more profitable than domestic sourcing, we need to assume that $\Gamma_{OO} > (w_N/\tau w_S)^{-(1-\eta)(\sigma-1)}$, while for some firms to possibly prefer foreign insourcing to foreign outsourcing, it is also necessary that $\lambda^{1-\sigma} > \Gamma_{OO}$. Note that these two conditions guarantee that both $\tilde{\varphi}_{OO}$ and $\tilde{\varphi}_{OV}$ are positive, but they are *not* sufficient to guarantee that $\tilde{\varphi}_{OV} > \tilde{\varphi}_{OO}$ or that the share of intrafirm trade in (6.6) falls strictly between 0 and 1 (more on this below). For the time being we shall sidestep this issue and assume that we are in an industry equilibrium in which all organizational forms are chosen by a positive measure of firms, as depicted in Figure 6.4 above.

Inspection of equations (6.6), (6.7) and (6.8) reveals that, holding constant the market demand level B, a decrease in governance costs (lower λ) decreases the threshold productivity level $\tilde{\varphi}_{OV}$, while leaving $\tilde{\varphi}_{OO}$ unchanged. As a result, it appears that the share of intrafirm trade tends to be higher in lower governance cost industries. On the other hand, a higher foreign outsourcing contractual efficiency Γ_{OO} increases $\tilde{\varphi}_{OV}$ and reduces $\tilde{\varphi}_{OO}$, thereby tending to reduce Sh_{i-f}. These predictions should be treated with caution, however, because these parameter changes will affect the residual demand level B as well.

In order to obtain sharper comparative statics, we next turn to the case in which the distribution of core productivity φ is Pareto with shape parameter $\kappa > \sigma - 1$, as in previous chapters. Such a parameterization is also necessary to tease out the effects of lower trade costs, lower Southern wages or lower headquarter intensity η, which are all associated with lower levels of *both* $\tilde{\varphi}_{OO}$ and $\tilde{\varphi}_{OV}$. Assuming a Pareto distribution of productivity, equations (6.6), (6.7) and (6.8), can be reduced to

$$Sh_{i-f} = \frac{\lambda^{1-\sigma}}{\Gamma_{OO}\left[\left(\frac{\tilde{\varphi}_{OV}}{\tilde{\varphi}_{OO}}\right)^{\kappa-\sigma-1} - 1\right] + \lambda^{1-\sigma}} \tag{6.9}$$

where

$$\frac{\tilde{\varphi}_{OV}}{\tilde{\varphi}_{OO}} = \left[\frac{f_{OV} - f_{OO}}{f_{OO} - f_{DO}} \times \frac{\Gamma_{OO} - (w_N/\tau w_S)^{-(1-\eta)(\sigma-1)}}{\lambda^{1-\sigma} - \Gamma_{OO}}\right]^{1/(\sigma-1)}. \tag{6.10}$$

Note that the share of intrafirm imported inputs is now only a function of parameters and the ratio of thresholds $\tilde{\varphi}_{OV}/\tilde{\varphi}_{OO}$, which itself is independent of the residual demand level B.

These features greatly simplify the characterization of Sh_{i-f}. First, it is now clear that lower governance costs λ or lower foreign outsourcing contractual efficiency Γ_{OO} are associated with a higher share of intrafirm imports, which are intuitive results. It is also apparent and obvious that lower fixed costs of foreign integration f_{OV} or higher fixed costs of offshore outsourcing f_{OO} will also tend to increase the share of intrafirm trade. Less trivially, equations (6.9) and (6.10) also reveal that the share of intrafirm imported inputs is decreasing in the term $(w_N/\tau w_S)^{(1-\eta)(\sigma-1)}$ and is thus increasing in headquarter intensity η, trade costs, and Southern labor costs. The extensive margin of offshoring is key for understanding these effects. Intuitively, decreases in these parameters lead firms that were sourcing domestically to select into offshoring, but as observed in Figure 6.4, these new foreign input purchases necessarily occur at arm's-length, hence decreasing the relative prevalence of intrafirm imports in overall imports. For the same reason, the share Sh_{i-f} is also decreasing in the fixed costs of domestic sourcing f_{DO}. Finally, and in analogy to our results in Chapters 2 and 4, the fact that firms engaged in foreign integration are more productive than those engaged in offshore outsourcing translates into a positive effect of productivity dispersion (a lower κ) on the share of intrafirm trade.

I have focused so far on equilibria of the type depicted in Figure 6.4 in which all organizational forms are chosen by a positive measure of firms. It

should be clear, however, that if λ becomes larger and larger, not only will the share of intrafirm trade decline, but if eventually $\lambda^{1-\sigma}$ becomes lower than Γ_{OO}, then no firm will find it optimal to engage in foreign integration and the share of intrafirm trade will be 0 regardless of the particular value of λ. Similarly, if Γ_{OO} becomes lower and lower, the share of intrafirm trade rises monotonically until Γ_{OO} exceeds a threshold level $\tilde{\Gamma}_{OO}$ at which $Sh_{i-f} = 1$ regardless of the particular value of $\Gamma_{OO} > \tilde{\Gamma}_{OO}$.[7]

Determinants of Foreign Outsourcing Efficiency

The result that the relative prevalence of intrafirm imported input purchases is increasing in the transaction costs associated with offshore outsourcing is intuitive but of little use in guiding an empirical analysis of the determinants of the relative prevalence of foreign integration in the data. Fortunately, in Chapter 4 we discussed at length the determination of arm's-length contractual efficiency Γ_{OO} (referred to as Γ_O in that chapter) in several types of environments featuring generalized bargaining power, constraints on ex-ante transfers, variable degrees of relationship-specificity, partial contractibility, multiple suppliers and a sequential production process. As summarized in Table 5.5, for instance, remember that Γ_{OO} is decreasing in the elasticity of demand σ, the level of financial constraints $(1-\phi)$, and the degree of input relationship-specificity ϵ, while it is increasing in the level of Southern contractibility μ_S and input substitutability ρ. The effect of headquarter intensity η on the level of transaction-costs is instead generally ambiguous.

When combining these comparative statics with those discussed above regarding the effect of Γ_{OO} and other parameters on Sh_{i-f} in equation (6.9), we can conclude that the share of intrafirm imports is shaped by the various 'deep' parameters of our global sourcing models in the following manner:

$$Sh_{i-f} = Sh_{i-f}\left(\underset{-}{\lambda}, \underset{-}{w_N/w_S}, \underset{+}{\tau}, \underset{-}{\kappa}, \underset{-}{\phi}, \underset{-}{\mu_S}, \underset{+}{\epsilon}, \underset{-}{\rho}, \underset{?}{\sigma}, \underset{?}{\eta}\right). \tag{6.11}$$

I will defer until Chapter 8 a discussion of the implementation of an empirical test of these predictions, but a general insight obtained in transaction-cost models of global sourcing is that the same type of parameters that in Chapter 4 were associated with *high* offshoring shares (e.g., a high w_N/w_S, ϕ, μ_S, or ρ, and a low τ or ϵ) will now tend to be associated with *low* intrafirm trade

[7] The threshold value is given by $\tilde{\Gamma}_{OO} = \varpi\lambda^{1-\sigma} + (1-\varpi)\left(w_N/\tau w_S\right)^{-(1-\eta)(\sigma-1)}$ where $\varpi = \left(f_{OO} - f_{DO}\right)/\left(f_{OV} - f_{DO}\right)$.

shares.[8] Similarly, as I will further elaborate in Chapter 8, the same type of interactions of industry and country characteristics that we demonstrated in Chapters 4 and 5 to affect positively the efficiency of offshore outsourcing, are now expected to exert a negative effect on intrafirm trade shares.

Downstreamness and Integration

As shown throughout this book, in the various versions of the global sourcing model, it is relatively easy to transition from the individual firm-level organizational decisions to predictions for the share of intrafirm imported inputs at the sectoral level. Unfortunately, the variant of the model with sequential production is a bit more difficult to handle because it yields predictions for the differential incentives to integrate different production stages by firms in the sector under study. As shown by Antràs and Chor (2013), a natural way to deal with this complication is to aggregate the decisions of all firms in all sectors related to the optimal sourcing of the individual stages of production. To give a specific example, the idea is to aggregate the decisions of all firms buying a specific input, say rubber, and then studying the extent to which those purchases are internalized.

In the presence of fixed costs of outsourcing and integration at the input level, Antràs and Chor (2013) then show that the model can be aggregated to deliver implications for the relative prevalence of integration of a particular input as a function of the position (i.e., downstreamness) of that input in the value chain. Furthermore, under the plausible assumption that the input-specific costs of integration are higher than those under outsourcing, the model delivers the implication that more productive firms will tend to integrate a larger share of the inputs they use in production, and that the share of firms integrating a particular stage is weakly increasing in the dispersion of productivity across the buyers of that input.

How does the downstreamness of an input affect the incentives of firms to integrate that input in the transaction-cost model? In order to answer this question, one can refer back to Chapter 4, where we discussed the choice between domestic sourcing and offshoring whenever the main contractual difference between these two modes was in their associated level of contractibility. There we argued that (contractually-insecure) foreign sourcing is particularly appealing in downstream stages in the sequential complements case, but particularly unappealing in those same downstream stages in the

[8] The overall effect of σ is nevertheless ambiguous because the negative effect of σ on Γ_{OO} is counterbalanced by its negative effect on $\lambda^{1-\sigma}$, as well as by the direct effects working via the terms in σ in (6.9) and (6.10).

sequential substitutes case. In a similar vein, in the Theoretical Appendix (section A.3), I develop a sequential model of production with a transaction-cost determination of firm boundaries. More specifically, under outsourcing, production and contracting decisions are as described in Chapter 4. Instead, under integration, governance costs inflate marginal costs by a factor $\lambda > 1$, but supplier investment levels can be set to maximize the full incremental contribution $\Delta R(v)$ in (4.29) rather than $(1 - \beta_O)\Delta R(v)$, and the final-good producer captures all surplus from the relationship.

As shown in the Theoretical Appendix, this sequential model delivers the prediction that downstreamness should have a negative effect on foreign integration relative to offshore outsourcing whenever inputs are sequential complements, while it should have a positive effect on foreign integration when inputs are sequential substitutes. Intuitively, in the sequential complements case, it is particularly important to eliminate (via integration) the contractual inefficiencies of upstream investments because of the positive trickle-down effect of high investments in those early stages. Conversely, in the sequential substitutes case, high values of upstream stages are now relatively detrimental, so internalization is least advantageous in those early stages.

Contractual Frictions in Domestic Sourcing

In our analysis we have so far simplified matters by assuming complete contracting in domestic input purchases, which made domestic integration a dominated strategy for firms. It should be clear, however, that when contracts are also incomplete in domestic transactions (so $\Gamma_{DO} < 1$), and both governance costs and wage differences are small, it may be the case that a subset of firms finds domestic integration optimal. Given our assumption on the ranking of fixed costs in equation (6.1), any equilibrium in which all four possible organization forms (domestic outsourcing, domestic integration, offshore outsourcing and offshore integration) are chosen by a positive measure of firms must satisfy the sorting depicted in Figure 6.5. This sorting pattern is analogous to that in Figure 6.4 but notice that domestic integration emerges as the most appealing option for a set of firms with productivity levels above those of firms choosing domestic outsourcing, but below those of firms choosing offshore outsourcing.

Following the approach in Antràs and Helpman (2004, 2008), one could use this variant of the model to study the determinants of the share of intrafirm input purchases in both domestic as well as offshore purchases. I will, however, continue to focus on the implications of the framework for the

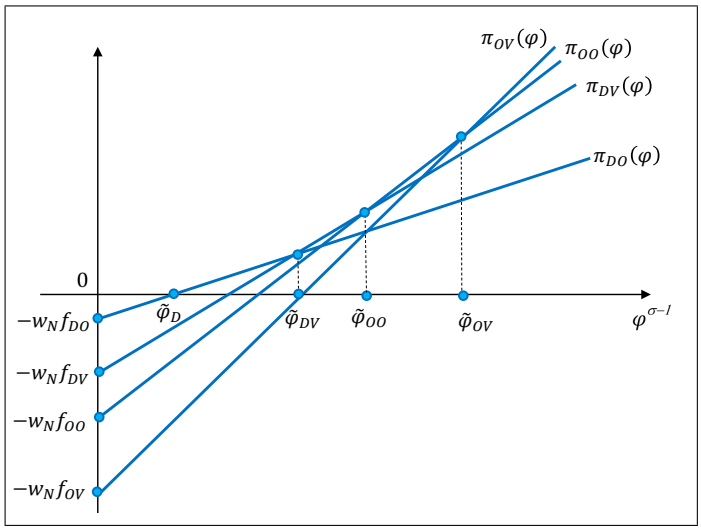

Figure 6.5: Equilibrium Sorting with Domestic Contractual Frictions

share of *offshore* input purchases that are transacted within firm boundaries. Assuming a Pareto distribution of productivity, this share continues to be given by equation (6.9), but the ratio $\tilde{\varphi}_{OV}/\tilde{\varphi}_{OO}$ is now slightly modified to

$$\frac{\tilde{\varphi}_{OV}}{\tilde{\varphi}_{OO}} = \left[\frac{f_{OV} - f_{OO}}{f_{OO} - f_{DV}} \times \frac{\Gamma_{OO} - \lambda^{1-\sigma}\left(w_N/\tau w_S\right)^{-(1-\eta)(\sigma-1)}}{\lambda^{1-\sigma} - \Gamma_{OO}}\right]^{1/(\sigma-1)}.$$

The only two differences relative to equation (6.10) is that f_{DV} replaces f_{DO} and there is an additional term $\lambda^{1-\sigma}$ in the numerator of the second term. These modifications have little bearing on the comparative statics discussed above and, in particular, those related to the effect of the 'deep' parameters in (6.11) remain unaltered.

When governance costs λ or the fixed costs of domestic integration are sufficiently large, it may be the case that no firm in the industry finds it optimal to vertically integrate domestic suppliers of inputs, and the equilibrium sorting pattern in the industry will be as in Figure 6.4 above. Notice, however, that domestic outsourcing now entails positive transaction costs captured by a contractual efficiency level $\Gamma_{OD} < 1$. In such a case, the ratio of thresholds $\tilde{\varphi}_{OV}/\tilde{\varphi}_{OO}$ in (6.10) now includes an additional term Γ_{OD} in the numerator of the second term:

$$\frac{\tilde{\varphi}_{OV}}{\tilde{\varphi}_{OO}} = \left[\frac{f_{OV} - f_{OO}}{f_{OO} - f_{DO}} \times \frac{\Gamma_{OO} - \Gamma_{OD}\left(w_N/\tau w_S\right)^{-(1-\eta)(\sigma-1)}}{\lambda^{1-\sigma} - \Gamma_{OO}}\right]^{1/(\sigma-1)}. \quad (6.12)$$

As shown in Chapter 5, however, and summarized in particular in Table 5.5, the deep parameters of the model affect the term Γ_{OO} and the ratio Γ_{OO}/Γ_{OD} in the same exact fashion, and thus the comparative statics emerging from this equilibrium will again coincide with the predicted signs summarized in (6.11).

Alternative Ranking of Fixed Costs

Up to this point, we have computed intrafirm trade shares under the assumption that the ranking of fixed costs is given by the inequalities in (6.1). Indeed, I believe it is natural to assume that (i) fixing the ownership structure, offshoring is associated with higher fixed costs than domestic sourcing, and (ii) fixing the location of input production, vertical integration is associated with higher fixed costs than outsourcing. It is less clear, however, that the fixed costs of offshore outsourcing are necessarily higher than those of domestic integration. One may wonder how our results would change if instead one assumed that the ranking of fixed costs was as follows

$$f_{OV} > f_{DV} > f_{OO} > f_{DO}.$$

Notice that in such a case, the only type of equilibria in which all organizational modes are chosen by some firms in the industry is of the type depicted in Figure 6.6. As is clear from the graph, the most productive firms continue to engage in foreign integration and intrafirm trade, while the least productive active firms make use of domestic outsourcing. The main novel feature of this equilibrium is that firms engaging in domestic integration now feature higher productivity levels than those outsourcing abroad. How important is this distinction for the comparative statics discussed above?

In this new type of equilibrium, the share Sh_{i-f} of intrafirm imported inputs over the total imported input purchases is now given by

$$Sh_{i-f} = \frac{\lambda^{1-\sigma} \int_{\tilde{\varphi}_{OV}}^{\infty} \varphi^{\sigma-1} dG(\varphi)}{\Gamma_{OO} \int_{\tilde{\varphi}_{OO}}^{\tilde{\varphi}_{DV}} \varphi^{\sigma-1} dG(\varphi) + \lambda^{1-\sigma} \int_{\tilde{\varphi}_{OV}}^{\infty} \varphi^{\sigma-1} dG(\varphi)},$$

where notice that, relative to the formula for the share in (6.6), the upper limit of the first integral of the denominator is now $\tilde{\varphi}_{DV}$ rather than $\tilde{\varphi}_{OV}$. The definitions of the thresholds $\tilde{\varphi}_{OO}$ and $\tilde{\varphi}_{OV}$ are also different from those under the previous sorting in Figure 6.5. Despite these differences, in the Theoretical Appendix I show that, when firm productivity φ is drawn from a Pareto distribution, this variant of the model continues to deliver the *exact* same comparative statics as those summarized in (6.11). Conversely,

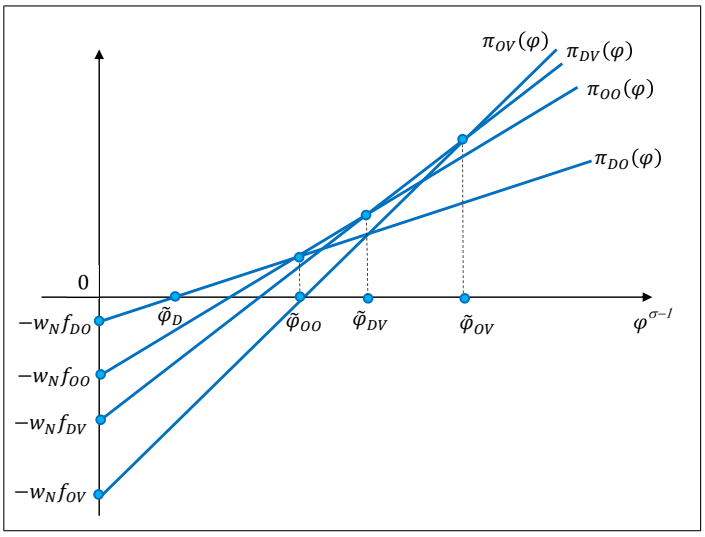

Figure 6.6: Equilibrium Sorting with an Alternative Ranking of Fixed Costs

the results in the property-rights model developed in the next chapter will prove to be a bit more sensitive to the assumed ranking of fixed costs across organizational forms, so with that in mind, in Chapter 8 I will study which ranking is most consistent with available empirical evidence.

An Overview of Other Applications of the Transaction-Cost Approach

The transaction-cost model of the internalization decision of multinational firms I have developed above does not originate in any specific paper in the literature. Nevertheless, it borrows heavily from previous work in this area and thus it is pertinent to briefly outline other contributions to the transaction-cost literature in international trade. My goal is not to offer an encyclopedic overview of the literature but rather to highlight some key aspects of the international organization of production that have been studied in the literature but from which I have abstracted in my discussion above.

An early application of the transaction-cost approach to international economics is offered by Ethier (1986). In his seminal paper, Ethier argues that the main difference between transacting within the boundaries of multinational firms and transacting at arm's length is that, in the latter case, headquarters cannot offer quality-contingent contracts to downstream producers or distributors (rather than the upstream suppliers in the models developed above). As a result, headquarters cannot always devise a contract

that ensures ex-post efficiency and extracts all surplus from their contracting partners. In those situations, the headquarters may be better off integrating the downstream producer. Interestingly, Ethier (1986) finds that integration is more attractive when differences in production costs across countries are small, which resonates with our result above that the share of intrafirm trade is lower the farther is $w_N/\tau w_S$ from one.[9]

The work of Grossman and Helpman (2002) is much more closely related to the transaction-cost model developed above. In fact, their framework is an important source of inspiration for the benchmark global sourcing model studied in this book. Grossman and Helpman (2002) offer an alternative general-equilibrium framework that also emphasizes a trade off between hold-up inefficiencies and governance costs. In their framework, suppliers undertake relationship-specific investments that enhance the value of a good sold by a final-good producer. Their framework is simpler as they focused on a closed-economy model with no producer heterogeneity and with no headquarter investments (or $\eta = 0$ in the model above). The Grossman and Helpman (2002) setup was later extended by the authors to open-economy environments with cross-country variation in the degree of contractibility of inputs in Grossman and Helpman (2003, 2005).

It is important to emphasize that the Grossman and Helpman framework is richer than the one developed in this book as it includes search frictions and an endogenous choice of the customization of inputs. More specifically, in our transaction-cost model, we have assumed that when a firm chooses outsourcing instead of integration, it can simply post a contract and pick an operator M from the set of firms applying to fulfill the order. Grossman and Helpman (2002) instead assume that the matching between stand-alone F agents and M operators is random and depends on the relative mass of each type of agents looking for matches. Furthermore, once a match is formed, agents are not allowed to exchange transfers prior to production. Other things equal, it is clear that these features will tend to reduce the attractiveness of outsourcing vis à vis integration, even holding constant supplier investments. Intuitively, search frictions and the lack of transfers inhibit the ability of final-good producers to fully capture the rents generated in production relative to a situation in which they can make take-it-or-leave-it offers to a perfectly elastic supply of M operators. Search frictions can generate much more subtle and interesting results when allowing the matching

[9] In subsequent work, Ethier and Markusen (1996) further explored the role of contractual frictions, trade costs and scale economies in shaping the horizontal integration decisions of multinational firms.

function governing the search process to feature increasing returns to scale. In such a case, Grossman and Helpman (2002) show that there may exist multiple equilibria with different organizational forms (or industry systems) applying in ex-ante identical countries or industries. Furthermore, the likelihood of an equilibrium with outsourcing is enhanced by an expansion in the market size, which increases the efficiency of matching in the presence of increasing returns to scale in the matching function.

Another distinguishing feature of the Grossman and Helpman (2002, 2003, 2005) papers is that they incorporate an organizational decision related to the degree to which suppliers customize their intermediate products to their intended buyers. This is related to the parameter ϵ in our global sourcing model, but Grossman and Helpman study the endogenous determination of such level of specificity by trading off the productivity gains associated with a higher level of customization with the increased contractual frictions associated with these. In that respect, their framework shares features with those in Qiu and Spencer (2002) and Chen and Feenstra (2008) (see also Spencer, 2005).

The papers by Grossman and Helpman were actually preceded by a noteworthy paper by McLaren (2000), who also developed a framework that allowed for interdependencies in the organizational decisions of firms within an industry. Rather than introducing search/congestion externalities of the type in the Grossman and Helpman frameworks, McLaren (2000) focused on the implications of market thickness for the ex-post division of surplus between F and M agents. In his framework, the thicker is the market for inputs, the larger is the ex-post payoff obtained by M producers since they are in a better position to find an alternative buyer for their customized inputs. Thicker downstream markets thereby alleviate hold-up inefficiencies. Crucially, however, the thickness of the market for inputs depends in turn on the extent to which final-good producers rely on outsourcing versus integration in their procurement decisions, since only firms engaged in outsourcing populate that market. McLaren (2000) demonstrates that this setup too generates the possibility of multiple equilibria and shows that trade opening, by thickening the market for inputs, may lead to a worldwide move towards more disintegrated industrial systems, thus increasing world welfare and leading to gains from trade quite different from those emphasized in traditional trade theory.

The vast majority of applications of the transaction-cost approach in the international trade literature have modeled the transaction costs emanating from incomplete contracts as taking the form of ex-ante inefficiencies associated with the suboptimal provision of relationship-specific investments. As noted earlier, Williamson's own work, particularly Williamson (1975), instead

emphasized the role of ex-post inefficiencies related to the inefficient termination or execution of contracts. For instance, in an uncertain environment, parties might need to adapt to ex-post situations that were not foreseen in an initial (incomplete) contract. In those situations, transaction-cost theory would posit that adaptation can be carried out more efficiently within firm boundaries than at arm's-length due to the useful role of authority in reducing the scope for opportunism and costly renegotiation. Costinot, Oldenski and Rauch (2011) develop a simple theory of multinational firm boundaries based on these ideas and show, both theoretically as well as empirically, that the propensity to integrate foreign suppliers is lower in more routine sectors.