

Outsourcing versus vertical integration: Ethier–Markusen meets the property-rights approach

James R. Markusen* and Yiqing Xie†

Early analyses of direct investment versus outsourcing focused on the existence of knowledge-based assets, knowledge being non-rivalled and non-excludable. Ethier was the first to formally model the consequences of non-excludability for the vertical integration versus outsourcing decision. Later authors took a different approach, modeling physical capital as fully excludable but relationship-specific. This paper further develops a model with both non-excludable knowledge capital and fully excludable physical capital. Results show that vertical integration tends to be chosen when (a) the technology is relatively knowledge intensive and/or when (b) knowledge and physical capital are strong complements.

Key words vertical integration, outsourcing, knowledge capital, excludability, property rights, knowledge intensive.

JEL classification F12, F23

Accepted 1 September 2013

1 Introduction

International business scholars have long been interested in the decisions of firms to produce abroad and their decisions on the “mode” of foreign production. Primary among these modes are foreign direct investment (FDI), in which the foreign producer is vertically integrated into the parent firm, and a contract or licensing arrangement with an independent foreign producer, now commonly referred to as outsourcing (OS). The foreign production and mode choice problems are really quite separate: the decision to produce abroad is a *location* decision, while the decision to keep production in-house is an *ownership* decision. The international business literature is very rich indeed, but it generally lacks formal models and often has to rely on corporate anecdotes, given the absence of detailed firm-level data (the absence of OS data persists today).¹

* Department of Economics, University of Colorado at Boulder, USA. Email: james.markusen@colorado.edu

† Department of World Economy and Institute of World Economy, School of Economics, Fudan University, Shanghai, China.

This paper has been prepared for a special issue of the *International Journal of Economic Theory* in honor of Professor Wilfred Ethier. Markusen and Ethier further developed Ethier’s earlier work when they both visited the Institute for Advanced Studies at the Hebrew University for 6 months in 1989, under a group grant obtained by Elhanan Helpman. We are/were grateful for this opportunity.

¹ Dunning (1977) is an early treatment that examines the location and outsourcing issues, though he referred to “internalization” to mean in-house production, the opposite of outsourcing. Caves (2007) provides a good summary and evaluation of international business research.

One thing that comes up repeatedly in the international business literature is that multinational firms (with owned foreign subsidiaries) are intensive in what we now broadly refer to as knowledge-based assets or capital. This is inferred from looking at information such as R&D to sales ratios, the white-collar share of the workforce, product newness and complexity, reputations, product differentiation measures, and so forth. Physical capital does not tend to be strongly correlated with multinationality.

Knowledge tends to have two characteristics associated with public goods: it is at least partially non-rival and non-excludable, though we have not seen these terms used in the international business literature we have read. The first of these properties creates an opportunity for a knowledge-intensive firm whereas the second creates a difficulty. The non-rival property means that, once created, knowledge can be applied in many production facilities without reducing its value in any single facility. This in turn gives firms a strong incentive to expand to multiple plants abroad, using the same knowledge at zero additional cost. But the non-excludability property means that a foreign firm, whether an owned subsidiary or an arm's-length contractor, cannot be prevented from learning the knowledge and either demanding a high return or going it alone.

Two early papers (Helpman 1984; Markusen 1984) noted the role of non-rival knowledge assets in determining the location decision, connecting the theory of the multinational with international trade theory. That is, the *non-rival* property is connected to the *location* decision. But the question of the mode of foreign production in these papers is simply assumed to be an owned subsidiary (vertical integration).

To the best of our knowledge, the first formal model that attacked the non-excludability problem was Ethier (1986). His general idea was that there is a difficulty trying to contract with a foreign agent: the firm needs to reveal what knowledge it has in the negotiation, but once it does so the (potential) agent absorbs some of that knowledge costlessly.

The basic consideration working against the arm's-length [outsourcing] alternative is the fact that in order to sell its information for its full value, the firm must convincingly indicate what it has to sell, thereby losing, at least in part, its monopoly advantage (Ethier 1986, p. 808).

The potential agent or contractor needs to see the blueprints, but once he does, the firm's advantage is partially lost: the agent can walk away but retaining the knowledge. *Non-excludability* is connected to the *ownership* decision. At least in the field of international economics, there seemed to be little interest in this topic in the next decade, with the exception of a paper by Horstmann and Markusen (1987) which focused on the existence and dissipation of reputation capital.

Ethier returned to this topic in a paper with Markusen (Ethier and Markusen 1996). There is a multi-period model with a new good every two periods. A foreign agent or firm can learn the knowledge in the first period, and then defect to start a rival firm in the second period, so the multinational must share rents with the foreign agent in the second period if the contract is to be preserved. This approach was further developed in Markusen (2001), with closely related papers by Fosfuri *et al.* (2001) and Glass and Saggi (2002) appearing at the same time. All of these papers focus on the non-excludability properties of knowledge-based assets.

A couple of years later, a number of papers appeared which took a quite different and complementary approach to the FDI versus OS decision. These included Antràs (2003, 2005), Antràs and Helpman (2004), and Grossman and Helpman (2002, 2004). These papers begin with an extreme polar assumption: capital assets are fully excludable physical capital. This assumption is combined with Williamson's (1975) notion of relation-specific investments, developed more formally by Grossman and Hart (1986) and Hart and Moore (1990). The two parties to an investment project, firm and foreign agent, must make up front investment in capital goods, for example, that have little or no

use outside their relationship. In the absence of complete and enforceable contracts, this leads to a holdup problem once the investments have been made. The firm can own the subsidiary, which allows it to avoid holdup, but this has the disadvantage of leading to agent (manager) shirking or lack of effort. The advantage of OS, where the agent owns the capital, is more high-powered incentives to the agent. Thus the FDI–OS decision trades off high-powered incentives versus holdup. Interesting empirical evidence is found in Feenstra and Hanson (2005).

As far as we know, the only formal attempt to integrate these two approaches is in Chen *et al.* (2012). Their model has both excludable physical capital and non-excludable knowledge capital. The multinational can own the physical capital (FDI), which reduces the holdup power that the agent or manager has once the latter has absorbed the knowledge capital after one period. But owning the physical capital removes the incentive for the manager to maintain that capital, the equivalent of the shirking in the property-rights approach. Outsourcing gives the agent the correct incentives, but leads to second-period holdup as in Ethier and Markusen (1996) and Markusen (2001). The response to this holdup problem is that the firm restricts the knowledge transfer in the first period. Thus FDI suffers from weak incentives for the manager, but OS leads to an inefficiently small knowledge transfer.

In this paper written in honor of Bill Ethier, we continue down the path set by Ethier and Markusen (1996) and Chen *et al.* (2012). Specifically, we dig more deeply into the issues of what sorts of firm technologies lead to FDI versus OS. Though there is some presumption in Chen *et al.* that more knowledge-intensive firms will choose FDI, that finding is unclear and relies on several assumptions which lack good economic intuition. Here, we add more structure where that paper ends. That structure clarifies the role of relative factor intensities, physical versus knowledge intensive, and also the role of absolute capital requirements.

We show in a base case that relatively knowledge-intensive firms choose FDI, while relatively physical-capital-intensive firms choose OS. The intuition lies in the tradeoff of the disadvantages of each mode. As just noted, the disadvantage of OS, due to holdup, is an inefficient transfer of knowledge. This becomes more costly relative to total returns as the firm becomes more knowledge intensive. Owning the physical capital under FDI protects the knowledge capital. The disadvantage of FDI is weak incentives to maintain physical capital, and this becomes more costly as the firm becomes more physical-capital intensive. Thus the relatively knowledge-intensive firm chooses FDI. There is no obvious relationship of the mode choice to the absolute physical-capital requirements of the firm: relative intensity is what matters.

This result is far from completely general, however. A second example is based on a strong complementarity between physical and knowledge capital. This case can produce a result that firms which have higher absolute physical-capital requirements but also relatively higher physical-capital intensity may choose FDI. As physical capital requirements grow, the relative knowledge intensity falls. But the loss in rents due to inefficient knowledge transfer under OS is larger for a large, physical-capital-intensive firm than for a small one. The idea here is that even if a large firm has only a small knowledge stock, that knowledge may be vital to the productivity of the physical capital. Specifically, the inefficient knowledge transfer reduces rents in proportion to the physical capital requirements, not in proportion to the knowledge intensity. While we acknowledge that this case is quite special, it emphasizes that complementarity can trump factor intensity.

In the model, knowledge capital is divided into a portion that is protected (codifiable and contractible) and a portion that is not protected and not contractible (tacit knowledge). In all cases examined, an increase in the ability of the firm to protect its knowledge capital will lead some portion of the firms to switch from FDI to OS.

The final section of the paper tries to tie the theory to empirical evidence. In particular, the model suggests that FDI should be preferred by firms with high values of Tobin's Q ; roughly the market

value of the firm divided by the book value of capital. The idea is that market value reflects the value of intangible knowledge capital, while the book value tends to undervalue these assets if at all. Thus a high Q corresponds to firms that are knowledge intensive and those firms are more likely to choose FDI than OS. A recent paper by Jinji *et al.* (2012) gives strong support for this relationship.

2 The hybrid knowledge and physical-capital model

As just noted, the basic model is close to Chen *et al.* (2012), and we follow that paper in outlining the specification in the next few paragraphs. Then we make some changes in the model details, altering some aspects in order to make the model suitable for the somewhat different questions asked here and simplifying some details where the complications needed for the direction taken in Chen *et al.* (2012) are not needed here. Similarly, we will simply refer to results proved in Chen *et al.* (2012) and not re-derive them, but will try to provide the intuition for those results.

A multinational firm (MNE) in the North, denoted by M , plans to produce a product in the South, owing to cost advantages of manufacturing there. There are two periods of production, $t = 1, 2$, and there is no discounting. Production in the South requires the services of a local agent/manager, denoted as A , and two types of asset services: physical capital and knowledge capital/intellectual property. For production to occur, a physical capital investment in the amount K must be undertaken at the beginning of $t = 1$. Within any given industry, the value of K is fixed and exogenous. Either M or A can undertake this investment and, by so doing, becomes the “owner” of the physical capital. The price of K is one. By contrast, only M owns the knowledge capital initially; the amount of knowledge capital that M owns is a fixed and exogenous amount equal to S .

At the beginning of $t = 1$, M makes a once-for-all choice between two possible organizational forms, FDI and OS. With FDI, M acquires (and owns) the physical capital used for production in the South and employs A under a sequence of one-period employment contracts to manage a production process utilizing M 's capital. A 's hiring occurs at the beginning of $t = 1$ and A 's employment contract is renegotiated at the beginning of $t = 2$. M also decides each period how much knowledge capital to transfer to A to be utilized in production.

With OS, A acquires (and owns) the physical capital, though it may be partially or entirely financed by M , as discussed shortly. Capital acquisition again occurs at the beginning of $t = 1$. M signs a one-period licensing contract with A that licenses an amount $s_1 \leq S$ of M 's knowledge capital to A for use in production at $t = 1$. This licensing agreement is renegotiated at the beginning of $t = 2$, and an amount s_2 is transferred at $t = 2$.

As in Grossman and Hart (1986), ownership of physical capital bestows control rights on the owner. Specifically, the owner can decide on the uses to which the capital is put and can exclude access to the capital for any other uses at any time. That having been said, effort from the agent is required to improve or maintain the efficiency (usefulness) of the physical capital, which will lead to the moral hazard problem that is a key element of the property-rights approach.

Knowledge (or intellectual) capital, by contrast, does not have the same excludability properties. The owner of knowledge capital may not be able to control to the same extent the uses to which the capital is put and to capture the returns that the knowledge capital generates. In essence, property rights to knowledge capital are harder to define and protect than is the case for physical capital. To capture these features of knowledge capital, we assume that, at $t = 1$, there exists a fraction $\sigma < 1$ of S that is “explicit knowledge” in the sense that its use can be defined in a $t = 1$ licensing agreement. In this sense, the fraction σ of M 's knowledge capital can be “owned” in the same way as physical capital is owned. A fraction $1 - \sigma$ of S is “tacit knowledge” and is not contractible at $t = 1$ in the

sense that its current and future uses cannot be controlled by M in a $t = 1$ licensing agreement. The value of σ is assumed to be exogenous to the firm and can be thought of as capturing either characteristics of the knowledge capital or a characteristic of the legal regime of the country in which M is contracting. While M may not transfer all of S at $t = 1$, there is no further holdup issue at $t = 2$ (time ends).

Note that this set of assumptions allows a simple characterization of the environment in which M and A operate. Specifically, the values of K , S , and σ completely define the environment—characteristics of the relevant industry/product—under consideration. The values of K and S give the physical relative to knowledge capital intensity of the industry, while σ gives features of the appropriability of M 's knowledge asset. Subsequent analysis will consider how variation in the economic environment—variation in (K, S, σ) —affects the observed pattern of OS and FDI.

We will try to be brief in developing the model, and not derive proofs that are shown in detail in Chen *et al.* (2012), but will try to provide the intuition as noted above. Let the gross surplus generated by the investment project over the two time periods be given by

$$U(K, s_1) + U(K, s_2), \quad s_1, s_2 \leq S, \quad (1)$$

where s_1 and s_2 are the amounts of knowledge capital transferred in periods 1 and 2. U is increasing in both arguments. It is also assumed that $U(K, 0) = 0$, so the agent has no alternative use for the capital if there is no knowledge transferred.

Let K denote the capital cost of the project as noted above. The agent needs to exert effort e , equal to zero or one, to maintain the capital, and effort is non-observable and non-contractible. It is assumed that if the agent exerts effort ($e = 1$) to use, care for, and maintain the capital properly then there is no depreciation or added cost in the second period. If the agent exerts no effort ($e = 0$), then a repair or replacement charge of δK must be incurred in the second period.

We assume that the party not owning capital at $t = 1$ is not able to raise the amount K sufficiently quickly to produce alone at $t = 2$ and assume that the knowledge is of no value to A without the physical capital. Thus owning the capital under FDI protects M 's knowledge capital at $t = 2$. But under OS where A owns the physical capital, these assumptions produce a (potential) specific relationship between M and A at $t = 2$. In this case, the $t = 2$ surplus is allocated based on the Nash bargaining solution.

There is a perfectly elastic supply of agents with opportunity cost $W \geq 0$ in each period. Both A and M are risk neutral and make choices to maximize expected income. We assume that, at $t = 1$, M has access to perfect capital markets, while A has only limited ability to borrow against future income to finance capital K and consumption W (i.e., W is both an opportunity cost and a subsistence requirement). Specifically, under outsourcing A can only cover a fraction $\rho \in (0, 1)$ of $t = 1$ costs ($K + W$) via the capital market. The remainder must be covered out of $t = 1$ revenues. As will be seen later, this inability on A 's part to borrow against future income limits M 's ability to extract $t = 2$ surplus from A via the initial licensing agreement at $t = 1$. It also forces M to bear some of the costs of A 's capital investment. Assumption 1 of Chen *et al.* (2012) gives some parameter restrictions that assure participation.

At the beginning of $t = 1$, M chooses between FDI and OS. If M chooses FDI, then M offers A an employment contract involving payment to A of W ; if M chooses outsourcing, M and A negotiate a contract involving a transfer to A of knowledge capital, s_1 , and a licensing payment from A to M of l_1 . In either case, should A accept the contract, A chooses an effort level e and M chooses the level of knowledge capital to use (this level is determined by the licensing contract under OS). Gross surplus $U(K, s_1)$ is then realized and payments are made.

At the beginning of $t = 2$, M offers A a second employment contract involving a payment W under FDI. If A accepts, M again chooses a level of knowledge capital utilization for A . Under outsourcing, M and A negotiate, via Nash bargaining, a second contract involving a transfer to A of intellectual capital, s_2 , and a licensing payment from A to M of l_2 . Gross surplus $U(K, s_2)$ is then realized and payments are made. In all cases, the equilibrium levels of e, s_1, s_2 are the result of subgame perfect Nash equilibrium strategy choices by A and M , and U is sufficiently large relative to K and W that in equilibrium it is profitable for M to choose either FDI or OS.

3 Equilibrium under FDI

Begin by analyzing the contract at $t = 2$. Since M owns the physical asset under FDI, M can control the use of K at $t = 2$. Thus M can prevent A from using S for purposes other than M 's project in $t = 2$. Consequently, A 's outside option is simply W and, since $U(K, s_2)$ is increasing in s_2 , M chooses knowledge capital transfer $s_2 = S$. As a result, the contract that M offers A involves a payment of W to A .

At $t = 1$, M will optimally choose $s_1 = S$, since $U(K, s_1)$ is increasing in s_1 and the payoff at $t = 2$ is independent of the value of s_1 . Further, since effort is costly for A and A 's compensation at $t = 2$ is independent of A 's effort, A will choose zero effort. Recognizing this, M offers a payment of W to A , knowing that an expenditure δK will have to be made at $t = 2$ to repair or replace depreciated capital. Profits for M (subscript m) under FDI (superscript f) in periods 2 and 1 will be

$$\pi_{m2}^f = U(K, S) - \delta K - W, \quad \pi_{m1}^f = U(K, S) - K - W, \quad (2)$$

so M 's two-period payoff from FDI will be

$$\Pi_m^f = 2U(K, S) - (1 + \delta)K - 2W. \quad (3)$$

4 Equilibrium under outsourcing

Under outsourcing, A will own the physical capital but may be able to pay for only a portion of it or none at all, as discussed above. Owning K gives A an incentive to maintain it, and we assume that the benefits of $e = 1$ (not incurring δK in the second period) outweigh the cost of effort. Owning K also gives A second-period holdup power. Here we will make a simplification that departs from the more complex treatment in Chen *et al.* (2012): a positive outside option for A in owning K in second-period bargaining only derives from the portion of the tacit knowledge transferred at $t = 1$. The codifiable portion σS is contractible and cannot be used independently by A in the event of a bargaining breakdown. Thus if only the codifiable part is transferred in period 1, A 's return in the event of a breakdown is $U(K, 0) = 0$ as assumed above. A may return to wage work in period 2 in the event of a breakdown.

Let us just assert a result proved in Chen *et al.* (2012) and not re-derive the proof, since we think the intuition is quite apparent:

$$s_1 = \sigma S, \quad s_2 = S. \quad (4)$$

In the first period, the multinational only transfers the codifiable portion of knowledge capital. There are two reinforcing reasons for this. First, if M transfers more than σS , that added amount is

non-contractible and M cannot charge for it. Second, transferring more than σS will give A more holdup power at $t = 2$. This gives A a larger share of the surplus at $t = 2$ and lower profits for M . Denoting the first-period license fee under outsourcing as l_1^o and A 's first-period profits as π_{a1}^o , we have

$$l_1^o = U(K, \sigma S), \quad \pi_{a1}^o = -\rho(K + W), \quad (5)$$

$$\pi_{m1}^o = U(K, \sigma S) - (1 - \rho)(K + W), \quad (6)$$

so that the sum of the two parties' earnings add to the total available surplus:

$$\pi_{m1}^o + \pi_{a1}^o = U(K, \sigma S) - (K + W). \quad (7)$$

Here we see the value of having A finance a portion $\rho > 0$ of first-period costs $(K + W)$. M can use this to "claw back" rents that A is going to receive in the second period due to holdup power.

Assume that in the event of a bargaining breakdown, M has no outside option while A can resort to wage work. Given the result that M only transfers σS in the first period, we have the result that A has no additional outside option from owning K other than wage work earning W . With M transferring the whole knowledge stock in the second period, Nash bargaining then maximizes the product

$$\max \text{ w.r.t. } \mu \quad (\mu U(K, S))((1 - \mu)U(K, S) - W), \quad (8)$$

where μ gives M 's equilibrium share of the total surplus U . The Nash bargaining solution then gives second-period payoff to M and A equal to

$$\pi_{a2}^o = W + (U(K, S) - W)/2 = U(K, S)/2 + W/2, \quad (9)$$

$$\pi_{m2}^o = U(K, S)/2 - W/2. \quad (10)$$

M 's profits under OS are then (6) + (10):

$$\Pi_m^o = U(K, \sigma S) + U(K, S)/2 - W/2 - (1 - \rho)(K + W). \quad (11)$$

5 Optimal mode choice

Both mode choices suffer from some joint inefficiency. FDI suffers from the moral hazard problem, reflected by the term δK in (3). Outsourcing suffers from the problem of inefficient knowledge transfer in period 1: $U(K, \sigma S) < U(K, S)$. In addition to these overall inefficiencies, OS transfers some rent to the agent through holdup. The relative attractiveness of FDI is given by subtracting (11) from (3).

$$\begin{aligned} \Pi_m^f - \Pi_m^o &= 2U(K, S) - (1 + \delta)K - 2W - U(K, \sigma S) \\ &\quad - U(K, S)/2 + W/2 + (1 - \rho)(K + W) \end{aligned} \quad (12)$$

We can rearrange (12) as follows:

$$\Pi_m^f - \Pi_m^o = [U(K, S) - U(K, \sigma S)] + [U(K, S)/2 - W/2] - \delta K - \rho(K + W). \quad (13)$$

FDI is chosen when (13) is positive, and OS is chosen when (13) is negative.

The first term in brackets in (13) is the gain under FDI from the more efficient transfer of knowledge (+). The second term in brackets is the gain to M under FDI from avoiding holdup that occurs under outsourcing (+). The third term, $-\delta K$, is the cost under FDI of moral hazard, the non-maintenance of capital (-). The fourth term, $-\rho(K + W)$, is the benefit under OS of getting A to finance first-period costs (-). Note that if there was no moral hazard and the agent was unable to finance any of the first-period costs, then FDI would be unambiguously preferred.

All of this seems pretty straightforward, and so now we turn to the question of what sort of industries or firms will choose one mode versus the other. The obvious hypothesis is that relatively S -intensive industries should choose FDI: the first bracketed term in (13) will be larger and the two final terms, both negative, should be smaller. Conversely, for a relatively physical-capital-intensive industry, the first (positive) term may be small and the final two (negative) terms should be large. In Chen *et al.* (2012), we put some rather abstract restrictions on the U function such that this would be true.

However, we were never very satisfied with this, and here is where we want to pick up where the earlier paper left off. We want to do a more satisfying job of characterizing what sorts of technologies lead to one choice or the other.

6 Characteristics of technologies that lead to FDI versus outsourcing

The difficulties with Equation (13) are apparent after a little reflection. The first difficulty is that the values of the $U(K, s_i)$ relative to the cost terms will depend on the relative factor intensities of the industry, but probably also on the absolute factor requirements. It looks as though an absolutely high K industry will choose OS, but we do not have a good idea if an industry with absolutely high K requirements is relatively S - or K -intensive. Boeing and Intel may have huge K requirements, but they may nevertheless be relative S -intensive. What we would like for empirical analysis is the result that relatively S -intensive industries choose FDI independently of their size, and it is not at all obvious what restrictions this will place on the technology in (13).

The second thing that may matter is the loss that inefficient knowledge transfer imposes in relation to the relative factor intensity in U . Perhaps an industry is relatively physical-capital intensive, but the small amount of knowledge capital is vital, and inefficient knowledge transfer due to holdup imposes a big loss: the first bracketed term in (13) is large. We can probably think of this as the question of whether K and S are complements or substitutes. If they are substitutes, then we may conjecture that a relatively K -intensive industry will choose OS. But if they are complements, then even a relatively K -intensive industry may choose FDI to minimize the losses of inefficient knowledge transfer.

In order to examine these issues more clearly, we will set $W = 0$ for the remainder of this section. It is not clear how W (perhaps the cost of the management team) should be related to the size of the firm: perhaps for example, it is simply linear in U . To focus on the roles of K and S , we simply set it to zero. Equation (13) then reduces to

$$\Pi_m^f - \Pi_m^o = \left[\frac{3}{2} U(K, S) - U(K, \sigma S) \right] - (\delta + \rho)K. \quad (14)$$

First, it seems that any reasonable assumptions about the effect of S on U will imply that (14) is increasing in S , holding K constant. Thus more S -intensive industries will tend to prefer FDI. The effect of increasing K , holding S constant, is also relatively clear: if K has a diminishing marginal

product in U , then the bracketed term in (14) will be strictly concave in K while costs are rising linearly in K : more K -intensive firms will prefer outsourcing, holding S constant.

But this seems a somewhat unsatisfactory way of thinking since, as we move across a line of industries, S and K are likely to increase together. Industries with larger K may or may not be more K -intensive. Suppose we line up all firms or industries according to their (absolute) K requirements. Allow this to be a continuum of firms/industries, noting that K and S are not choice variables of the firms, they are varying across firms but exogenous to the individual firms. Taking the total derivative of U , we have

$$\frac{dU(K, S)}{dK} = \frac{\partial U(K, S)}{\partial K} + \frac{\partial U(K, S)}{\partial S} \left[\frac{dS}{dK} \right]. \quad (15)$$

The last bracketed term reflects how the relative factor intensity changes as we move to a firm with a higher absolute K requirement. Equation (15) can be written in proportional change terms as follows:

$$\frac{K}{U} \frac{dU}{dK} = \frac{K}{U} \frac{\partial U}{\partial K} + \frac{S}{U} \frac{\partial U}{\partial S} \left[\frac{K}{S} \frac{dS}{dK} \right]. \quad (16)$$

Suppose that U has “constant returns”; that is, as we move to a firm with a proportionally higher K requirement and S “endowment”, the surplus of that firm increases in the same proportion. If, in addition, S increases in strict proportion to K as we move across the continuum, then firms with higher absolute K requirements will have a higher U in strict proportion to K . Equation (14) will be linear in K , and will be either positive or negative for all firms in the continuum.

This is not realistic, but it does help clarify why (14) does not offer us a very clear empirical prediction about a firm’s choice of FDI versus OS in relation to its absolute and relative factor intensities. Let us therefore sharpen things by looking at a special case. First, assume that a firm’s surplus has “constant returns” in its K and S requirements; specifically, assume that

$$U = K^\alpha S^{1-\alpha} \quad (\text{surplus linear in industry “size”}). \quad (17)$$

Second, assume a systematic relationship between a firm’s absolute K requirement and its S characteristic (its existing S endowment):

$$S = K^\beta, \quad \begin{aligned} \beta < 1: & \text{ bigger firms, relatively more } K\text{-intensive,} \\ \beta > 1: & \text{ bigger firms, relatively more } S\text{-intensive.} \end{aligned} \quad (18)$$

Substituting (18) into (17), we get

$$U(K, S) = K^\alpha S^{1-\alpha} = K^\alpha K^{\beta(1-\alpha)} = K^\gamma, \quad \gamma \equiv \alpha + \beta(1 - \alpha), \quad (19)$$

$$U(K, \sigma S) = \sigma^{1-\alpha} K^\alpha S^{1-\alpha} = \sigma^{1-\alpha} K^\alpha K^{\beta(1-\alpha)} = \sigma^{1-\alpha} K^\gamma. \quad (20)$$

The relative advantage of FDI in (14) then becomes

$$\begin{aligned} \Pi_m^f - \Pi_m^o &= [(3/2)U(K, S) - U(K, \sigma S)] - (\delta + \rho)K \\ &= K^\gamma [3/2 - \sigma^{1-\alpha}] - (\delta + \rho)K, \end{aligned} \quad (21)$$

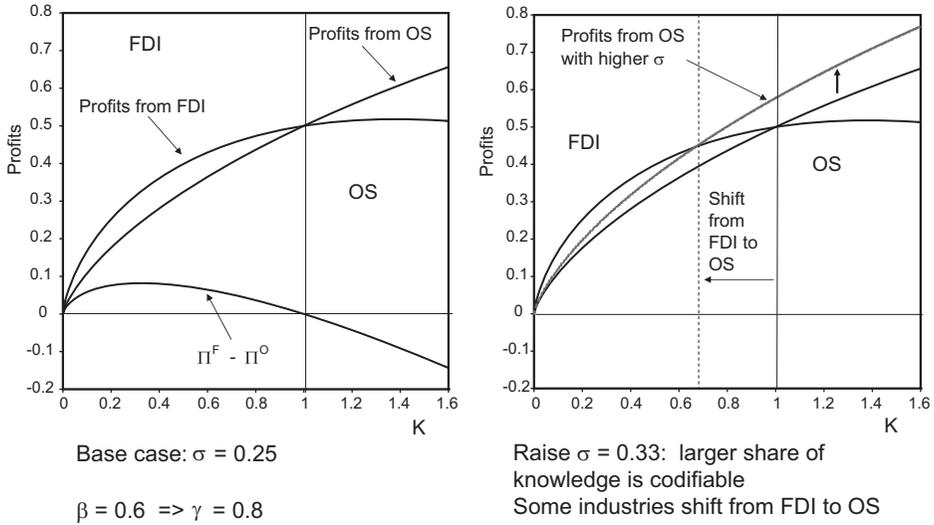


Figure 1 Case 1: $\beta < 1$. Higher K , higher relative K/S ; industries with large K requirements are physical-capital intensive.

where

$$\gamma \equiv \alpha + \beta(1 - \alpha), \quad \gamma < 1 \text{ iff } \beta < 1.$$

Equation (21) will be concave in K if firms with absolutely higher K requirements are also relatively K -intensive (β and $\gamma < 1$), and convex in K if firms with higher K requirements are relatively S -intensive (β and $\gamma > 1$).

Figures 1 and 2 present two cases, the first with β and $\gamma < 1$, and the second with β and $\gamma > 1$. We choose a simple normalization $\alpha = \delta = \rho = 0.5$, and $\sigma = 0.25$, such that (21) is zero at $K = 1$.

In Figure 1, we choose $\beta = 0.6$, implying $\gamma = 0.8$. The results in the left-hand panel show that firms with absolutely higher K requirements choose OS. However, these firms are relatively more K -intensive (K/S rises with K). Hence we do get the result that FDI is chosen by S -intensive (low K) firms while OS is chosen by K -intensive firms. That is the result we are looking for, and it emphasizes that there is no obvious relationship between firm size and mode choices.

The right-hand panel of Figure 1 shows a simple comparative-statics experiment in which we raise the value of σ from 0.25 to 0.33: a larger share of the firm's knowledge is codifiable and contractible. This does not affect the profits from FDI, but it does raise the profits from OS by reducing the cost of holdup. This shifts the switch point between FDI and OS to the left, implying a set of firms will now switch from FDI to OS.

In Figure 2, we choose $\beta = 1.5$, implying $\gamma = 1.25$. The results in the left-hand panel show that firms with an absolutely higher K choose FDI. But in contrast to Figure 1, high K firms are now knowledge capital intensive since S/K rises with K . So in both cases 1 and 2 (Figures 1 and 2), relatively more knowledge-intensive firms choose FDI. The right-hand panel of Figure 2 shows the same comparative-statics experiment as in Figure 1: we raise the share of knowledge that is codifiable and contractible from $\sigma = 0.25$ to $\sigma = 0.33$. Although we now move the mode-choice boundary to the right, the result is qualitatively the same as in Figure 1: some firms switch from FDI to OS.

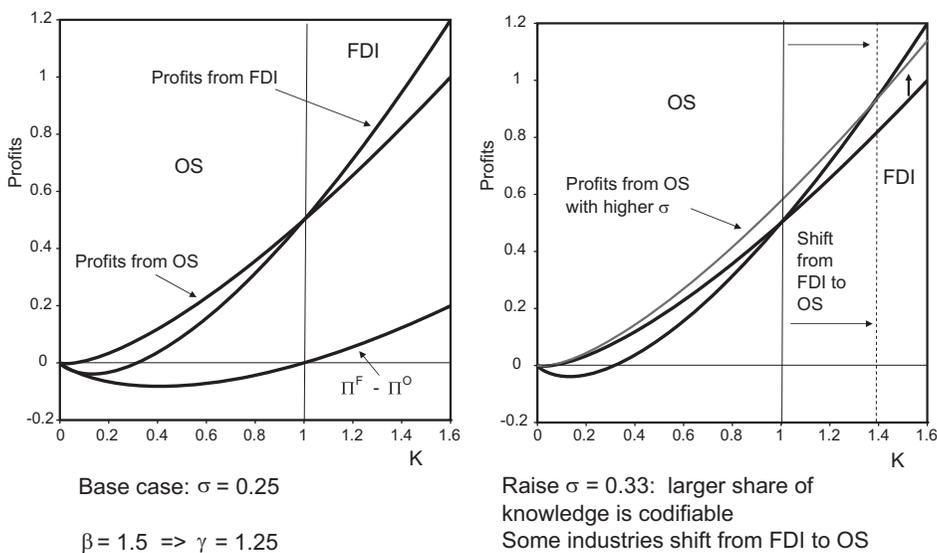


Figure 2 Case 2: $\beta > 1$. Higher K , higher relative S/K ; industries with large K requirements are knowledge intensive.

These results certainly seem intuitive within the context of the model: relatively more knowledge-intensive firms choose FDI, and higher intellectual property protection and contract enforcement will favor OS. We will have more to say about this in reviewing the empirical evidence later in the paper.

7 Strong complementarity between K and S

It is not of course clear how this simple example of technologies across firms generalizes. In fact, we can show a case that pulls the other way in this section. This occurs when there is a strong complementarity between K and S . It might be that firms with a higher K requirement are relatively less S -intensive. But simply looking at factor shares does not tell us how important S is: perhaps it is a small part of value added, but reducing it through holdup and inefficient knowledge transfer may have a very big effect on reducing output.

Here we concentrate on another very special case to make our point. Suppose that U and the S - K relationship across industries are given by

$$U = \min [K, S], \quad S = K + 1 \Rightarrow S/K = 1 + 1/K \text{ falls with } K, \tag{22}$$

such that firms with absolutely higher K are relatively more K -intensive. The relative attractiveness of FDI is then given by

$$\begin{aligned} \Pi_m^f - \Pi_m^o &= \left[\frac{3}{2} U(K, S) - U(K, \sigma S) \right] - (\delta + \rho)K \\ &= \frac{3}{2} \min [K, K + 1] - \min [K, \sigma(K + 1)] - (\delta + \rho)K. \end{aligned} \tag{23}$$

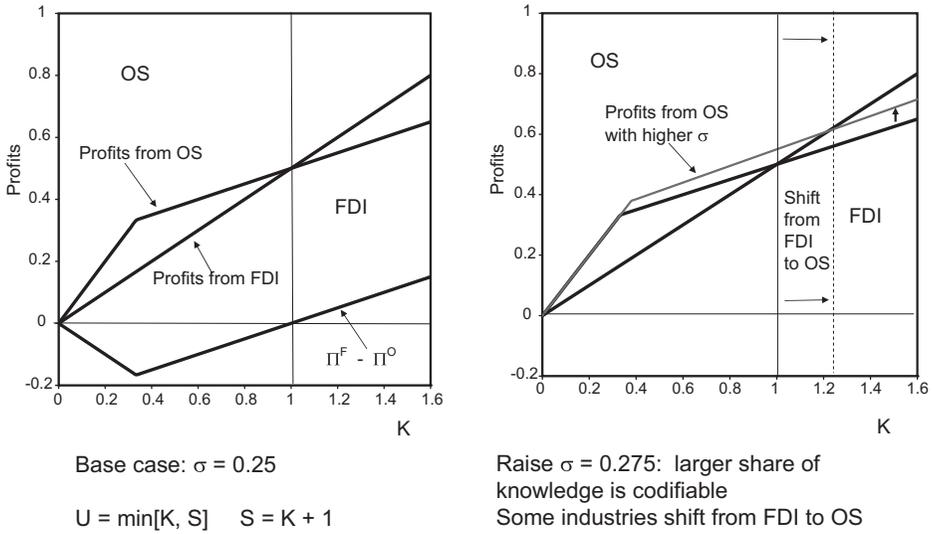


Figure 3 Case 3: K, S complements. Higher K , higher relative K/S ; industries with large K requirements are physical-capital intensive.

The first and third terms are linear in K . However, given $\sigma < 1$, the middle term has a switch point at a critical value of K :

$$K < \frac{\sigma}{1 - \sigma} \Rightarrow \min [K, \sigma(K + 1)] = K, \quad \frac{dU}{dK} = 1, \tag{24}$$

$$K > \frac{\sigma}{1 - \sigma} \Rightarrow \min [K, \sigma(K + 1)] = \sigma(K + 1), \quad \frac{dU}{dK} = \sigma < 1. \tag{25}$$

At the switch point, the additional rents earned by a firm a little further to the right on the continuum falls from 1 to σ . Holdup and inefficient knowledge transfer have a bigger impact on high K , high K/S firms.

An outcome is shown in Figure 3, using the same value of $\sigma = 0.25$ and same values for δ and ρ as in Figures 1 and 2. The switch point in the curve for profits from OS is at $1/3$ in this example. Now it is the case that high K/S firms choose FDI and knowledge-intensive firms choose OS. The reason, again, is that holdup and therefore inefficient knowledge transfer is costless for low K , high S/K firms, but is very costly for high K , high K/S firms. The right-hand panel of Figure 3 shows the effect of increasing σ from 0.25 to 0.275. While case 3 reverses the role of relative factor intensities, the increase in the share of knowledge that is codifiable and contractible results in some firms shifting from FDI to OS as in the previous section.

We surely make no claim of generality in these examples. Rather, our point is to develop some intuition as to what aspects of technologies are important in determining the mode choice between FDI and OS.² However restrictive, three things are clear from our results. First, it is straightforward to construct cases where knowledge-intensive firms tend to choose FDI and there is no necessary

² Using $U = \min[K, \sigma S]$ with $S = K^\beta$ and $\beta < 1$ generates a non-monotonic profit difference: OS is chosen at low values of K , FDI at middle values of K , and OS again at high values of K .

relationship between firm size and the mode choice. Second, this result can be reversed with strong complementarities between knowledge capital and physical capital. Holdup can then be very costly for a firm even if the knowledge-capital share of production is small. Third, increases in intellectual property protection and contract enforcement lead to more OS.

8 Model predictions and empirical evidence

The model offers a number of empirical predictions, some of which are briefly outlined in Chen *et al.* (2012). We will run through these briefly, and then spend more time on a narrow set of these, relating them to currently available empirical findings.

The model offers three predictions that relate to characteristics of firms and markets. First, the results suggest that measures of Tobin's Q should be positively related to a firm's choice of FDI over OS. Second, there is a prediction that the results may relate to product-cycle evidence and to Antràs (2005) in particular. If firms become less knowledge-capital intensive over time as technology becomes standardized, we should see firms switching from FDI to licensing or OS. There are certainly many observations in the international business literature about this, such as Taiwanese entrepreneurs buying out their former owners and becoming suppliers to the latter. Third, there are likely some interesting ideas about the level and intensity of competition. If more intense competition weakens holdup power, then firms are more likely to choose OS.

There are also some predictions that relate to characteristics of host economies and their institutions. Fourth, the state of capital market development in a host economy can be captured by our parameter ρ , the share of first-period capital that a local firm can finance. The higher the state of capital-market development, the higher ρ and therefore more firms choose OS. Fifth, more advanced institutions in the host country, including intellectual property protection (IPR) and contract enforcement, can be thought of as a higher value of σ , the share of knowledge capital that is protected and contractible. More advanced institutions then imply more OS. Sixth, the level of local managerial skills and entrepreneurship can matter. With some imagination, we could think of W as reflecting the skills and opportunity cost of a local manager (though better managers are also more productive). A higher value of W encourages OS.

The empirical literature on multinational firms' mode choice has explored different aspects of our theoretical predictions and given support to varying degrees. In what follows, we will largely confine our attention to evidence on whether or not knowledge intensity is correlated with OS, whether measured by Tobin's Q or by other evidence. Second, we will examine some work on the relationship between host-country institutions such as IPR in relation to the FDI–OS decision. We caution that this is not simple, primarily for the reason that we do not have good measures of OS. There is a fair bit of evidence on the exporting versus FDI decision, but much less on the FDI versus OS decision.

Given the lack of OS data, most existing empirical work could only focus on how knowledge intensity affects the parent firm's FDI choice. In order to define the FDI mode choice, some authors (Keller and Yeaple 2013; Awokuse *et al.* 2012; Jinji *et al.* 2012) use direct information on whether or not the parent firm has set up a foreign affiliate either at firm level or at relatively aggregate industry level. Others (Antràs 2003; Yeaple 2006; Costinot *et al.* 2009; Antràs and Yeaple 2013) try to capture FDI behavior by using some indirect information such as intra-firm trade flows, specifically the share of intra-firm imports in total US imports.

It is quite difficult to find a good empirical variable to represent knowledge intensity as defined in our theoretical model. So there are different ways to measure knowledge intensity in the literature.

The most commonly used ones are R&D intensity, advertising intensity, skilled-labor intensity and information technology capital share (computer share); see Antràs (2003), Yeaple (2006), Awokuse *et al.* (2012), Antràs and Yeaple (2013), and Keller and Yeaple (2013). Costinot *et al.* (2009) use average “routineness” of production (the importance of problem-solving) and also control for R&D intensity and skilled-labor intensity to analyze the effects of knowledge capital. Jinji *et al.* (2012) calculate each publicly traded firm’s Tobin’s Q as their knowledge intensity variable.

No matter what combination of FDI variable and knowledge intensity variable is chosen, empirical evidence suggests that multinational firm tends to choose to set up its own foreign subsidiary when the knowledge intensity of the production process is high. It is quite clear from the data that knowledge intensity increases the incentive for firms to choose FDI mode.

One advantage of Japanese firm-level data is that it gives information on both sales of foreign affiliates and OS, so it allows Jinji *et al.* (2012) to show how knowledge intensity (Tobin’s Q) affects a firm’s FDI–OS decision. They find that a firm with a higher Tobin’s Q (higher knowledge intensity) favors FDI over OS, which could serve as a direct proof of our first theoretical prediction.

Some of the empirical analyses on FDI–OS decisions relate these two mode choices to firm-level heterogeneity in productivity. This type of research usually applies firm-level data, including Tomiura (2007), for Japan, Kohler and Smolka (2009) for Spain, and Xie and Markusen (2013) for Chile. In spite of different countries, these papers all find that less productive firms usually prefer OS while more productive firms are more likely to choose FDI. If we believe that knowledge capital is one of the key factors that drive up productivity (see Pakes and Schankerman 1984; Doraszelski and Jaumandreu 2011) so that higher productivity is caused by higher knowledge intensity of a firm, this specific productivity and mode choice relationship could support our theory indirectly.

Smith (2001) checks how foreign patent rights (FPR) affect US firms’ mode choice. Yang and Maskus (2001) and Maskus *et al.* (2005) analyze the effects of patent rights on licensing and FDI. Keller and Yeaple (2013) include the country-level IPR index from Park (2008) as a control variable when they test the knowledge transfer of multinational firms. All papers find that multinational firms are more willing to outsource their production processes abroad under stronger FPR or IPR protection. Since better IPR protection also indicates a more desirable host-country environment, it may encourage firm-level FDI behavior especially on the extensive margin, according to Smith (2001) and Maskus *et al.* (2005).

It seems better to analyze the effect of contract efficiency to compare the OS decision to FDI. Arora (1996) suggests that bundling know-how with complementary inputs could increase the efficiency of contract and know-how transfer through OS. This is consistent with our approach, in which owning the complementary physical capital protects the firm’s knowledge capital by lowering the holdup cost. Bernard *et al.* (2010) create a product contractibility measure which is based on the degree of intermediation for the product; and Antràs and Yeaple (2013) apply a contractibility measure by using judicial quality at country level introduced by Nunn (2007). Both measures could capture the knowledge capital contract efficiency problem (σ) in our model. They observe negative correlation between FDI behaviors and contract efficiency from US data. These empirical results are consistent with our theoretical finding (the fifth prediction) that larger share of not contractible “tacit knowledge” in a production technology makes multinational firms favor FDI over OS.

9 Summary

Ethier (1986) was, in our view, the first to try to deal formally with the non-excludability of knowledge as a key determinant of the choice between FDI and OS, following the less formal work of a number

of international business scholars (who referred to OS by its converse name, internalization—now in turn usually referred to as vertical integration). A number of other papers followed in the late 1990s and early 2000s. But then the focus among trade economists shifted to fully excludable physical capital using the off-the-shelf “property-rights” approach.

This paper takes a further step in trying to combine these approaches in a single model, and to use that model to derive empirical predictions. Empirical work in this area has been and continues to be hampered by a lack of data on OS, licensing and foreign contracting in the firm-level datasets that have allowed for a better grasp of the FDI versus exporting decision. Hopefully, scholars will find a way forward on this problem.

References

- Antràs, P. (2003), “Firms, contracts, and trade structure,” *Quarterly Journal of Economics* **118**, 1375–418.
- Antràs, P. (2005), “Incomplete contracts and the product cycle,” *American Economic Review* **95**, 1054–71.
- Antràs, P., and E. Helpman (2004), “Global sourcing,” *Journal of Political Economy* **112**, 552–80.
- Antràs, P., and S. R. Yeaple (2013), “Multinational firms and the structure of international trade,” *Handbook of International Economics* (in press).
- Arora, A. (1996), “Contracting for tacit knowledge: The provision of technical services in technology licensing contracts,” *Journal of Development Economics* **50**, 233–56.
- Awokuse, T. O., K. E. Maskus, and Y. An (2012), “Knowledge capital, international trade, and foreign direct investment: A sectoral analysis,” *Economic Inquiry* **50**, 707–23.
- Bernard, A. B., J. B. Jensen, S. Redding, and P. K. Schott (2010), “Intra-firm trade and product contractibility,” *American Economic Review Papers and Proceedings*, **100**, 444–8.
- Caves, R. E. (2007), *Multinational Enterprise and Economic Analysis*, 3rd edition, Cambridge: Cambridge University Press.
- Chen, Y., I. J. Horstmann, and J. R. Markusen (2012), “Physical capital, knowledge capital, and the choice between FDI and outsourcing,” *Canadian Journal of Economics* **45**, 1–15.
- Costinot, A., L. Oldenski, and J. Rauch (2009), “Adaptation and the boundary of multinational firms,” *Review of Economics and Statistics* **93**, 298–308.
- Doraszelski, U., and J. Jaumandreu (2011), “R&D and productivity: Estimating endogenous productivity,” Working Paper, Harvard University and Universidad Carlos III de Madrid.
- Dunning, J. H. (1977), “The determinants of international production,” *Oxford Economic Papers* **25**, 289–336.
- Ethier, W. J. (1986), “The multinational firm,” *Quarterly Journal of Economics* **101**, 805–33.
- Ethier, W. J., and J. R. Markusen (1996), “Multinational firms, technology diffusion and trade,” *Journal of International Economics* **41**, 1–28.
- Feenstra, R. C., and G. H. Hanson (2005), “Ownership and control in outsourcing to China: Estimating the property-rights theory of the firm,” *Quarterly Journal of Economics* **120**, 729–61.
- Fosfuri, A., M. Motta, and T. Ronde (2001), “Foreign direct investments and spillovers through workers’ mobility,” *Journal of International Economics* **53**, 205–22.
- Glass, A. J., and K. Saggi (2002), “Multinational firms and technology transfer,” *Scandinavian Journal of Economics* **104**, 495–513.
- Grossman, G. M., and E. Helpman (2002), “Integration versus outsourcing in industry equilibrium,” *Quarterly Journal of Economics* **117**, 85–120.
- Grossman G. M., and E. Helpman (2004), “Managerial incentives and the international organization of production,” *Journal of International Economics* **63**, 237–62.
- Grossman, S. J., and O. D. Hart (1986), “The costs and benefits of ownership: A theory of vertical and lateral integration,” *Journal of Political Economy* **94**, 691–719.
- Hart, O. D., and J. Moore (1990), “Property rights and the nature of the firm,” *Journal of Political Economy* **98**, 1119–58.
- Helpman, E. (1984), “A simple theory of trade with multinational corporations,” *Journal of Political Economy* **92**, 451–71.
- Horstmann, I. J., and J. R. Markusen (1987), “Licensing versus direct investment: A model of internalization by the multinational enterprise,” *Canadian Journal of Economics* **20**, 464–81.

- Jinji, N., X. Zhang, and S. Haruna (2012), "Does Tobin's q matter for firms' choices of globalization mode?," RIETI Discussion Paper Series 11-E-061.
- Keller, W., and S. Yeaple (2013), "The gravity of knowledge," *American Economic Review* **103**, 1414–44.
- Kohler, W., and M. Smolka (2009), "Global sourcing decisions and firm productivity: Evidence from Spain," CESifo Working Paper No. 2903, CESifo Group, Munich.
- Markusen, J. R. (1984), "Multinationals, multi-plant economies and the gains from trade," *Journal of International Economics* **16**, 205–26.
- Markusen, J. R. (2001), "Contracts, intellectual property rights, and multinational investment in developing countries," *Journal of International Economics* **53**, 189–204.
- Maskus, K. E., K. Saggi, and T. Puttitanun (2005), "Patent rights and international technology transfer through direct investment and licensing," K. E. Maskus and J. H. Reichman, eds, *International Public Goods and Transfer of Technology under a Globalized Intellectual Property Regime*, 265–81, Cambridge: Cambridge University Press.
- Nunn, N. (2007), "Relationship-specificity, incomplete contracts, and the pattern of trade," *Quarterly Journal of Economics* **122**, 569–600.
- Pakes, A., and M. Schankerman (1984), "An exploration into the determinants of research intensity," Z. Griliches, ed., *R&D, Patents and Productivity*, 209–32, Chicago: University of Chicago Press.
- Park, W. (2008), "International patent protection," *Research Policy* **37**, 761–6.
- Smith, P. (2001), "Patent rights and bilateral exchange: A cross-country analysis of U.S. exports, FDI, and licensing," *Journal of International Economics* **52**, 411–39.
- Tomiura, E. (2007), "Foreign outsourcing, exporting, and FDI: A productivity comparison at the firm level," *Journal of International Economics* **72**, 113–27.
- Williamson, O. E. (1975), *Markets and Hierarchies: Analysis and Antitrust Implication: A Study in the Economics of Internal Organization*, New York: Free Press.
- Xie, Y., and J. R. Markusen (2013), "Exporting, licensing, FDI and productivity choice: Theory and evidence from Chilean data," Working Paper, Fudan University and University of Colorado at Boulder.
- Yang, G.F., and K. E. Maskus (2001), "Intellectual property rights and licensing: An econometric investigation," *Review of World Economics (Weltwirtschaftliches Archiv)* **137**, 58–79.
- Yeaple, S. (2006), "Foreign direct investment, and the structure of U.S. trade," *Journal of the European Economic Association* **4**, 602–11.