

# Grossman–Hart (1986) Goes Global: Incomplete Contracts, Property Rights, and the International Organization of Production

Pol Antràs\*  
Harvard University

I survey the influence of Grossman and Hart's (1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," 94 *Journal of Political Economy* 691–719.) seminal paper in the field of International Trade. I discuss the implementation of the theory in open-economy environments and its implications for the international organization of production and the structure of international trade flows. I also review empirical work suggestive of the empirical relevance of the property-rights theory. Along the way, I develop novel theoretical results and also outline some of the key limitations of existing contributions. (*JEL* D23, F10, F12, F14, F21, F23, L22, L23).

## 1. Introduction

In 1986, the year of publication of Sandy Grossman and Oliver Hart's seminal paper on the boundaries of the firm, the state-of-the-art mobile phone, the Motorola DynaTAC 8000x, resembled a large brick, weighed close to two pounds, and was sold at \$3995 a piece (or a whopping \$8,235 in current US dollars). That same year, IBM released its first laptop computer, the 13-pound IBM PC Convertible, featuring an Intel 80C88 CPU that ran at 4.77MHz, 256 kBytes of RAM, and a narrow 10-inch LCD screen, all for "just" \$2000 (or the equivalent of \$3820 today).

The last 25 years have witnessed an information and communication technology (ICT) revolution that has transformed the world economy. The exponential growth in the processing power and memory capacity of computers (as exemplified by Moore's law) and the equally exponential growth in the network capacity of optical fiber have led to a dramatic fall in the cost of processing and transmitting information at long distances. One of the manifestations of this ICT revolution has been the gradual

---

\*This survey was prepared for the "Grossman and Hart at 25" conference held in Brussels on June 24–26, 2011. I am grateful to Lucia Antras, Davin Chor, Arnaud Costinot, Oliver Hart, and an anonymous referee for very helpful comments and to Elhanan Helpman for many conversations on this topic over the years. I am also grateful to Davin Chor, Federico Díez, Nathan Nunn, Heiwai Tang, and Dan Treffer for kindly sharing their data with me, and to Ruiqing Cao for excellent research assistance.

disintegration of production processes (or “slicing of the value chain”) across borders. More and more firms now organize production on a global scale and choose to offshore parts, components, or services to producers in foreign and often distant countries.

Although this trend is significant enough to be salient in aggregate statistics, it is often best illustrated through particular examples.<sup>1</sup> Apple’s iPad 2 tablet is a case in point. Its slim and sleek exterior hides a complex manufacturing process combining components provided by multiple suppliers located in various countries. Apple does not disclose information on its input providers, but teardown reports (such as those published by [isuppli.com](http://isuppli.com) and [ifixit.com](http://ifixit.com)) have shed light on the global nature of the iPad 2 production process. The tablet itself is assembled in China (and by the end of 2011 also in Brazil) by Taiwan-based Foxconn. The displays are believed to be manufactured by LG Display and, more recently, by Samsung, both of which are based in South Korea. The distinctive touch panel is produced by Wintek, a Taiwan-based company that also owns plants in China, India, and Vietnam, while the case is provided by another Taiwanese company, Catcher Technologies, with operations in Taiwan and China. A third important component, the battery pack, also originates in Taiwan and is sold by Simplo Technologies and Dynapack International. Apart from these easily identifiable parts, the iPad 2 incorporates a variety of chips and other small technical components provided by various firms with R&D centers in developed economies and manufacturing plants (under various organizational structures) worldwide. A non-exhaustive list includes (again) Korea’s Samsung, which is believed to manufacture the main processor (designed by Apple) and possibly the flash memory, Japan’s Elpida contributing the SDRAM, Germany’s Infineon and US Qualcomm both supplying 3G modules, and Italo-French STMicroelectronics, Japan’s AKM Superconductors, and US TAOS each contributing key sensors.

When designing their global sourcing strategies, one of the key organizational decisions faced by firms concerns the extent of control they want to exert over their foreign production processes. Ownership of foreign assets is one of the key methods to enhance such control as exemplified by Intel’s decision in 1997 to offshore a significant part of its worldwide production of microprocessors to a \$300-million wholly-owned manufacturing plant in Costa Rica. Other firms, such as Nike or Apple, also rely heavily on offshore manufacturing, but choose to subcontract production to non-affiliated producers around the world, while keeping within firm boundaries only the design and marketing stages of production.

Why do some firms find it optimal to exert a tight control over their foreign production operations while others choose not to do so? Or

1. For aggregate evidence on the increase in vertical fragmentation, see [Feenstra and Hanson \(1996\)](#); [Campa and Goldberg \(1997\)](#); [Hummels et al. \(2001\)](#); [Yeats \(2001\)](#). This trend also featured prominently in the [WTO’s 2008 World Trade Report](#).

paraphrasing the opening sentences of Grossman and Hart's 1986 article (referred to as GH, hereafter), "what is a multinational firm? And what are the determinants of how vertically or laterally integrated the activities of the multinational firm are?" The purpose of this article is to provide an overview of recent literature in international trade that has attempted to provide answers to these questions by applying some of the insights of the influential property-rights theory first explicated in GH and further developed in Hart and Moore (1990) and Hart (1995).

The relevance of incomplete-contracting theories of integration is not particularly difficult to motivate in international environments in which goods or services are exchanged across borders. Global production networks necessarily entail contracting relationships between agents located in countries with heterogeneous legal systems and contracting institutions.<sup>2</sup> A natural difficulty in contract disputes involving international transactions is determining which country's laws are applicable to the contract being signed. Even when they are competent (in a legal sense), local courts may be reluctant to enforce a contract involving residents of foreign countries, especially if such enforcement would entail an unfavorable outcome for local residents. Although there have been coordinated attempts to reduce the contractual uncertainties and ambiguities associated with international transactions, such as the signing of the United Nations Convention on Contracts for the International Sale of Goods (CISG), many countries have opted out of these agreements (most notably, the UK). Other forms of arbitration, such as those provided by the International Chamber of Commerce in Paris, are also available but are rarely used in practice. The limited amount of repeated interactions and lack of collective punishment mechanisms often associated with international transactions also makes implicit contracts less sustainable in those environments. As summarized by Rodrik (2000) "ultimately, [international] contracts are often neither explicit nor implicit; they simply remain incomplete."

The detrimental effects of imperfect contract enforcement on international trade flows are particularly acute in transactions involving intermediate inputs, as those tend to be associated with longer time lags between the time the order is placed (and the contract is signed) and the time the goods or services are delivered (and the contract executed), and they also often entail significant relationship-specific investments and other sources of lock-in on the part of both buyers and suppliers. For instance, suppliers often tailor their manufacturing production to the needs of particular buyers and have difficulty placing those goods in world markets should the buyer decide not to abide by the terms of the contract. Similarly, buyers undertake significant investments whose value

2. As an example, Djankov et al. (2003) document that the total duration of a legal procedure aimed at collecting a bounced check ranges from seven days in Tunisia to 1003 in Slovenia.

can be severely diminished by incompatibilities, production line delays, or quality debasements associated with suppliers not following through with their contractual obligations.

In *Poorly Made in China*, Paul Midler describes his misadventures as an offshoring consultant in China, where his command of Chinese made him a valuable asset for American companies seeking suppliers in that country. Midler describes numerous last-minute pricing maneuvers and clever manipulations of quality undertaken by Chinese suppliers, attempting to extract surplus from the Western buyers they transact with.<sup>3</sup> As Midler puts it, “‘Price go up!’ was the resounding chorus heard across the manufacturing sector” (year 2009: 184), which is very much reminiscent of the classical hold-up problem modeled in GH. Furthermore, Midler illustrates how potential solutions to this hold-up problem tend to be ineffective in China. For instance, relational contracting does not appear to diminish the frustrations of Western companies, and if anything, tends to increase them, something he labels the “Reverse Frequent Flyer” effect. Furthermore, and consistently with one of the fundamental assumptions in GH, ownership of Chinese suppliers does not eliminate opportunistic behavior on the part of these producers, or as he puts it, there is no “Joint Venture Panacea”.

The remainder of the article is divided into three sections and a final section with concluding remarks. In Section 2, I develop a variant of the partial equilibrium framework in GH, along the lines of Antràs (2003), which has served as the basis for most applications of the property-rights theory in international environments. While the literature has generally assumed particular functional forms to derive results on the key determinants of the integration decision, I show that some of the key predictions of the model are robust to more general assumptions on preferences and technology. Later in the section, I outline several extensions of this framework that have been developed in the literature, including the introduction of liquidity constraints, multiple suppliers, partial contractibility, and productivity heterogeneity across firms.

In Section 3, I summarize different ways in which this partial-equilibrium property-rights model has been adapted in general-equilibrium, open-economy environments. Even restricting oneself to partial-equilibrium environments, considering the location of different parts of the production process significantly complicates the organizational decisions of firms. It is thus natural that the literature has largely focused on simple environments with either a limited number of countries (often two) or with particularly convenient symmetry assumptions across firms and countries. As a payoff to these simplifications, these frameworks deliver sharp implications for how the organizational decisions of firms aggregate up to easily observable statistics, such as the intrafirm component

3. The lack of enforceability of contracts is illustrated by the Chinese old saying that “signing a contract is simply a first step in negotiations.”

of international trade across sectors and countries. I argue, however, that future contributions should incorporate more realistic features, particularly when the goal is to develop theoretical frameworks that guide empirical analysis using *firm-level data*.

The empirical implementation of open-economy property-rights frameworks is precisely the focus of Section 4. Empirically validating the property-rights theory poses at least two important challenges. First, the theory's predictions are associated with subtle characteristics of marginal returns to investments that are generally unobservable in the data (see Whinston 2003). Second, data on the integration decisions of firms are not readily available. Admittedly, the international trade literature has not made an awful lot of progress addressing the first concern (although it has not ignored it). Conversely, data on international transactions are particularly accessible due to the existence of official records of goods crossing borders. Furthermore, some countries collect detailed data on whether international trade flows involve related parties or non-related parties, thus generating hundreds of thousands of observations *per year* describing the relative prevalence of integration across products and countries. In Section 4, I describe the rich variation observed in US intrafirm import data and explicitly discuss the pros and cons of using this source of data to test the property-rights theory. I also review several papers in the empirical literature on intrafirm trade by graphically illustrating their key findings and how they have been interpreted in light of the GH framework. Finally, I briefly discuss a few recent contributions using international firm-level data sources and suggest some avenues for future research.

In Section 5, I offer some concluding remarks and outline the broader influence of GH in the field of international trade.

Before proceeding any further, it is important to mention some topics that are omitted in this survey. First, it should be emphasized that GH's property-rights approach has not been the only theory of the firm applied to understand multinational firm boundaries. The literature has understood since Hymer's (1960) seminal PhD thesis that the issue of control is essential to understanding the nature of the multinational firm, and several early contributions adopted the transaction-cost approach of Coase (1937) and Williamson (1975, 1985) to shed light on some of the key determinants of the boundaries of the multinational firm.<sup>4</sup> The first general-equilibrium implementation of the transaction-cost approach is due to Ethier (1986), with important subsequent contributions by Ethier and Markusen (1996); McLaren (2000); and Grossman and Helpman (2002, 2003).<sup>5</sup> Second, the international trade literature has also concerned itself with other organizational decisions of multinational firms, such as

4. See, among others, Casson (1979); Dunning (1981) or Rugman (1981).

5. See also Qiu and Spencer (2002) and Chen and Feenstra (2008) for related frameworks. As a curiosity, Ethier's (1986) article cites a 1984 working paper version of GH, and thus appears to be the first published paper in international trade to cite GH.

the allocation of decision rights among employees and the optimal compensation of worker effort. For instance, [Marin and Verdier \(2009\)](#) and [Puga and Trefler \(2010\)](#) have applied the [Aghion and Tirole \(1997\)](#) authority framework in general-equilibrium, open-economy environments, while [Grossman and Helpman \(2004\)](#) have studied how optimal incentive schemes interact with multinational firm boundaries. Reviewing these contributions in detail is outside the scope of this article, though I will touch upon them towards the end of Section 3.<sup>6</sup> Similarly, I will refrain from delving into the foundations of incomplete contracting frameworks since this is not more of an issue in open-economy environments than it is in closed-economy ones, and since it has been discussed in depth elsewhere in the literature (see, for instance, the articles in the 1999 Symposium in the *Review of Economic Studies*).

## 2. A Property-Rights Model

In this section, I develop a simple variant of GH along the lines of [Antràs \(2003\)](#).

### 2.1 Model Setup

**2.1.1 Environment.** Consider a situation in which only the manager of a firm  $F$  has access to a technology for converting a specialized intermediate input or component  $m$  into a differentiated final good. The manager  $F$  is also in charge of providing headquarter services  $h$ , which raise the marginal production of  $m$ . Given an amount  $m$  of components and an amount  $h$  of headquarter services, sale revenue is given by  $R(h, m)$  with  $R_h > 0$ ,  $R_m > 0$ ,  $R_{hm} > 0$ ,  $R_{hh} < 0$ ,  $R_{mm} < 0$ , and  $R(h, 0) = R(0, m) = 0$ . The manager  $F$  needs to contract with an operator of a manufacturing plant (denoted by  $M$ ) for the provision of  $m$ . Production of  $h$  and  $m$  require investments on the part of  $F$  and  $M$ , respectively.  $F$  obtains  $1/c_h$  units of  $h$  for each unit of investment, while  $M$  obtains  $1/c_m$  units of  $m$  for each unit of investment. Investments are made simultaneously at some date  $t = 1$  and the inputs are obtained at a later date  $t = 2$ . The inputs are tailored specifically to the other party in the transaction and are useless or incompatible in alternative production processes.<sup>7</sup> Finally,  $F$  converts inputs  $h$  and  $m$  into the final good at a final date  $t = 3$ . For simplicity, agents do not discount the future between  $t = 1$  and  $t = 3$ .

6. These topics have been treated in some detail in surveys by [Markusen \(1995\)](#); [Spencer \(2005\)](#); [Helpman \(2006\)](#); [Antràs and Rossi-Hansberg \(2009\)](#), as well as in the classic book by [Caves \(1996\)](#).

7. Of course, partial relationship-specificity would suffice for the results. Furthermore, in the presence of search frictions or time-to-build constraints, a lock-in effect might arise even in the absence of customization of inputs, as producers might be unable to costlessly combine their inputs with those of other producers.



2.1.2 Contracting. The managers  $F$  and  $M$  get together at some initial stage  $t=0$  to negotiate a contract. As in GH, before investments in  $h$  and  $m$  are made, the only contractibles are the allocation of residual rights (i.e., the ownership structure) and a lump-sum transfer between the two parties.<sup>8</sup> I will consider below more general environments in which contracts on a wider set of variables are enforceable. Without a binding contract detailing the terms of exchange, parties are left to negotiate these terms (i.e., a payment by  $F$  to  $M$  for the provision of  $m$ ) after these inputs have been produced at stage  $t=2$ . As in GH, it is assumed that this ex-post agreement is fully enforceable, that  $F$  and  $M$  have symmetric information at that stage, and that the negotiation outcome can be approximated by the symmetric Nash bargaining solution. The significance and ramifications of these assumptions have been discussed at length in the literature, and although they remain controversial, they continue to be standard in the literature.<sup>9</sup>

2.1.3 Property Rights and Bargaining Power. Following GH's property-rights approach to the theory of the firm, and contrary to the Coase–Williamson approach, it is assumed that the space of contracts and the nature of the ex-post negotiations between  $F$  and  $M$  are independent of the ownership structure decision at stage  $t=0$ . The ex-post distribution of surplus is, however, sensitive to the chosen organizational structure. To see this, notice that given our assumptions, when  $F$  and  $M$  remain unaffiliated entities, a case I shall refer to throughout as *outsourcing*, a contractual termination leaves both agents with a zero payoff because the inputs are useless unless combined together. In this framework the same would be true if  $M$  were to integrate  $F$  because  $F$ 's human capital has been assumed essential for the production of the final good. For this reason, and because it does not appear to feature prominently in the data, I hereafter abstract from considering the possibility of (forward) integration of  $F$  by  $M$ . The case of backward integration, which I refer to throughout simply as *integration*, is more consequential. This is because when  $F$  integrates the production of  $m$ , it effectively purchases the residual rights of control over this input. As a result, if  $M$  refuses to trade after these inputs have been produced,  $F$  now has the option of selectively firing  $M$  and seizing the amount of  $m$  already produced. To ensure positive ex-post gains from trade, it is assumed, however, that  $F$  cannot use the input

8. In practice, contractual disputes are often triggered by concerns about substandard quality of inputs or incompatibilities of these inputs with other parts of the production process, both of which are seldom verifiable by third parties. Furthermore, the particular nature of the required investments in  $h$  and  $m$  might be difficult to specify in a comprehensive contract.

9. Following GH, I assume that both agents have symmetric primitive bargaining in the ex post negotiations. The international trade literature has for the most part considered the generalized Nash bargaining solution.

without  $M$  as effectively as it can with the cooperation of  $M$ , so firing  $M$  results in a loss of a fraction  $1 - \delta$  in sale revenue.

**2.1.4 Discussion.** At this point, it may be useful to compare the assumptions I have made with those in GH. First, note that the above framework is more general than GH in two respects: it allows investments to be complementary in creating surplus and it also considers the possibility that one agent's investments might affect the other agent's disagreement payoffs, thereby avoiding the need to narrowly interpret investments as being human capital investments. As pointed out by Whinston (2003), these features are indispensable in order to be able to comfortably use the model to interpret integration decisions in the real world.

Despite these generalizations, it should be clear that the above framework is more restrictive than GH in at least three dimensions. First, ownership of physical assets (buildings, machines) has been associated above with ownership of the inputs that are produced with those assets. This is a narrow interpretation of the role of asset ownership, but it is a useful assumption to make when one is trying to simplify the exposition of the partial-equilibrium side of the model. Second, although in principle, the parties could agree at  $t = 0$  on the allocation of ownership rights over two assets ( $h$  and  $m$ ), as pointed out above, it is never optimal to allocate ownership rights over  $h$  to  $M$ . Again, this is largely an expositional assumption that draws attention away from situations that do not appear essential in general equilibrium, open-economy applications of the GH framework. Third, the model above places much more structure on how investments affect inside and outside options than the original GH framework does. To be more precise, given the assumptions above, the payoffs obtained by  $F$  and  $M$  in the ex-post bargaining are proportional to a common aggregator of  $h$  and  $m$ , with the ownership structure decision simply affecting the shares obtained by each agent. In particular,  $F$  obtains a fraction  $\beta_k$  of revenue  $R(h, m)$  and  $M$  obtains the remaining  $1 - \beta_k$ , where

$$\beta_k = \begin{cases} \beta_O = \frac{1}{2} & \text{if } F \text{ outsources to } M \\ \beta_V = \delta + \frac{1}{2}(1 - \delta) & \text{if } F \text{ integrates } M \end{cases} \quad (1)$$

Obviously, this is a significant point of departure from GH, but as I will shortly demonstrate, it will deliver a particularly sharp representation of one of the key insights from the property-rights theory, while sidestepping Whinston's (2003) criticism about the robustness of GH's insights to the presence of complementary investments.

**2.1.5 Formulation of the Problem.** Having discussed our assumptions, I next offer a succinct formulation of the "organizational problem" solved by  $F$  and  $M$ . Notice that given the existence of ex ante transfers, firms will agree at  $t = 0$  on the ownership structure (outsourcing or integration) that maximizes the joint payoff of  $F$  and  $M$ . This surplus is of course affected by the



investments in  $h$  and  $m$  undertaken by  $F$  and  $M$ ; these investments are chosen simultaneously and noncooperatively by these two agents at  $t=1$  with the aim of maximizing their ex post payoffs at  $t=2$ . In sum, the key organizational decision  $k \in \{V, O\} - V$  for vertical integration and  $O$  for outsourcing – solves

$$\begin{aligned} \max_{k \in \{V, O\}} \quad & \pi_k = R(h_k, m_k) - c_h \cdot h_k - c_m \cdot m_k \\ \text{s.t.} \quad & h_k = \arg \max_h \{ \beta_k R(h, m_k) - c_h \cdot h \} \\ & m_k = \arg \max_m \{ (1 - \beta_k) R(h_k, m) - c_m \cdot m \}, \end{aligned} \quad (\text{P1})$$

where  $\beta_k$  is given in equation (1).

## 2.2 Characterization

Notice that program (P1) boils down to choosing a value of  $\beta \in \{\beta_V, \beta_O\}$  to maximize joint profits. As suggested by [Antràs and Helpman \(2004, 2008\)](#), a pedagogically useful way to characterize the optimal choice of ownership structure is to consider the hypothetical case in which  $F$  and  $M$  could freely choose  $\beta$  from the continuum of values in  $[0, 1]$ . Intuitively, integration will tend to be more attractive in situations in which the joint-profit maximizing share of surplus accruing to  $F$  is large, while outsourcing will tend to be preferred when this share is low.

Straightforward differentiation of the objective function in program (P1) delivers

$$\frac{d\pi_k}{d\beta} = (R_h - c_h) \frac{dh}{d\beta} + (R_m - c_m) \frac{dm}{d\beta}, \quad (2)$$

which, plugging the first-order condition associated with the constraints in program (P1) and setting  $d\pi_k/d\beta=0$ , delivers the following joint-profit maximizing division of surplus  $\beta^*$ :

$$\frac{\beta^*}{1 - \beta^*} = \frac{\eta_{R,h} \cdot \xi_{h,\beta}}{\eta_{R,m} \cdot (-\xi_{m,\beta})}, \quad (3)$$

where  $\eta_{R,j} \equiv jR_j/R$  is the elasticity of surplus to investments in input  $j=h, m$  and  $\xi_{j,\beta} \equiv \frac{dj}{d\beta} \frac{\beta}{j}$  is the elasticity of investment in  $j$  to changes in the distribution of surplus  $\beta$ . In words, equation (3) implies that:<sup>10</sup>

*Proposition 1.* The (hypothetical) optimal share of revenue allocated to an agent is increasing in the elasticity of revenue with respect to that agent's investment and in the elasticity of that agent's investment with respect to changes in the distribution of surplus.

10. I choose to represent equation (3) in terms of elasticities rather than in terms of marginal returns and marginal effects on investment, i.e.,  $\frac{\beta^*}{1-\beta^*} = \frac{R_h \cdot (dh/d\beta)}{R_m \cdot (-dm/d\beta)}$ , to facilitate a comparison with previous results derived in the literature and discussed below.

Because in the property-rights theory the only way to shift surplus between agents is via the allocation of ownership rights, the first result above is reminiscent of one of the key results in GH, namely that *ex ante* efficiency dictates that ownership of assets, and thus residual rights of control, should be assigned to the party whose investment contributes relatively more to the value of the relationship. The relative importance of an agent's investment is captured here by the elasticity of revenue with respect to that agent's investment.

Equation (3) highlights, however, that the responsiveness of investments to changes in bargaining power is also an important determinant of the ownership decision. What determines this responsiveness? One might worry that if this elasticity depends on the division of surplus  $\beta$  and on features of the revenue function, the overall effects of the revenue elasticities  $\eta_{R,h}$  and  $\eta_{R,m}$  might be less clear-cut than as suggested in equation (3). Totally differentiating the first-order conditions associated with the constraints in program (P1) we indeed find that

$$\xi_{h,\beta} = \frac{-R_h(1-\beta)R_{mm} - R_m\beta R_{hm}}{(1-\beta)h(R_{hh}R_{mm} - (R_{hm})^2)}; \quad (4)$$

$$\xi_{m,\beta} = \frac{\beta R_m R_{hh} + R_h(1-\beta)R_{hm}}{(1-\beta)m(R_{hh}R_{mm} - (R_{hm})^2)}. \quad (5)$$

These equations indicate that the optimal division of surplus, and thus the optimal allocation of ownership rights, depends on subtle properties of the revenue function  $R(h, m)$ , not just on its partial derivatives.<sup>11</sup> Although as pointed out above, the assumptions we have made about how investments affect inside and outside options are rather restrictive, these results resonate with those expositied by Whinston (2003).

To make some progress, the literature has typically assumed particular functional forms for the revenue function  $R(h, m)$ . A particularly widely used formulation, starting with the work of Antràs (2003), is to assume

$$R(h, m) = Ah^{\eta_h}m^{\eta_m}, \quad (6)$$

with  $A > 0$ ,  $0 < \eta_h < 1$ ,  $0 < \eta_m < 1$ , and  $\eta_h + \eta_m < 1$ . This specification may seem *ad hoc*, but it follows directly from two assumptions that are fairly common in modern international trade models, namely that preferences feature a constant-elasticity of substitution across varieties within an industry (c.f., Dixit and Stiglitz, 1977) and that inputs are combined in production according to a Cobb–Douglas technology (see Section 3 for more details). It may be argued that the Cobb–Douglas assumption is particularly restrictive so I will address this concern below. In any case,

11. Note that even the effect of  $\beta$  on investments is of ambiguous sign. The denominator in (4) and (5) is necessarily positive if the second-order conditions for the choice of  $h$  and  $m$  in a complete contracting environment are to be met, but the sign of the numerators is shaped by the relative concavity and complementarity of  $R(h, m)$ .

with the revenue function in equation (6), the revenue elasticities  $\eta_{R,h}$  and  $\eta_{R,m}$  are pinned down by the parameters  $\eta_h$  and  $\eta_m$ , respectively, and the ratio of elasticities  $\xi_{h,\beta}/\xi_{m,\beta}$  turns out to depend only on these parameters as well. Equation (3) then takes a particularly simple form:

$$\frac{\beta^*}{1 - \beta^*} = \sqrt{\frac{\eta_h/(1 - \eta_h)}{\eta_m/(1 - \eta_m)}}. \quad (7)$$

I will associate throughout the parameters  $\eta_h$  and  $\eta_m$  with the terms *headquarter intensity* and *component intensity*, respectively. It is then clear from equation (7) that the joint-profit maximizing share of revenue assigned to an agent is increasing in the relative importance of that agent's investment, as measured by the elasticity of revenue to that investment. In terms of the choice of ownership structure, this isoelastic example delivers the following sharp result (see, for instance, [Antràs and Helpman, 2008](#), for a proof):

*Proposition 2.* Let firm revenues be given by equation (6) and let  $\eta_h = \alpha v$  and  $\eta_m = \alpha(1 - v)$ , with  $0 < \alpha < 1$  and  $0 < v < 1$ . Then there exists a unique threshold  $\hat{v} \in (0, 1)$  such that for all  $v > \hat{v}$ , integration dominates outsourcing (or  $k^* = V$ ), while for  $v < \hat{v}$ , outsourcing dominates integration (or  $k^* = O$ ). Furthermore, the threshold  $\hat{v}$  is independent of the cost parameters  $c_h$  and  $c_m$ .

Hence, integration is optimal for headquarter intensities above (or component intensities below) a given threshold, while outsourcing is chosen for headquarter intensities below (or component intensities above) that threshold.

Even though the specification in equation (6) has been widely used in the literature, one might still be concerned that the results in Proposition 2 heavily rely on the Cobb–Douglas assumption implicit in equation (6). It can be shown, however, that the result readily generalizes to the case in which the revenue function is a general homogenous of degree  $\alpha \in (0, 1)$  in  $h$  and  $m$ .<sup>12</sup> In Appendix A, it is shown that in that case, equation (3) simplifies to:

$$\frac{\beta^*}{1 - \beta^*} = \sqrt{\frac{\eta_{R,h} \alpha(1 - \eta_{R,m}) + (1 - \alpha)(\sigma_{h,m} - 1)\eta_{R,m}}{\eta_{R,m} \alpha(1 - \eta_{R,h}) + (1 - \alpha)(\sigma_{h,m} - 1)\eta_{R,h}}}, \quad (8)$$

where  $\eta_{R,h}$  and  $\eta_{R,m}$  again denote the revenue elasticities of headquarter services and components, respectively, and  $\sigma_{h,m}$  is the elasticity of substitution between headquarter services  $h$  and the input  $m$  in revenue. Simple differentiation then confirms that for any constant  $\sigma_{h,m} > 0$ ,  $\beta^*$  continues

12. This would be the case, for instance, if the inverse demand faced by the final-good producer is homogenous of degree  $\alpha_r - 1 < 0$  in output—as with Dixit–Stiglitz preferences—and the production function combining  $h$  and  $m$  is homogenous of degree  $\alpha_q \in (0, 1]$ . In such a case, we would have  $\alpha = \alpha, \alpha_q$ .

to be increasing in  $\eta_{R,h}$  and decreasing in  $\eta_{R,m}$ , and as a result it continues to be efficient to allocate residual rights of control and thus “power” to the party whose investment has a relatively larger impact on surplus.<sup>13</sup> In other words, the prediction of the model that integration is more attractive in headquarter-intensive sectors than in component-intensive sectors appears robust.

## 2.3 Extensions

Before discussing the open-economy implementation of the above model, I briefly describe some of the extensions of the benchmark model above that have been considered in the literature. These extensions have been developed in open-economy, general-equilibrium environments, but it is more convenient to discuss them here and refer back to them in Section 3. For simplicity, I develop these extensions one at a time, though they could be readily incorporated in a unified framework. Also, I focus throughout on revenue functions akin to equation (6), which implicitly assume that  $F$  faces a demand schedule with a constant price elasticity and that headquarter services/headquarter services and the bundle of supplier inputs feature a unit elasticity of substitution. Presumably, the results I am about to discuss would still hold under more general environments (such as those discussed at the end of the previous section), but I will not attempt to verify this here.

**2.3.1 Financial Constraints.** Following GH, I have assumed so far that the choice of organizational form is always *ex ante* efficient, in the sense that, at  $t=0$ , both parties can freely exchange lump-sum transfers and this ensures that a joint-profit maximizing ownership structure is chosen. In practice, however, it is not clear that firms can easily resort to nondistortionary transfers in their initial negotiations. For instance, some firms might be financially constrained and might have difficulties raising the amount of cash that would be needed to compensate their counterparty for choosing a particular ownership structure, and this may lead to an inefficient choice of organizational form. I next build on Basco (2010) to illustrate how financial constraints shape the choice of ownership structure within the version of GH developed above. A more complete treatment of how liquidity constraints shape organizational choices in other variants of the GH model can be found in Aghion and Tirole (1994), Legros and Newman (2008) and Carluccio and Fally (forthcoming).

In anticipation of the market structure discussed below in Section 3, I focus on situations in which  $F$  has full bargaining power *ex ante* (i.e.,  $F$  makes a take-it-or-leave-it offer to  $M$ ) and the initial contract calls for a positive transfer from  $M$  to  $F$ . The key new feature is that  $M$  is now

13. It should be noted that if the revenue function does not feature a constant elasticity of substitution between  $h$  and  $m$ , then  $\sigma_{h,m}$  will of course be endogenous to the parameters of the model.

assumed to only be able to pledge to external financiers at most a share  $\phi$  of the income it receives from transacting with  $F$ , which remember is given by  $(1 - \beta_k)R(h_k, m_k) - c_m \cdot m_k$  under organizational mode  $k \in \{V, O\}$ .<sup>14</sup> When financial constraints bind, the optimal ownership structure from the point of view of  $F$  now solves program (P1) but with the objective function now given by

$$\pi_k = \beta_k R(h_k, m_k) - c_h h_k + \phi[(1 - \beta_k)R(h, m) - c_m m_k]. \quad (9)$$

Following the same steps as in Section (2.2), we find that the (hypothetical) profit maximizing division of surplus  $\beta^*$  for  $F$  is given by

$$\frac{\beta^*}{1 - \beta^*} = \phi \frac{\eta_{R,h} \cdot \xi_{h,\beta}}{\eta_{R,m} \cdot (-\xi_{m,\beta})} + (1 - \phi) \frac{\beta^*}{1 - \beta^*} \frac{1}{\eta_{R,m} \cdot (-\xi_{m,\beta})},$$

which naturally reduces to equation (3) when  $\phi = 1$  and financial constraints disappear. When the revenue function further has the isoelastic form in equation (6), we can use equations (4) and (5) to obtain:

$$\frac{\beta^*}{1 - \beta^*} = \phi \frac{\eta_h (1 - \beta^*) - \eta_m}{\eta_m \beta^* - \eta_h} + (1 - \phi) \left( \frac{\eta_h (1 - \beta^*) - \eta_m}{\eta_m \beta^* - \eta_h} + \frac{1 - \eta_m}{\eta_m} \right). \quad (10)$$

A few things are worth highlighting about equation (10). First, and quite obviously, we have that when  $\phi = 1$ , financial constraints disappear and equation (10) simplifies to equation (7). Second, note that the right-hand-side is a weighted sum of two decreasing functions of  $\beta^*$ , with the weights being  $\phi$  and  $1 - \phi$ , and with the second term being necessarily higher than the first one. This immediately implies that  $\beta^*$  is necessarily, higher, the lower is  $\phi$ . In words, the desired division of surplus is more tilted towards  $F$ , the tighter are financial constraints, reflecting the fact that  $F$  now uses integration not only to balance incentives but also to extract surplus from  $M$ . Third, and in a related manner, the optimal share of ex post surplus accruing to  $F$  is positive even when headquarter intensity is negligible, i.e.,  $\beta^* = (1 - \eta_m)(1 - \phi)/(1 - \phi(1 - \eta_m)) > 0$  when  $\eta_h = 0$ . Finally, it is easily verified that the positive effect of  $\eta_h$  and negative effect of  $\eta_m$  on the profit-maximizing value division of surplus  $\beta^*$  continues to hold for any value of  $\phi \in (0, 1]$ .<sup>15</sup>

14. For consistency with the other parts of the model, one should not interpret this financial contract as a revenue-sharing agreement, because sale revenues are assumed to be non-contractible. A typical way to rationalize this formulation of credit constraints is to introduce limited commitment on the part of the supplier; if the supplier can default on external investors and still retain a share of sale revenue, then the size of the initial loan will indeed be proportional to expected revenue.

15. Only in the limiting case  $\phi \rightarrow 0$ , when  $M$  has no ability to transfer cash to  $F$  at stage 0, we have that  $\beta^* \rightarrow 1 - \eta_m$ , and the effect of headquarter intensity vanishes. The intuition behind this result is that when  $\phi = 0$ , the objective function of  $F$  coincides with that in the first constraint of program (P1). Hence, the effect of changes in  $\eta_h$  working through the choice of input  $h$  have no first-order effect on  $F$ 's choice of ownership structure. Of course,  $\eta_h$  could still

As in the model without financial constraints, one can show that there continues to exist a headquarter-intensity cutoff  $\hat{\eta}_h \in [0, 1)$  such that for all  $\eta_h > \hat{\eta}_h$ , integration dominates outsourcing (or  $k^* = V$ ), while outsourcing dominates integration (or  $k^* = O$ ) for  $\eta_h < \hat{\eta}_h$  whenever  $\hat{\eta}_h > 0$ . The key new features brought about by financial constraints are (see Appendix A for a proof):

*Proposition 3.* The cutoff  $\hat{\eta}_h$  is lower the larger is  $\phi$ , implying that integration is more prevalent the tighter the financial constraints. Furthermore, for large enough  $\phi$ , integration may be optimal for any value of  $\eta_h \in (0, 1)$ , i.e.,  $\hat{\eta}_h = 0$ .

As in the model without financial frictions, I show in Appendix A that  $\hat{\eta}_h$  is decreasing in  $\eta_m$ , and thus integration continues to be more prevalent, the less important are supplier investments.

**2.3.2 Multiple Suppliers.** So far, I have focused on situations in which  $F$  is concerned only with the provision of one input. Acemoglu et al. (2007) consider a more realistic environment in which production requires multiple intermediate inputs. Although their framework does not allow for headquarter services, it is straightforward to incorporate them into the analysis. With that in mind, suppose that the environment is as before, but  $F$ 's production process now entails the combination of headquarter services and a unit measure of components, each provided by a different manager. Denote by  $\mathbf{m} \equiv \{m(i)\}_{i \in [0,1]}$  the (infinitely-dimensional) vector of investments by suppliers. I impose the following functional form, which will serve to illustrate the role of input substitutability on the integration decision:

$$R(h, \mathbf{m}) = Ah^{\eta_h} \left( \int_0^1 m(i)^\rho di \right)^{\frac{\eta_m}{\rho}}. \quad (11)$$

A few comments are in order. First, equation (11) is a direct generalization of the specification in equation (6), with  $m$  being now replaced by a constant-elasticity-of-substitution aggregator of the continuum of inputs. Second, the new parameter  $\rho \in (0, 1)$  governs the substitutability of the different inputs provided by suppliers; when  $\rho \rightarrow 1$ , these inputs become perfect substitutes, while when  $\rho \rightarrow 0$ , they are all essential in production. Third, equation (11) imposes complete symmetry across production stages, both in terms of substitutability patterns as well as in terms of the importance of these suppliers' investments in production. It would be interesting to incorporate asymmetries into the framework, but I shall not attempt to do so here.

---

affect  $F$ 's profits via its effects on  $M$ 's choice of input  $m$ , but the latter effect turns out to be zero with the isoelastic revenue function in (6).



As before, the initial contract between  $F$  and its suppliers only includes an organizational structure (i.e., which suppliers are integrated and which are not) and a set of lump-sum transfers across agents, on which I place no constraints in this section. Without a binding contract governing the ex post trade in inputs, the agents in the model are left to (multilaterally) bargain over the division of surplus at  $t=2$ , when the inputs have been produced. Despite being infinitesimally small, individual suppliers can use the threat of withholding their respective inputs from  $F$  to extract surplus. Acemoglu et al. (2007) follow Hart and Moore (1990) in using the Shapley value to determine the division of ex post surplus between  $F$  and its suppliers.<sup>16</sup> It can be verified that, in a symmetric equilibrium in which no supplier is integrated, a particular supplier's  $i$  payoff in the ex post bargaining is given by (see Appendix A):

$$s_O(h, m(-j), m(j)) = \frac{\eta_m}{\eta_m + \rho} A h^{\eta_h} m(-i)^{\eta_m} \left( \frac{m(i)}{m(-i)} \right)^{\rho} = \frac{\eta_m}{\eta_m + \rho} R(h, \mathbf{m}), \quad (12)$$

where  $m(-i)$  represents the (symmetric) investments of all suppliers other than  $i$  and where in the second equality we have used the fact that supplier investments will be symmetric in equilibrium. The final-good manager  $F$  then captures the residual share  $\rho/(\eta_m + \rho)$  of revenue. Notice that the more substitutable are inputs in production (the higher is  $\rho$ ), the lower is the share of revenues that accrues to suppliers, as their ability to hold up  $F$  is lower in that case.

Consider next the polar case in which all suppliers are integrated by  $F$ . Assume that in such a case, suppliers cannot withhold the full value of their marginal contribution to revenue, but only a share  $1 - \delta$ , as in our benchmark model above.<sup>17</sup> This results in  $t=2$  payoffs for suppliers equal to

$$s_V[h, m(-i), m(i)] = \frac{\eta_m(1 - \delta)}{\eta_m + \rho} A h^{\eta_h} m(-i)^{\eta_m} \left( \frac{m(i)}{m(-i)} \right)^{\rho} = \frac{\eta_m(1 - \delta)}{\eta_m + \rho} R(h, \mathbf{m}), \quad (13)$$

leaving  $F$  with the residual share  $(\eta_m \delta + \rho)/(\eta_m + \rho)$  of revenue.

Solving for all suppliers' production levels as well as  $F$ 's provision of headquarter services, one can obtain joint profits of  $F$  and all suppliers in terms of the parameters of the model, including  $\eta_h$ ,  $\eta_m$ , and  $\rho$ , and the

16. A complication arises from the fact that we now have a continuum of agents bargaining over surplus. Acemoglu et al. (2007) resolve this issue by considering a discrete-player version of the game and computing the *asymptotic* Shapley value of Robert J. Aumann and Shapley (1974).

17. Acemoglu et al. (2007) consider an alternative formulation in which suppliers withhold a share  $1 - \delta$  of their intermediate input. This generates analogous predictions for how input substitutability shapes the integration decision, but the proofs are much more cumbersome in that case.

ownership structure decision as captured by  $\delta$ . Even in the presence of multiple suppliers, it can be shown again that the ratio of joint profits under integration relative to those under outsourcing continue to be increasing in the relative importance of headquarter services as captured by  $\eta_h$ , and thus integration is again predicted to be more prevalent in headquarter intensive sectors.<sup>18</sup> The main new result that emerges from the modeling of multiple suppliers is the following role of input substitutability in shaping the integration decisions of  $F$  (see Appendix A for a proof):

*Proposition 4.* There exists a unique threshold  $\hat{\rho} \in (0, 1)$  such that for all  $\rho < \hat{\rho}$ , integration of all suppliers dominates outsourcing, while for  $\rho > \hat{\rho}$ , outsourcing dominates integration. Hence, integration is more prevalent the higher the complementarity (or the lower the substitutability) across supplier inputs.

The intuition behind this result is as follows. When there is a high degree of technological complementarity across inputs, the ex post payoff of  $F$  under outsourcing tends to be relatively low (note, in particular, that  $F$ 's payoff under outsourcing is 0 when  $\rho \rightarrow 0$ ) and the choice of headquarter services is particularly distorted. In such cases, vertical integration is particularly attractive because it helps restore the incentives of  $F$  to provide these headquarter services. Conversely, when  $\rho$  is high, suppliers face a particularly acute hold-up problem since their inputs are highly substitutable with each other; in those situations, strengthening the bargaining power of suppliers via an outsourcing contract constitutes the profit-maximizing organizational mode.

The variant of the model with multiple suppliers that I have developed assumes that all stages of production are performed simultaneously and that  $F$  negotiates with all suppliers also simultaneously. Antràs and Chor (2011) consider the case in which the production process is sequential in nature and the relationship-specific investments made by suppliers in upstream stages can affect the incentives of parties involved in later downstream stages. As a result, they show that  $F$  might have differential incentives to integrate suppliers along the value chain, and might end up outsourcing some inputs and integrating others, even when the production function treats these inputs symmetrically as in equation (11). In other words, the “downstreamness” of an input becomes a determinant of the ownership structure decisions related to that input. Whether upstream stages are more or less likely to be integrated than downstream stages turns out to depend critically on the relative size of the parameters  $\eta_m$

18. Conversely, the effect of  $\eta_m$  on the relative profitability of integration and outsourcing is less clear-cut. The reason for this is that as  $\eta_m$  goes down, the relative importance of suppliers' investment goes down, but their bargaining strength is also diminished, thereby aggravating the hold-up problem. It can be shown, however, that in the neighborhood of  $\eta_m \rightarrow 0$ , integration is necessarily more profitable than outsourcing.

and  $\rho$ . When inputs are sufficiently close complements ( $\rho < \eta_m$ ), the optimal choice involves the outsourcing of upstream stages and the integration of downstream stages, while the converse is true when inputs are sufficiently close substitutes ( $\rho > \eta_m$ ).<sup>19</sup>

**2.3.3 Partial Contractibility.** The models above have assumed that none of the aspects of production, except the allocation of ownership rights and a lump-sum transfer, are contractible before productive investments have taken place. This is obviously an unrealistic assumption even in the international environments we will be considering shortly, so it is important to discuss the implications of relaxing it. For that purpose, and to simplify matters, let us go back to the case in which there is a unique supplier with whom  $F$  contracts and revenue is given by the simpler function in equation (6). Following Acemoglu et al. (2007) and Antràs and Helpman (2008), we now allow the inputs  $h$  and  $m$  to be produced by combining a set of input-specific components or services  $\{x_j(i)\}_{i \in [0,1]}$  for  $j = h, m$ , each at a marginal cost  $c_j$ , according to

$$j = \exp \left[ \int_0^1 \log x_j(i) di \right], \quad j = h, m. \quad (14)$$

To capture partial contractibility, assume that the components related to input  $j$  in the range  $[0, \mu_j]$ ,  $[0, \mu_j]$ ,  $0 \leq \mu_j \leq 1$ ,  $j = h, m$ , are now contractible in the sense that the characteristics of these activities can be fully specified in advance in an *enforceable* ex ante contract (though these investments take place at  $t = 2$ , simultaneously with the noncontractible ones). Notice that the parameters  $\mu_h$  and  $\mu_m$  capture the level of contractibility of headquarter services and components, respectively. Because the terms of exchange of some of the inputs are not determined by the initial contract,  $F$  and  $M$  will again negotiate the price of those exchanges at  $t = 2$ . Even though each party is bound to provide the contractually stipulated levels for the contractible components, they can threaten to withhold part of the noncontractible ones, which in light of the specification in equation (14) can significantly impact revenue. Following GH and in analogy with the benchmark model above, assume that suppliers can entirely withhold those noncontractible components in case of a contractual breach when  $F$  outsources the production of  $m$  to  $M$ . Conversely, in the case of integration,  $F$  can selectively fire  $M$  in case of a contractual disagreement and use the components in production (since it has ownership rights over them), but as in the benchmark model we continue to assume that this entails an efficiency loss of a fraction  $1 - \delta$  of revenue.

19. Intuitively, outsourcing elicits high levels of investment from upstream suppliers. The complementarity of upstream with downstream inputs in the case  $\rho < \eta_m$  in turn alleviates the underinvestment problem for downstream suppliers, and the firm introduces fewer distortions by integrating downstream to retain a larger share of the realized output and enhance the investments in headquarter services.

Given the symmetry in the model, it is straightforward to verify that in equilibrium there will be symmetric investments  $x_{jn}$  in all noncontractible activities and  $x_{jc}$  in all contractible activities for  $j = h, m$ . This allows writing the revenue function as

$$R(h, m) = Ax_{hc}^{\mu_h \eta_h} x_{mc}^{\mu_m \eta_m} x_{hn}^{(1-\mu_h)\eta_h} x_{mn}^{(1-\mu_m)\eta_m} = \tilde{A} x_{hn}^{(1-\mu_h)\eta_h} x_{mn}^{(1-\mu_m)\eta_m}.$$

Clearly, this revenue function is identical to equation (6) except that the revenue shifter  $A$  is now replaced by  $\tilde{A} = A(x_{hc})^{\mu_h \eta_h} (x_{mc})^{\mu_m \eta_m}$ , while  $\eta_h$  and  $\eta_m$  are now replaced by  $(1 - \mu_h)\eta_h$  and  $(1 - \mu_m)\eta_m$ , respectively. From the analysis above, it is then clear that the integration decision now depends on the relative magnitude of the terms  $(1 - \mu_h)\eta_h$  and  $(1 - \mu_m)\eta_m$ , that is, on the relative intensity of the noncontractible inputs provided by  $F$  and  $M$ , respectively. Furthermore, from equation (7), the (hypothetical) share of revenue that would optimally be allocated to  $F$  is given by

$$\frac{\beta^*}{1 - \beta^*} = \sqrt{\frac{(1 - \mu_h)\eta_h / (1 - (1 - \mu_h)\eta_h)}{(1 - \mu_m)\eta_m / (1 - (1 - \mu_m)\eta_m)}}, \quad (15)$$

and is increasing in  $\eta_h$  and  $\mu_m$ , and decreasing in  $\eta_m$  and  $\mu_h$ .<sup>20</sup> As Antràs and Helpman (2008), one can show the following sharper result:

*Proposition 5.* There exists a unique headquarter-intensity cutoff  $\hat{\eta}_h \in (0, 1)$  such that profits are higher under outsourcing for  $\eta_h < \hat{\eta}_h$  and higher under integration for  $\eta_h > \hat{\eta}_h$ . Furthermore, the cutoff  $\eta_{hc}$  is higher the larger  $\mu_h$  is and the smaller  $\mu_m$  is.

This result highlights that, with partial contractibility, the integration decision not only depends on the level of headquarter intensity, but also on the degrees of contractibility of the different inputs, with distinct effects for different types of inputs. An improvement in contracting possibilities for components enhances integration (since  $F$  is less dependent on the incentive effects of outsourcing to elicit investments from  $M$ ), while an improvement on the contractibility of headquarter services has the opposite effect and makes integration less attractive.

**2.3.4 Organizational Fixed Costs and Producer Heterogeneity.** The frameworks we have considered so far treat the decisions of  $F$  and its suppliers independently of the decisions of other firms in  $F$ 's industry. In reality, firms within industries tend to face similar demand conditions and tend to operate with similar production technologies. Yet, even within narrowly

20. It may seem surprising that we do not need to concern ourselves with solving for the level of contractible investments and computing overall profits in order to determine the optimal bargaining share in equation (15). Note, however, that the level of contractible investments is irrelevant for the choice of  $\beta^*$  because these investments are set at  $t = 0$  to maximize joint surplus and thus the envelope theorem washes out these terms in the first-order condition in equation (2).

defined industries, we observe significant variation in the organizational decisions of firms, including the location of their production stages but also the extent of control they want to exert over those processes. What explains these differences? I next follow [Antràs and Helpman \(2004\)](#) in showing how the mere existence of either heterogeneity in the (Hicks-neutral) productivity level or in the demand level faced by firms can give rise to heterogeneous organizational decisions among firms sharing a common headquarter intensity level. A useful implication of generating such heterogeneity is that the key parameters of the model not only affect the incentives of firms to integrate or not particular production processes (something that is rarely observed in the data), but they now also lead to smooth changes in the set of firms choosing different organizational modes, thus generating comparative static predictions for how changes in these parameters affect *aggregate* measures of the prevalence of integration in particular sectors.

In order to illustrate these features, I consider two simple modifications of the above framework. First, the revenue function now includes a productivity (or demand shifter)  $\theta$ ,

$$R(h, m) = A\theta h^{\eta_h} m^{\eta_m}, \quad (16)$$

so that firms with higher values of  $\theta$  generate larger amounts of revenue for a given size of investments in  $h$  and  $m$ , perhaps because they combine these inputs more efficiently or perhaps because consumers are particularly keen on the manner in which these inputs are combined. Let  $\theta$  be distributed across firms according to the cumulative density function  $G(\theta)$ . The second new feature is the presence of organizational fixed costs  $f_k$ ,  $k = V, O$ , that vary by ownership structure. Following [Antràs and Helpman \(2004, 2008\)](#), I focus on the plausible case in which integration entails higher fixed costs than outsourcing,  $f_V > f_O$ , though most of the results discussed in this survey only require that these organizational fixed costs be distinct from each other.<sup>21</sup> With these assumptions it is straightforward to verify the following result (see [Antràs and Helpman 2004, 2008](#), for a proof):

*Proposition 6.* There exists a unique headquarter-intensity cutoff  $\hat{\eta}_h \in (0, 1)$  and a unique productivity level  $\hat{\theta}$ , such that if  $\eta_h < \hat{\eta}_h$  all firms outsource the production of  $m$ , while if  $\eta_h > \hat{\eta}_h$ , all firms with  $\theta < \hat{\theta}$  outsource the production of  $m$ , while all firms with  $\theta > \hat{\theta}$  firm integrate it. In the latter case, the share of firms integrating the production of  $m$  is increasing in the level of headquarter intensity and decreasing in the level of component intensity.

21. There is an obvious tension between this assumption and the spirit of GH's approach that associates different ownerships structures only with different allocations of assets, while holding technology and contracting constant. Still, differences in these organizational fixed costs are likely to be relevant in practice so the literature has incorporated them into the analysis.

Intuitively, even though integration might be the organizational mode that maximizes operating profits whenever  $\eta_h > \hat{\eta}_h$ , firms with relatively low revenue levels might not be able to amortize the relatively high fixed costs associated with integration and are thus left to obtain input  $m$  via outsourcing contracts.

### 3. The Boundaries of Multinational Firms

In this section, I overview how the property-rights theory has been used to draw multinational firm boundaries and thereby shed light on important aspects of the international economy. I first discuss how the benchmark partial-equilibrium model developed above is modified by the possibility of international trade across borders and later discuss how the framework can be embedded in industry equilibrium and also in general equilibrium. The bulk of the section is devoted to describing the novel results that emerge from such an analysis.

We begin by considering an  $N$ -country version of the benchmark model in the previous section, but now allow  $F$  to locate different parts of the production process in different countries. The production process we described above entails three stages (production of  $h$ ,  $m$ , and the final good), thus giving rise to several potential locational decisions. Let us denote by  $L$  the set of possible locational decisions and by  $\ell \in L$  a particular one. For example,  $\ell$  could entail production of headquarter services and of the final good in  $F$ 's Home country and production of  $m$  in a foreign country. Notice that different location choices will in general entail different values for the key parameters of the model. In particular, trade barriers (technological or man-made), and cross-country differences in production costs and in institutions imply that different locational choices can be associated with different values for the parameters  $c_h$ ,  $c_m$ ,  $f_V$ ,  $f_O$ , and  $\beta_V$ , as well as for the revenue function  $R(h, m)$ . How do these generalizations affect the way firms organize production? In analogy with (P1), the optimal ownership structure  $k^*$  and the optimal locational choice  $\ell^*$  now solve the following program:

$$\begin{aligned} \max_{k \in \{V, O\}, \ell \in L} \quad & \pi_k^\ell = R(h_k^\ell, m_k^\ell) - c_h^\ell \cdot h_k^\ell - c_m^\ell \cdot m_k^\ell - f_k^\ell \\ \text{s.t.} \quad & h_k^\ell = \arg \max_h \{ \beta_k^\ell R(h, m_k^\ell) - c_h^\ell \cdot h \} \\ & m_k^\ell = \arg \max_m \{ (1 - \beta_k^\ell) R(h_k^\ell, m) - c_m^\ell \cdot m \} \end{aligned} \quad (\text{P2})$$

Notice that even in this stylized model the cardinality of the firms' choice set can be very large. There are three potential production stages, two possible ownership structures and  $N$  countries, thus giving rise to at least  $2N^3$  possible organizational modes.<sup>22</sup> For  $N=5$ , this generates 250 possible combinations, while for  $N=100$ , there are two million

22. This assumes that headquarter services and final-good production are always integrated (which is consistent with the property-rights theory since  $F$  is essential for those



combinations. To reduce the dimensionality of the problem, the literature has typically followed one of two approaches. One of them involves imposing symmetry across countries in most parameters of the model, while the other one entails focusing on two-country environments. Below, I will sketch an example of each of these two approaches.

### 3.1 Headquarter Intensity and Comparative Advantage

Antràs (2003) considers a general-equilibrium model of trade in which consumers in  $N$  countries have identical preferences and spend a constant share of their income on a continuum of differentiated varieties in two sectors  $Y$  and  $Z$ . The setup is similar to the transaction-cost model in Grossman and Helpman (2002). Identical Dixit–Stiglitz subutility functions in each sector give rise to a demand function for a particular variety  $\omega$  in sector  $s$  of the form

$$q_s(\omega) = \frac{\mu_s E}{\int_0^{n_s} p_s(\omega)^{-\alpha/(1-\alpha)} dj} p_s(\omega)^{-1/(1-\alpha)} = A_s p_s(\omega)^{-1/(1-\alpha)}, \quad (17)$$

where  $\mu_s$  is the share of aggregate spending  $E$  spent in sector  $s$ ,  $n_s$  is the measure of varieties available to consumers, and  $\alpha$  governs how substitutable varieties are within sectors. Because firms take their demand shifter as given, firm revenue can simply be expressed as  $R_s(\omega) = A_s^{1-\alpha} q_s(\omega)^\alpha$  (see equation [17]), where  $A_s$  is treated as a constant by firms. Production of differentiated varieties is as described in Section 2, with final-good production of each variety being controlled by a different manager  $F$  who is also in charge of providing headquarter services but needs to contract with a production plant manager  $M$  for the provision of input  $m$ . Production of inputs can be located in any of the  $N$  countries in the world. Antràs (2003) focuses on the case in which  $h$  and  $m$  produce output according to a constant-returns-to-scale Cobb–Douglas production function, with the elasticity of output to headquarter services given by a sector-specific constant, which I denote by  $\eta_s$ . The production technology is identical for all firms within an industry. Notice that firm-level revenue takes the isoelastic form in equation (6).

To simplify the complexities inherent in the general problem (P2), Antràs (2003) considers the case in which countries differ *only* in their relative factor endowments and thus in their (autarky) factor costs. In particular, he rules out trade costs across countries, lets contract incompleteness and the efficiency loss parameter  $\delta$  be identical in all countries, and also assumes that organizational fixed costs are independent of ownership structure and feature the same factor intensity as variable costs (i.e., they combine  $h$  and  $m$  under the same Cobb–Douglas aggregator as these enter the firm’s production function).

---

stages), and that the production of  $m$  takes place in only one country. Relaxing these assumptions would only increase the complexity of the firm problem.

The combination of these assumptions makes the problem (P2) particularly easy to solve because the ownership structure and location decisions can be treated independent from each other. In particular, the ownership structure decision is characterized by Proposition 2 above with  $F$  managers *worldwide* choosing to integrate their suppliers if  $\eta_s > \hat{\eta}$ , and outsource to them if  $\eta_s < \hat{\eta}$ , the key being that  $\hat{\eta}$  is independent of production costs (and thus factor prices). On the other hand, the location decision boils down to choosing the location of input production that minimizes the marginal cost of provision of inputs. Antràs (2003) assumes that when  $F$  invests in  $h$  for a production plant located in country  $\ell$ , it needs to hire local factors  $\ell$  and thus faces the production costs in that country, thus implying that the location decision simply solves  $\min_{\ell} \{ (c_h^{\ell})^{\eta_s} (c_m^{\ell})^{1-\eta_s} \}$ .<sup>23</sup>

In order to complete the description of the general-equilibrium of the world economy, one needs to specify the market structure in each industry as well as the factors of production that firms employ in producing inputs. Antràs (2003) considers a Helpman–Krugman (1985) model with monopolistic competition and free entry driving profits down to zero.<sup>24</sup> He assumes that inputs are produced with physical capital and labor and that the production of headquarter services is more capital intensive than that of components. The latter is the key assumption of the article as it introduces a positive correlation between the abstract concept of headquarter intensity and an observable variable, namely capital intensity. Antràs (2003) justifies this assumption on empirical grounds, arguing that cost-sharing practices of multinational firms in their relations with independent subcontractors tend to be associated with physical capital investments rather than with labor input choices. We shall come back to the plausibility of this assumption in Section 4.

Given the positive association between headquarter intensity and capital intensity, Proposition 2 then implies a higher prevalence of integration in capital-intensive sectors. To complete the characterization of the equilibrium, one need only impose that factor markets clear country by country and that world income equals world spending. Antràs (2003) shows that if relative factor endowments are not too different across countries, free trade will equalize factor prices worldwide, but aggregate production patterns and bilateral trade flows across countries will be fully determined.

23. The fact that technology involves increasing returns to scale ensures that, with free trade, each input will be produced in only one location.

24. In order to ensure that no agent earns rents in equilibrium, Antràs (2003) allows  $F$  managers to make take-it-or-leave-it offers to  $M$  managers at  $t=0$ . The ex ante transfer between  $F$  and  $M$  thus ensures that  $M$  managers end up with a zero net payoff. On the other hand, free entry by  $F$  managers implies that the expected operating profits from producing a differentiated variety in a given industry exactly cover the fixed organizational costs associated with that industry. An alternative approach is developed in Grossman and Helpman (2002), where a process of search is specified and free entry leads all agents to anticipate a zero net payoff *in expectation* despite the absence of ex ante transfers between agents.

For example, labor-abundant countries will end up with the same wage level as capital-abundant countries, but factor-market clearing ensures that they will end up producing a disproportionate amount of the world-wide production inputs in the labor-intensive industry.

Antràs (2003) derives explicit formulas for overall bilateral trade flows as well as their intrafirm component and highlights two main predictions from the model. First, in a cross-section of industries, the share of intrafirm imports in total imports should be increasing in the capital intensity in production of the exporting industry.<sup>25</sup> Second, in a cross-section of countries, the share of intrafirm imports in total imports should be increasing in the aggregate capital–labor ratio of the exporting country, as labor-abundant countries tend to export small amounts of capital-intensive goods.

### 3.2 Heterogeneity and Global Sourcing

In the model in Antràs (2003), the organizational decisions of firms in the world economy and their implications for certain aggregate variables, such as the intrafirm component of trade across industries and countries, are easy to characterize due to the strong symmetry assumptions made regarding the nature of contracting across countries. These “tractability” assumptions are the bread and butter of applied theorists seeking to construct models to understand certain qualitative features of the world. Yet these same assumptions often result in models that are not particularly useful tools for empirical analysis. As an example, Antràs’ (2003) model clarifies that a variable that is well known to be key in understanding the pattern of trade, namely capital intensity, may also be a key determinant of the integration decision of multinational firms and of the intrafirm component of trade, but it does so by making assumptions that render the model unsuitable for empirical analyses of the global sourcing strategies of firms or for empirical studies of the *various* determinants of the share of intrafirm trade.

In this section, I briefly describe the frameworks in Antràs and Helpman (2004, 2008), which incorporate several sources of heterogeneity while keeping the analysis manageable by focusing on a particularly simple two-country, North–South model. The key elements of the model are as follows. Consumers in both countries demand the output of one homogenous-good sector and  $J$  differentiated-good sectors. Preferences are quasilinear in the homogenous good and feature a constant elasticity of substitution between differentiated varieties within a sector and also between varieties in different sectors (with the latter elasticity assumed higher). This preference structure delivers a demand function for a particular variety in industry  $j$  that is analogous to that in equation (17), but

25. Strictly speaking, the model predicts that this share should be 0 for all industries with capital intensity below  $\hat{\eta}$  and 1 for all industries with capital intensity above this threshold.

with a demand shifter that is only a function of the aggregate consumption in the sector.

Firm behavior is a variant of the general program (P2). It is assumed that all  $F$  managers reside in the North and that the final good and headquarter services are always produced in that country. The location decision thus reduces to the choice of where to produce  $m$ , i.e.,  $\ell \in \{N, S\}$ . Producers in each country face a perfectly elastic supply of a unique factor of production, labor. Wage rates are fixed in general equilibrium by the “outside” homogenous good sector and technology in that sector is such that  $w^N > w^S$ . The final good is produced according to a Cobb–Douglas technology in  $h$  and  $m$  that features a productivity shifter  $\theta$ . Given these assumptions, the revenue function is of the type introduced above in equation (16). The productivity parameter  $\theta$  is firm specific and drawn from a Pareto distribution with shape  $z$ , i.e.,  $G(\theta) = 1 - (b/\theta)^z$  for  $\theta \geq b > 0$ , while the elasticity of output with respect to  $h$ , denoted again by  $\eta$ , is common to all firms within a sector, but may vary across sectors. Production of intermediate inputs requires one unit of labor per unit of output in the country where they are produced. International trade in components is costly and  $\tau$  units of  $m$  need to be shipped from the South for one unit to arrive to the North. Provided that this transport cost is low enough, these assumptions imply that  $c_h^N = c_m^N = w^N > \tau w^S = c_m^S$ .

Production also involves different types of fixed costs, which are all defined in terms of Northern labor. First,  $F$  needs to incur a fixed cost  $f_E$  of entry, upon which the productivity parameter  $\theta$  is revealed to him or her, as in Melitz (2003). If  $F$  decides to remain in the market, additional fixed organizational costs need to be incurred. As discussed above and in more detail in Antràs and Helpman (2004), these fixed organizational costs are likely to vary depending on whether  $m$  is sourced in the North or in the South, and on whether it is insourced or outsourced. In particular, a natural ranking of these fixed costs is

$$f_V^S > f_O^S > f_V^N > f_O^N.$$

In words, fixed organizational costs are higher when  $M$  is located in the South regardless of ownership structure, and given the location of  $M$ , the fixed organizational costs are higher when  $M$  is integrated than when it is not.

In terms of the contracting parameters, Antràs and Helpman (2004) consider environments in which no aspects of  $h$  and  $m$  are contractible regardless of where production takes place but there is cross-country variation in the efficiency loss parameter  $\delta$ , with  $\delta^N > \delta^S$ , and thus  $\beta_V^N > \beta_V^S$ , reflecting better legal protection in the North. Antràs and Helpman (2008) extend the analysis to incorporate partial contractibility (along the lines of the model developed in Section 2.3.3), and allow the degree of contractibility to be a function of both the type of input and the country where production takes place.

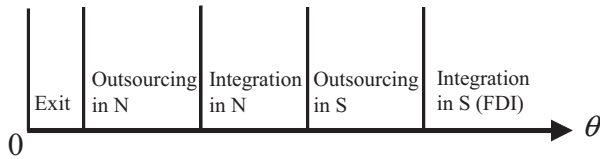


Figure 1. Sorting in a Headquarter Intensive Sector (Antràs and Helpman 2004, 2008).

Given the assumptions of the model, the choice of an organizational form faces two types of tensions. In terms of the location decision, the South entails relatively lower variable costs, but relatively higher fixed costs. As in the work of Melitz (2003), it is clear that the firm-specific productivity parameter  $\theta$  will crucially affect the firm's participation in foreign sourcing. In terms of the ownership structure decision, integration improves efficiency of variable production when the intensity of headquarter services is high (as implied by Proposition 2), but involves higher fixed costs. Hence, the integration decision will crucially depend on both  $\eta$  and  $\theta$ .

Antràs and Helpman (2004) show that the model can easily generate equilibria featuring multiple organizational forms within an industry. In particular, in sufficiently headquarter-intensive sectors it is possible for the least productive firms to exit the market upon observing their productivity, and for four nonempty (and connected) subsets of the remaining firms to choose each of the four possible organizational forms. The most productive firms in the North engage in foreign insourcing (or FDI in the South), the next most productive firms undertake foreign outsourcing, the next subset insource domestically, and the least productive firms among the surviving ones outsource domestically. This pattern is illustrated in Figure 1.

Antràs and Helpman (2004) also use the model to study the various determinants of the relative prevalence of these different organizational forms. As hinted above, the predictions of their model move well beyond those derived in Antràs (2003) and have served as the springboard for empirical studies of the determinants of the global sourcing strategies of firms. In particular, the model predicts that, in a cross-section of industries, the share of intrafirm imports of components in total imports of components should be higher in industries with higher headquarter intensity (higher  $\eta$ ), higher productivity dispersion (lower  $z$ ), and higher transport costs or import tariffs (higher  $\tau$ ). Furthermore, these parameters shape the relative prevalence of domestic insourcing and domestic outsourcing in the same manner (even for the case of trade frictions  $\tau$ ).

When extending the model to allow for partial contractibility, Antràs and Helpman (2008) find that an improvement in contractual institutions in South raises the share of firms engaged in offshoring in that country, but it can reduce the relative prevalence of either FDI or offshore

outsourcing if it affects disproportionately the contractibility of headquarter services or components, respectively. This result relates to our previous Proposition 5, which highlighted that the relative prevalence of alternative organizational forms crucially depends on the degree to which contractual incompleteness affects the inputs controlled by the final-good producer or by his or her suppliers.

### 3.3 Brief Overview of Other Contributions

Before discussing the empirical implementation of the property-rights theory of multinational firm boundaries, I briefly overview other significant contributions to the literature and I later identify what I view as important gaps in the literature.

In Section 2.3.1, I discussed the effects of financial or liquidity constraints on the integration decision. Naturally, the results derived there have implications for how firm boundaries are pinned down in open-economy environments. For instance, [Basco \(2010\)](#) and [Carluccio and Fally \(forthcoming\)](#) develop general-equilibrium, open-economy models in which, consistently with Proposition 3, multinationals are more likely to integrate suppliers located in countries with poor financial institutions. Furthermore, both papers predict that the effect of financial development should be especially large when trade involves complex goods, and both provide independent empirical evidence supporting this prediction.

As emphasized by [Legros and Newman \(2008\)](#), in the presence of financial constraints, equilibrium firm boundaries will also depend on the relative ex ante bargaining power of each party and their ability to exchange lump-sum transfers. This idea has been fruitfully applied in open-economy environments by [Conconi et al. \(forthcoming\)](#) and [Alfaro et al. \(2010\)](#) who show that vertical integration should be relatively more prevalent in industries in which (relative) prices are high, perhaps due to import-protecting trade policies. Intuitively, in their setup, which builds on [Hart and Holmström \(2010\)](#) and [Legros and Newman \(2009\)](#), ownership decisions are not ex ante optimal, but instead trade off the pecuniary benefits of coordinating production achieved under integration and the managers' private benefits of operating in their preferred ways associated with nonintegration. Consequently, the higher the industry price, the higher are the monetary benefits of integration and thus the more attractive this option is. [Alfaro et al. \(2010\)](#) provide evidence of a positive association between import tariffs and domestic integration decisions. [Diez \(2010\)](#) finds similar evidence in a cross-section of US industries when looking at intrafirm trade flows, but interprets the result in light of the [Antràs and Helpman \(2004, 2008\)](#) models, which as mentioned above, also predicts a positive effect of imports tariffs on foreign integration. I will illustrate his empirical results in Section 4.

The insights of the property-rights theory have also been applied to dynamic, general-equilibrium models of international trade with the goal of understanding how ownership decisions vary along the lifecycle



of a product or input. Antràs (2005), for instance, develops a model in which the incomplete nature of contracts governing international transactions limits the extent to which the production process can be fragmented across borders, thereby generating the emergence of Vernon-type product cycles, with new goods being initially manufactured in North (where product development takes place), and only later (when the goods are mature) is manufacturing carried out in South. Antràs (2005) also draws the boundaries of multinational firms and shows that the model gives rise to a new version of the product cycle in which, consistently with empirical evidence, manufacturing is shifted to the South first within firm boundaries, and only at a later stage to independent firms in the South.<sup>26</sup>

Throughout this section, I have restricted myself to reviewing papers that can be considered “direct offspring” of GH in the sense that they adopt variants of the property-rights approach to drawing firm boundaries in open-economy environments. As I will spell out in more detail in Section 5, the incomplete-contracting framework of GH has had a much broader impact in the international trade field via the influence of other theoretical frameworks that are themselves ‘direct offspring’ of GH. For instance, in the presence of incomplete contracts, another important organizational decision of firms concerns the allocation of decision rights among employees. In particular, in the presence of noncontractible effort decisions by workers, managers face a trade-off between granting decision rights to workers or keeping these to themselves. The former option has the benefit of providing workers with “initiative,” which may lead to higher effort, but delegation may result in decisions that are not necessarily optimal from the point of view of the manager. Avoiding delegation (i.e., exerting “authority”) tends to inhibit the initiative of workers but entails more control over the course of production. This trade-off was first formalized by Aghion and Tirole (1997) and has been applied to general-equilibrium frameworks by Marin and Verdier (2009) and Puga and Trefler (2010).

### 3.4 The Road Ahead

In the year 2036, we will be celebrating the 50th anniversary of Grossman and Hart’s article. How will GH shape the field of international trade in the next 25 years? As I will try to convey in the next section, there are good reasons to believe that the bulk of new work on multinational firm boundaries will be empirical in nature. I believe, however, that the theoretical frameworks that have been developed to date are still too rudimentary to be taken to the data in their current form. To a large extent this is due

26. Despite the dynamic nature of the model, Antràs (2005) assumes that the game played by managers can be treated as a static one and thus abstracts from an analysis of reputational equilibria. Corcos (2006) studies such dynamic contracts by applying the relational contract approach of Baker et al. (2002).

to the fact that these models were designed to guide empirical work based on industry-level data, while future work is more likely to make use of firm-level data sets. The theoretical frameworks in [Antràs and Helpman \(2004, 2008\)](#) generate a rich set of predictions, but they are fairly low-dimensional (featuring one supplier per firm and only two countries) and impose substantial symmetry in order to highlight a few key industry-level predictions. For example, the assumption that headquarter intensity varies across industries but not across firms within an industry is clearly counterfactual (see [Corcos et al., forthcoming](#), for French evidence). Future models should also incorporate multiple inputs and suppliers as in the model developed in Section 2.3.2, while allowing for some heterogeneity among these inputs.<sup>27</sup> Introducing such new sources of heterogeneity into the sort of partial equilibrium models that we started with in Section 2.1 is relatively straightforward. The challenge for future theoretical work is to introduce these features in a way that still permits characterizing the open-economy general equilibrium of such models, a feature that might be particularly relevant for performing counterfactual exercises when one is estimating these models structurally.

Another potentially fruitful area for future research relates to the study of the effects of the nonexcludable nature of knowledge on the internalization decision. Following GH, existing work has arguably focused too much on the distortionary effect of contractual incompleteness on ex ante investments. In practice ex post inefficiencies, including the possibility of technological expropriation by suppliers or licensees are often highlighted as being key to the internalization decisions of multinational firms. Previous attempts to incorporate a notion of nonexcludable knowledge into general equilibrium models of the multinational firm have adopted a transaction-cost approach and thus shed little light on how foreign integration circumvents the dissipation of knowledge to local producers. A more satisfactory approach would entail the application of the property-rights approach and of the notion of “access” developed by [Rajan and Zingales \(2001\)](#).<sup>28</sup>

#### 4. Taking the Property-Rights Theory to International Data

As I argued in the Introduction, the growing importance of international production networks, involving transactions both within and across multinational firms, is a key factor in understanding the recent impact of the property-rights theory in the field of international trade. Although the trigger for the development of new theoretical models was empirical in nature, it is fair to say that this branch of the literature has yet not come full circle in providing definitive empirical tests of these models.

27. The recent papers by [Antràs and Chor \(2011\)](#); [van Biesebroeck and Zhang \(2011\)](#); [Schwarz and Suedekum \(2011\)](#) constitute first steps in that direction.

28. See [Chen et al. \(2008\)](#) and [Ponzetto \(2009\)](#) for early attempts along these lines.

Several well-crafted papers have offered different pieces of evidence that are consistent with the property-rights theory, but the power of such tests remains fairly low.

In great part, these limitations are due to the fact that empirically validating the property-rights theory poses at least two important challenges. First, data on the integration decisions of firms are not readily available, and thus researchers are often left to test these theories with industry- or product-level data, which do not always allow one to appropriately control for unobservable characteristics of firms that might be driving the patterns observed in the data. Second, the predictions from the property-rights theory are associated with subtle features of the environment (such as the relative value of the marginal return to noncontractible, relationship-specific investments) that, by their own nature, are generally unobservable in the data (see [Whinston 2003](#)).

Admittedly, the contributions in the international trade literature have not made too much progress regarding this second hurdle, though the issue has not been completely ignored, as I will highlight later in this section. With regard to the first challenge on data availability, however, an advantage of researchers in our field is that data on international transactions are particularly accessible due to the widespread existence of official records of goods and services crossing borders. For instance, there exist fairly detailed data on US intrafirm trade at the six-digit Harmonized System classification of the product shipped (of which there are over 5000 categories) and at the origin/destination country level. This amounts to hundreds of thousands of observations *per year* on the relative prevalence of integration versus nonintegration. In the next section, I will discuss some of the pros and cons to using these data to test the property-rights theory and will also graphically illustrate some of the empirical patterns that emerge from the data and how they relate to the models we discussed above.

Beyond these product-level datasets from official statistics, a few researchers have made use of firm-level data sets (with different levels of representativeness) that contain detailed information on the sourcing strategies of firms in different countries. In the second part of this section, I will provide more details on these data sets and on how they have been used to test the property-rights theory of the multinational firm. I will conclude this section by offering some thoughts on future avenues for empirical research in the area.

#### 4.1 Tests with Product-Level Intrafirm Trade Data

A significant share of empirical tests of the property-rights theory of the multinational firm have used data from official import and export merchandise trade statistics, which in some countries identify whether transactions involve related or nonrelated parties. Most tests use the “US Related Party Trade” data collected by the US Bureau of Customs and Border Protection, so I will focus on discussing some of the

advantages and disadvantages that are associated with using this data source. Later, I will briefly discuss some special features of a similar data set compiled by the Customs General Administration of the People's Republic of China, which has also been used in the literature.

4.1.1 Pros and Cons of Product-Level Sources. Several features of these official statistics make them particularly attractive to empirical researchers. First, for some countries, notably the United States, intrafirm trade data are publicly available and easily downloadable from government websites.<sup>29</sup> Second, the data are of high quality and are not subject to sampling error, since several quality assurance procedures are performed, and the data offer a complete picture of the sourcing strategies of firms in a country. Third, there is a large amount of variation in the data: the share of US intrafirm imports over total US imports varies widely across products and origin countries, and there also exists significant variation in that share across products within exporting countries and across exporting countries within narrowly defined products. This is an important feature of the data so I will document it below. Fourth, by including information on all industrial sectors, these data make it easier to spot certain *fundamental* factors that appear to shape whether international transactions are internalized or not independently of the sector one studies. A fifth advantage of using these comprehensive data sets is that by covering a wide range of sectors, countries and time periods, they offer the potential to exploit exogenous changes in sector characteristics (due perhaps to technological change) or in institutional characteristics of exporting or importing countries (due, for instance, to institutional reforms) to better identify some of effects predicted by the property-rights theory.

It is important, however, to also be upfront about some of the limitations of using this type of data. First, there is an obvious tension in using product-level data, which aggregate the decisions of various firms, to test the validity of theories of firm boundaries. Although some of the available data are sufficiently disaggregated to ensure that each observation aggregates the transactions of only a handful of firms (or perhaps even just one firm), the inability to control for idiosyncratic firm-level characteristics remains problematic. Second, the data are reported based on the sector or industry category of the good being transacted and do not contain information on the sector that is purchasing the good or on whether the good is an intermediate inputs or a final good (though the level of

29. The U.S. data are publicly available at: <http://sasweb.ssd.census.gov/relatedparty/>. This website permits downloading the data at the six-digit NAICS level. The finer six-digit Harmonized System (HS) data are available from the US Census for a fee. Other researchers (e.g., Zeile 1997; Antràs 2003) have constructed intrafirm trade using direct investment data from the US Bureau of Economic Analysis or BEA (see [http://www.bea.gov/iTable/index\\_MNC.cfm](http://www.bea.gov/iTable/index_MNC.cfm)), but the publicly available data do not feature a fine industry disaggregation. The underlying firm-level data used to construct this data set are available only to researchers affiliated with the BEA and have not been used to test the property-rights theory.

disaggregation of the data often allows to make informed choices about whether or not the goods are inputs). Third, in related party transactions, the data do not typically report which firm is owned by whom, i.e., whether integration is backward or forward, and also do not provide information on the extent of control (or ownership share) of the parent company.<sup>30</sup> A fourth and final concern in using these data to study the global sourcing strategies of firms based in a given country (say the United States) is that this country's trade statistics will only capture those sourcing decisions that entail goods being shipped back to that country, while in practice some large firms have production networks in which parts and components are shipped across foreign locations (within and across firm boundaries) and then only shipped back to the home country after being assembled abroad (as is the case of the iPad 2 discussed in the Introduction). For this reason, US intrafirm imports generally underrepresent the involvement of US multinational firms in global sourcing strategies, though it is not obvious how this phenomenon biases the results of empirical studies using these data.

**4.1.2 Some Features of US Intrafirm Trade Data.** Before I discuss in detail how US intrafirm trade data have been employed to test the property-rights theory, it is worth providing some brief descriptive statistics that demonstrate the empirical relevance of intrafirm trade and illustrate how the share of intrafirm trade varies across products and countries. Throughout the section, I focus on data on US imports of goods for the year 2010, except when using the more disaggregated six-digit Harmonized System data, for which the most recent year available (to me) is 2005.

In 2010, intrafirm imports of goods totaled \$922 billion and constituted a remarkable 48.6% of total US imports of goods (\$1899 billion), thus indicating the importance of multinational firms for US trade. The share of US intrafirm imports varies widely across countries. On the one end, intrafirm imports equal 0 for 13 countries and territories (including Cuba and North Korea), which all record very low volumes of exports to the United States, while the share of intrafirm trade reaches a record level of 98.9% for US imports from Mauritius. Leaving aside communist

30. The US data define related-party trade as including import transactions between parties with various types of relationships including “any person directly or indirectly, owning, controlling or holding power to vote, 6 percent of the outstanding voting stock or shares of any organization.” In practice, extracts from the confidential direct investment data set collected by the BEA suggest that intrafirm trade is generally associated with one of the entities having a controlling stake in the other entity. Furthermore, the same source indicates that about two-thirds of total US intrafirm imports are accounted for by imports shipped by overseas affiliates to their US parents, rather than by imports shipped to US affiliates by their foreign parent group. Nunn and Treffer (2008) use information from this same data set to restrict the sample to countries for which at least two-thirds of intra-firm US imports from the country are imported by US parents.

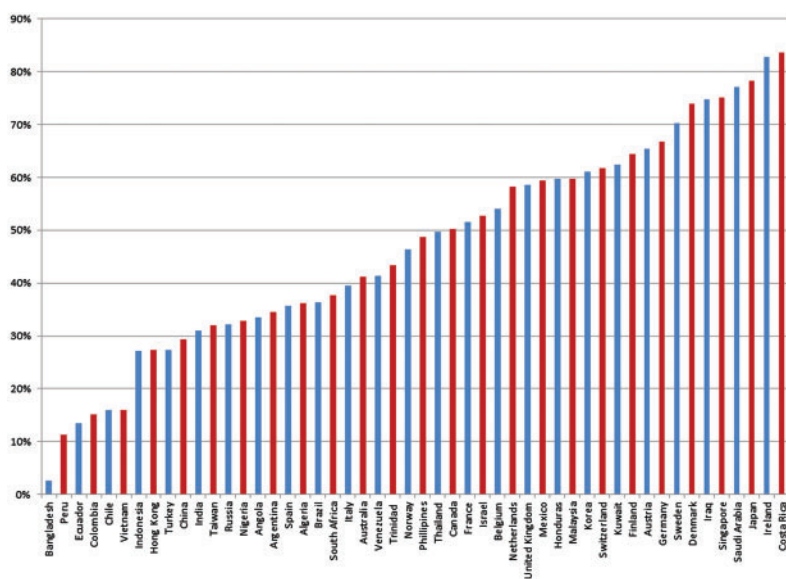


Figure 2. Share of US Intrafirm Imports for Largest 50 US Exporters in 2010.

dictatorships and tropical islands and focusing on the 50 largest exporters to the United States, Figure 2 illustrates that the share of intrafirm trade still varies significantly across countries, ranging from a mere 2.7% for Bangladesh to 83.6% for Costa Rica.

Similarly, the share of intrafirm trade varies widely depending on the type of product being imported. Again, the raw data contain infrequently traded goods with shares close to 0 and 100, but even when focusing on the top 25 six-digit NAICS industries by importing volume, one observes in Figure 3 significant variation in the share of intrafirm trade, ranging from shares under 8% for US imports of woman's and girls' blouses, shirts, and other outerwear to over 95% for imports of finished autos and light duty vehicles as well as of finished heavy duty trucks.

The very high intrafirm trade share of these finished vehicle industries (which are associated with exports from foreign manufacturing plants to US wholesale affiliates) highlights one of the potential limitations of the data mentioned above, namely the fact that it combines data on both intermediate input sectors and final good sectors. It is important to emphasize, however, that the question of why finished vehicles are rarely imported at arm's-length is one of tracing firm boundaries (though applying in manufacturer–distributor relationships rather than in supplier–manufacturer relationships), and thus it is not obvious that one would want to discard these type of observations when testing the property-rights theory of multinational firms. In any case, the deeper the level of disaggregation in the data, the easier it is to distinguish the nature of the products being imported.



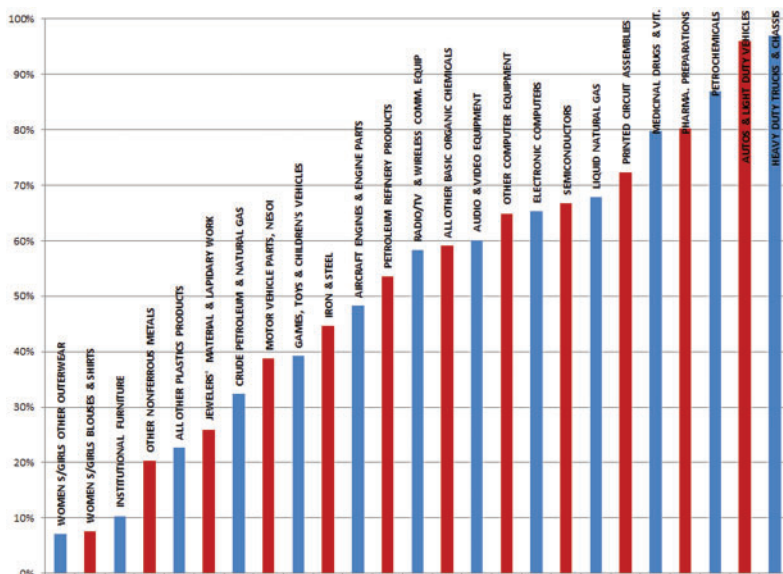


Figure 3. Share of US Intrafirm Imports for Top 25 Importing Industries (NAICS6) in 2010.

I next illustrate this point by moving from the publicly available six-digit NAICS classification (with around 450 distinct industry categories) to the six-digit Harmonized System classification (which contains over 5000 distinct products).<sup>31</sup> Continuing to focus on the case of vehicle imports, in Figure 4, I report the share of intrafirm trade for the 76 six-digit subcategories of the two-digit HS industry 87 (“Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof”). Again, the variation in the share of intrafirm trade within the sector is obvious to the eye and again essentially ranges from 0% to 100%.<sup>32</sup>

In Figure 5, I restrict the sample to a subset of these 76 subcategories, namely those that fall under the four-digit sector 8708 (“Parts and accessories of motor vehicles”) thereby focusing on intermediate input imports. Though the share of intrafirm trade now only ranges from 34.4% for “road wheels and parts” to 78.7% for “gear boxes for motor vehicles”, it is clear that firms based in the United States seem to source different auto parts under quite different ownership structures.

As a final illustration of the richness and variation in the data, in Figure 6 I take one of the six-digit HS industries depicted in Figure 5,

31. I thank Nathan Nunn and Dan Trefler for allowing me to use their data to illustrate the rich variation in the data. The empirical results in the rest of the article use the publicly available six-digit NAICS data.

32. A verbal description of these industry codes is available from the following website: <http://www.foreign-trade.com/reference/hscode.htm>



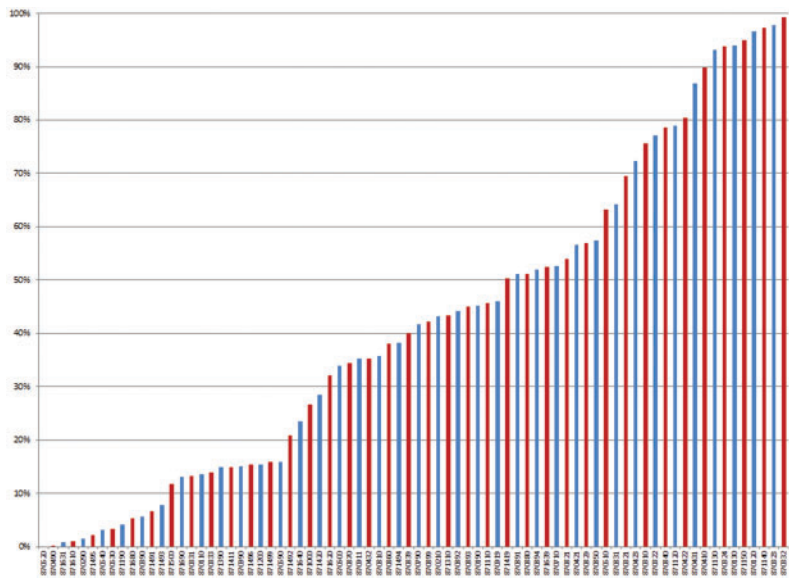


Figure 4. Variation in the Share of Intrafirm Trade within HS 87 (Vehicles, exc. Railways, and Parts) in 2005.

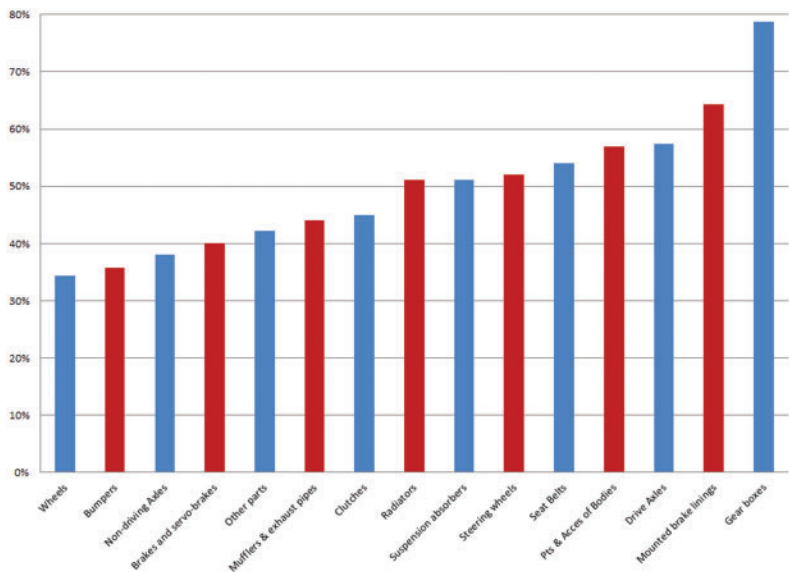


Figure 5. Variation in the Share of Intrafirm Trade within HS Sector 8708 (Auto Parts) in 2005.

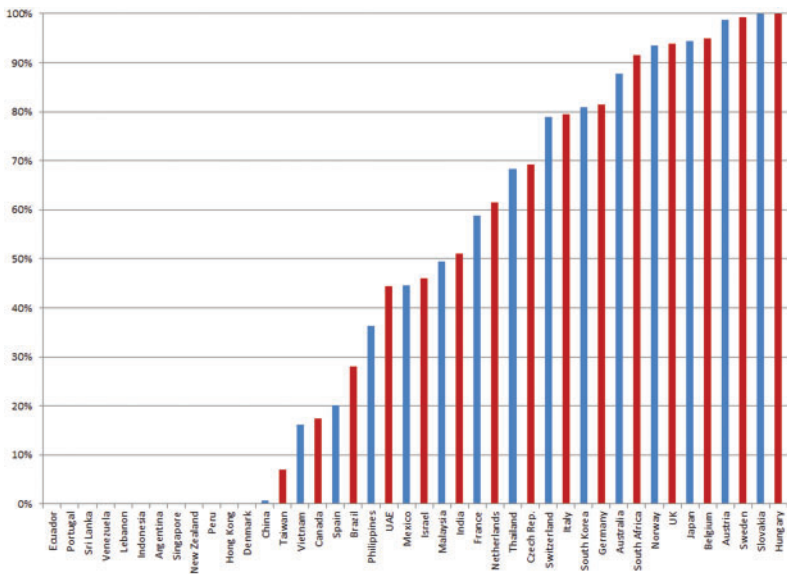


Figure 6. Variation in the Share of US Intrafirm Imports within HS 870810 (Bumpers) in 2005.

namely HS 870810 (“Bumpers and parts thereof for motor vehicles”), and report the share of intrafirm trade for all 42 countries with positive exports to the United States in that sector. As is clear from the graph, even when focusing on a narrowly defined component, a similar pattern to that in Figure 2 emerges, with US-based firms appearing to source particular inputs quite differently depending on the location from which these products are bought. Imports from 12 of the 42 countries are exclusively transacted at arm’s-length, while two countries (Slovakia and Hungary) sell bumpers to the United States exclusively within multinational firm boundaries. The remaining 28 countries feature shares of intrafirm trade fairly uniformly distributed between 0% and 100%.

4.1.3 Intrafirm Trade and the Property-Rights Theory: Empirical Strategies and Findings. Having described some of the basic sources of variation in intrafirm data, I next turn to describing how researchers have attempted to use the data to assess the empirical validity of the property-rights theory of the firm. The key robust implication I have highlighted in Sections 2 and 3 is that the relative prevalence of integration should be higher in relationships that feature high headquarter intensity, i.e., when noncontractible, relationship-specific investments carried out by headquarters are disproportionately more important than those undertaken by suppliers. A key question is then: how do we measure headquarter intensity in the data?

A first attempt at dealing with this issue was offered by Antràs (2003). As mentioned above, his property-rights theory of the multinational firm assumed that the investments provided by headquarters are more physical capital intensive than those provided by suppliers. In his framework, all investments are noncontractible and fully relationship-specific and thus the model generates a positive correlation between the *unobservable* headquarter intensity and physical capital intensity. The assumptions needed to make that connection are strong, so I will work on relaxing them below. Even when making these assumptions, however, one is still faced with some open questions. First, which capital intensity matters for the integration decision? And second, how does one measure it in the data?

As pointed out above, intrafirm trade data do not identify the industry or sector purchasing the imported goods, and thus one cannot easily construct a measure of capital intensity based on the relative investments of the importing sector and those of the exporting sector. Following Antràs (2003), most researchers using intrafirm trade data have associated headquarter intensity with the capital intensity of *the product being imported*. This is because in Antràs' (2003) general-equilibrium model, factors of production are internationally immobile so the headquarter's capital investments are undertaken in the location of the supplier division or firm, thus implying that these investments will be embodied in the intermediate input being shipped back to the headquarter's home country. With that strict interpretation, using the imported good sector's capital intensity is justified, yet it is important to bear in mind that, in the real world, headquarters often undertake other noncontractible relationship-specific investments that may not be embodied in the good being imported. I will return to this point below.

Now even when settling on the use of the capital intensity of the product being imported as a proxy of headquarter intensity, one still needs to measure it in the data. Ideally, one would construct measures of capital intensity (such as the ratio of physical capital expenditures to labor input expenditures) at the same level of disaggregation at which intrafirm trade data are available, namely at the product- and country-of-origin-level. Unfortunately, this is typically infeasible because industry-level data on capital intensity at an acceptable level of disaggregation are only available for a handful of countries. It should be emphasized that this limitation is not specific to the use of intrafirm trade data and also applies to certain empirical tests of the Heckscher–Ohlin model of trade (see, for instance, Romalis 2004). A standard solution to this problem is to work with data from just one country, typically the United States for reasons of availability and data quality, and to impose the same capital intensity to all countries exporting a particular good. This assumption is typically justified by appealing to the absence of capital intensity reversals.

In sum, a typical proxy for the headquarter intensity associated with US imports in a given good or sector is the physical capital intensity (i.e., the ratio of physical capital to employment) in that good or sector in the

United States. It may be argued that Antràs' (2003) assumption that head-quarters and suppliers' investments differ *only* in their capital intensity is too restrictive. With that in mind, the literature has offered alternative measures of headquarter intensity based on other measures of factor intensity, such as skill intensity (the ratio of nonproduction workers to production workers) or R&D intensity (the ratio of R&D expenditures to sales). The idea behind these proxies is that headquarters are particularly likely to be involved in investments related to skilled workers (e.g., training) or to R&D outlays.

Having described the measurement of headquarter intensity, I next turn to discuss some of the key findings in the literature. Rather than simply enumerating the results in previous papers, I will attempt to illustrate some of the key ones with scatter plots.

Antràs (2003) begins his paper by showing that in a cross-section of manufacturing sectors, there exists a striking positive correlation between the share of intrafirm trade in that sector and its physical capital intensity in the United States (see Figure 7). The data in Antràs (2003) include, however, only 23 fairly aggregated industries, and thus one might be skeptical of the robustness of such correlation.<sup>33</sup> This concern is addressed in Figure 8, in which I correlate the share of intrafirm trade with the three measures of headquarter intensity mentioned above (capital intensity, R&D intensity, and skill intensity), while using the much more disaggregated US Customs data on intrafirm trade described above (which have only become available in recent years). The intrafirm trade shares are computed for imports from all exporting countries in a given six-digit NAICS, while the headquarter intensity measures are computed using data from the US Census of Manufactures (in the case of capital and skill intensity) and from Orbis (in the case of R&D intensity).<sup>34</sup> All variables are averaged over the period 2000–05. As is clear from the left panel of Figure 8, the positive correlation between the share of intrafirm trade and capital intensity is robust to the use of this much more detailed data set. The fit is not as good as in the case of Antràs (2003) but the relationship is highly statistically significant. Furthermore, as documented by Nunn and Trefler (2011) and Bernard et al. (2010), this relationship is

33. Antràs' (2003) data are course because he used public extracts from the direct investment data set collected by the BEA. Yeaple (2006) performs similar tests using the confidential BEA data set, which increases the number of sectors from 23 to 51 and also allows him to perform the analysis for different sets of countries. He finds that the effect of capital intensity is robust to the finer disaggregation of the data though it appears much stronger when focusing on less developed and emerging economies than when restricting the sample to developed countries. His empirical analysis as well as Antràs' have now been superseded by a new wave of empirical work using the much more disaggregated US customs intrafirm trade data.

34. I am grateful to Davin Chor, Nathan Nunn, Dan Trefler, and Heiwai Tang for making their data available to me. I add 0.001 to the measure of (log) R&D intensity to avoid throwing away a large number of observations with zero R&D outlays.

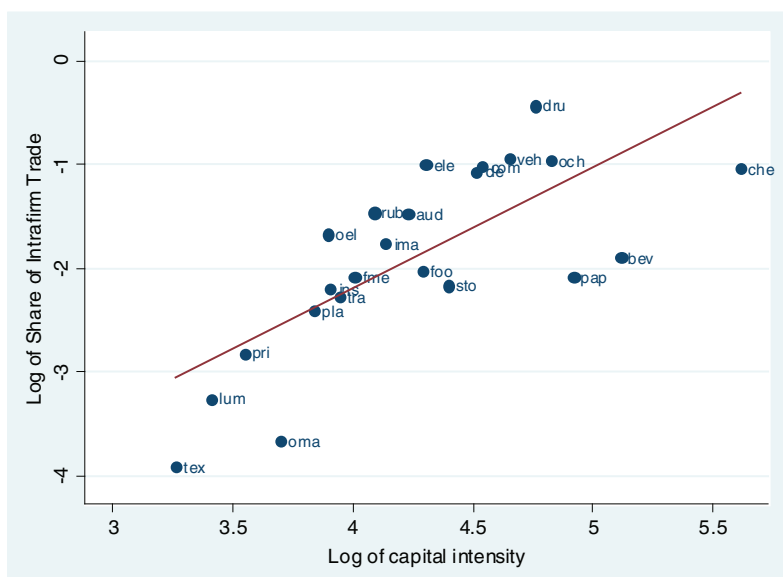


Figure 7. The Share of Intrafirm Trade and Capital Intensity by Sector in Antràs (2003).

robust to the inclusion of various industry controls and to the addition of country-fixed effects in regressions that exploit both the industry as well as the exporting country variation in the data.

The remaining two panels of Figure 8 show the correlations between the share of intrafirm trade and R&D and skill intensity. The positive correlation between these variables is also strongly statistically significant and visible to the naked eye, and as in the case of capital intensity, the effect of these variables continues to hold conditional on a wide set of industry covariates and is also robust to the inclusion of fixed effects in specifications exploiting the cross-country dimension of the data.

There are various reasons why one should be cautious in interpreting these results as empirically validating the property-rights theory of multinational firm boundaries. First of all, the US physical capital, skill, and R&D intensity measures are imperfect proxies for headquarter intensity as they only capture imperfectly the relative importance of the noncontractible, relationship-specific investments carried out by headquarters and their suppliers. Nunn and Trefler (2011) point out, for instance, that standard measures of capital intensity embody several investments that are fairly easy to contract on or that are not particularly relationship specific. If the property-rights theory is correct, one would then expect investments in specialized equipment to be much more relevant for the integration decision than investments in structures or in nonspecialized equipment (such as automobiles or computers), which tend to lose little value when not used in the intended production process. As found by

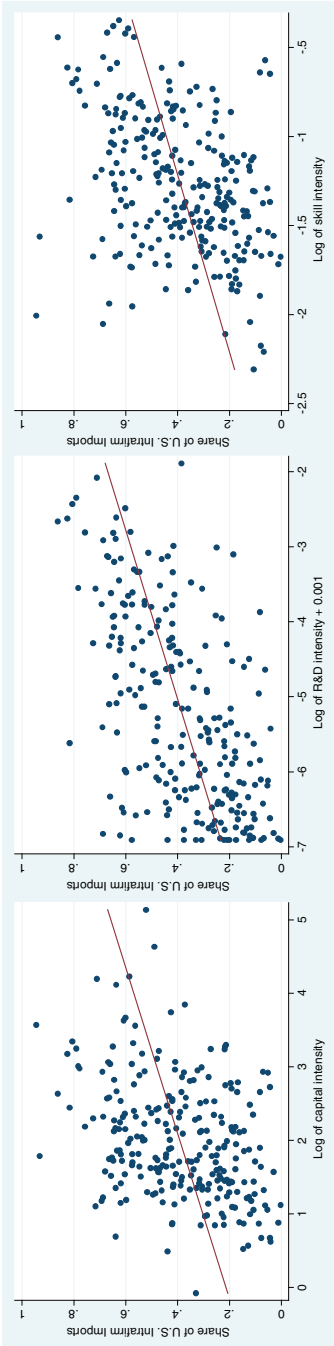


Figure 8. The Effect of Headquarter Intensity on the Share of the US Intrafirm Imports.

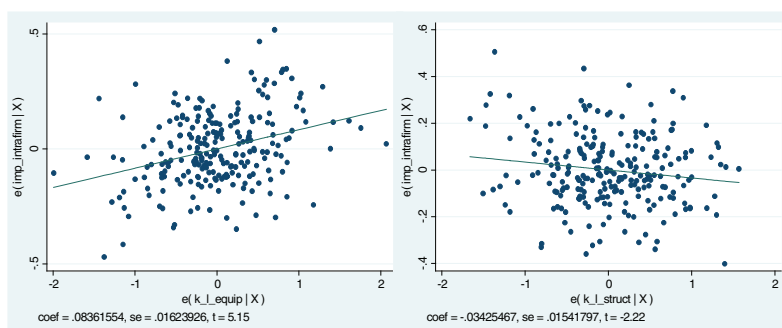


Figure 9. The Effect of Capital Intensity and Its Components on the Share of the US Intrafirm Imports.

Nunn and Trefler (2011), Figures 9 and 10 confirm that this is indeed what one observes in the data when using disaggregated measures of capital intensity from the US Census of Manufactures. In particular, the two panels of Figure 9 depicts the partial effect of capital-equipment intensity (i.e., the ratio of capital equipment expenditures to worker wages) and of capital-structures intensity (i.e., the ratio of expenditures on capital structures to worker wages) in cross-industry regressions that also control for skill intensity and R&D intensity. As is clear from the figure, the positive effect of capital-equipment intensity on the share of intrafirm trade is highly statistically significant, whereas the effect of capital-structures intensity appears to be negative (and also statistically significant). Figure 10 further breaks down the effect of capital equipment intensity into three components and shows that the effect observed in the left panel of Figure 9 is not driven by expenditures on computers and data processing equipment or on automobiles and trucks, which would be problematic for the theory. In fact, the effect of expenditures on automobiles and trucks appears to have a statistically significant *negative* effect on the share of intrafirm trade, which resonates with the results in Antràs and Helpman (2008) and Section 2.3.3 indicating a negative effect of headquarter services contractibility on the integration decision.<sup>35</sup>

35. Another reason that makes standard proxies for headquarter intensity problematic is the fact that they use only information on the capital, skill, or R&D intensity of the “selling” industry, i.e., of the good or sector being imported. Antràs and Chor (2011) acknowledge this limitation and use US input–output data to build measures of headquarter intensity that reflect the capital intensity of the “average buying” industry rather than of the selling one. This has only a minor effect on the estimates, though admittedly this might have to do with the fact that data limitations (i.e., insufficient disaggregation) prevent one from satisfactorily identifying the precise buying industry associated with different US import purchases. For the same reason, alternative tests using measures of capital, skill, or R&D intensity of both the buying and selling industry, as developed for instance in Acemoglu et al. (2010), are typically infeasible using US intrafirm import data.



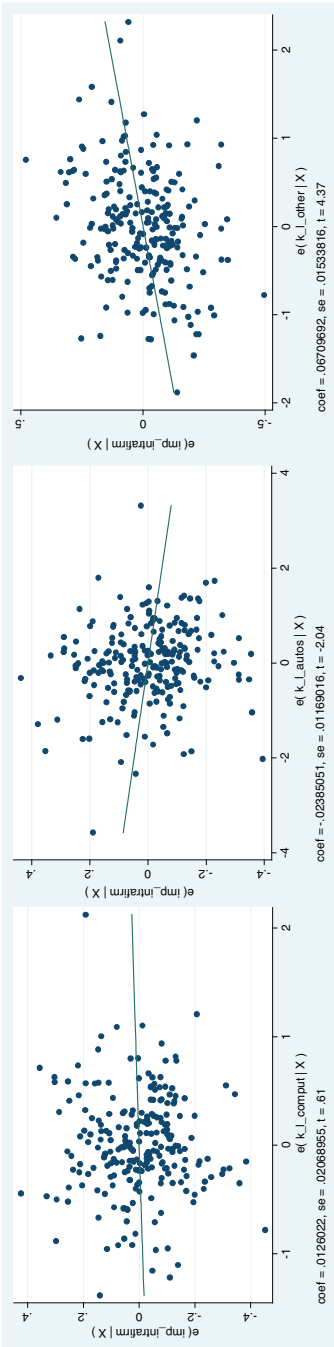


Figure 10. The Effect of Equipment Capital Intensity and Its Components on the Share of U.S. Intrafirm Imports.

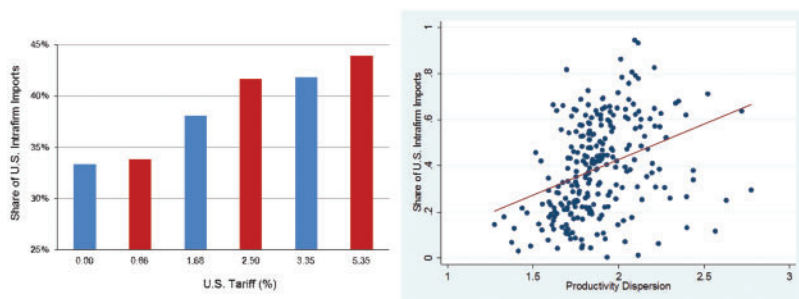


Figure 11. The Implications of Heterogeneity for the Determinants of Intrafirm Imports.

A second reason to be cautious about the findings discussed above relates to the relatively low statistical power of these tests of the property-rights theory. In other words, the patterns one observes in the data are consistent with the property-rights theory but they are not necessarily inconsistent with alternative theories of firm boundaries. For instance, the significance of R&D intensity for the integration decision of multinational firms could be viewed as a validation of transaction-cost theories that emphasize the importance of the nonexcludable nature of knowledge in shaping multinational firm boundaries (see, for instance, [Ethier and Markusen 1996](#)).

In order to alleviate this concern, the literature has attempted to provide more elaborate tests of the theory that relate to the richer set of implications that arise in the property-rights theory of the multinational firm with intraindustry heterogeneity developed by [Antràs and Helpman \(2004, 2008\)](#). Remember, for instance, that these frameworks predict that the share of intrafirm imports should not only increase in headquarter intensity, but should also be positively affected by trade frictions and by the degree of productivity dispersion within industries. [Figure 11](#) confirms the empirical validity of these predictions. In the left panel of the [Figure 11](#), I graphically summarize the results in [Díez \(2010\)](#), who unveils a positive association between the share of intrafirm trade and US tariffs in a cross-section of industries. In particular, the figure sorts industries into bins according to their tariff value (with the first bin containing all industries for which US tariff are 0, and the remaining industries sorted into quartiles), and reports the median share of intrafirm trade in those industries.<sup>36</sup> The right panel of [Figure 11](#) depicts a positive correlation between the share of intrafirm trade and productivity dispersion, as measured by

36. [Díez \(2010\)](#) performs a more complete empirical analysis at the country and industry level that confirms the positive association between the prevalence of intrafirm trade and US tariffs. He also finds a negative correlation between US intrafirm imports and foreign tariffs and shows that it can be reconciled with a variant of the [Antràs and Helpman \(2004\)](#) framework.

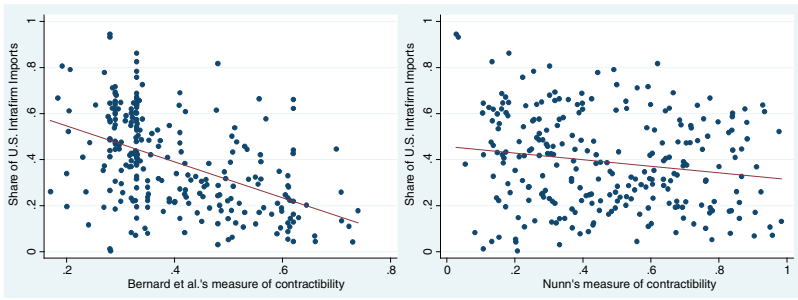


Figure 12. The Effect of Contractibility on the Share of US Intrafirm Imports.

Nunn and Trefler (2008), who also show that the effect of productivity dispersion is higher the higher the quintile of headquarter intensity, a prediction that is again consistent with the property-rights framework in Antràs and Helpman (2004, 2008).<sup>37</sup>

A particularly promising way to discriminate the property-rights theory of the multinational firm against alternative theories of firm boundaries consists of exploiting the implications of the theory for the effect of contractibility on the share of intrafirm trade. As discussed above, the work of Antràs and Helpman (2008) highlights that the effect of contractibility on the prevalence of integration depends crucially on the degree to which contractual incompleteness stems from noncontractibilities in the inputs controlled by the final-good producer or by his or her suppliers. If production processes in certain sectors are particularly noncontractible because of the nature of the investments carried out by headquarters, then the theory would predict that the share of intrafirm trade should be negatively affected by the level of these sectors' contractibility. Conversely, if the source of noncontractibilities stems from the nature of the supplier's activities, the theory may instead predict a positive correlation between the share of intrafirm trade and contractibility, a result that is hard to reconcile with transaction-cost theories of multinational firm boundaries.<sup>38</sup> Though the theory generates sharp predictions for how the source of noncontractibilities affects the share of intrafirm trade, a natural challenge for empiricists is to find appropriate proxies for these different sources of contractual frictions. Figure 12 depicts the correlation between two proposed measures of contractibility and the share of intrafirm trade.

37. The Nunn and Trefler (2008) measure of productivity dispersion corresponds to the standard deviation of log exports reported by the US Department of Commerce for each of the constituent HS10 products from each US location to each destination country in the year 2000.

38. The qualifier "may" in the previous sentence comes from the fact that improvements in contractibility also affect the participation of firms in international trade and may reduce the share of intrafirm trade on that account. See Antràs and Helpman (2008) for details.

The left panel uses the measure of contractibility proposed by Bernard et al. (2010), which is a weighted average of the wholesale employment share of firms importing goods in a particular sector, the idea being that contracting is likely to be easier for products passing through intermediaries such as wholesalers. The right panel uses Nunn's (2007) measure of contractibility (see Nunn and Treffer 2008), which corresponds to the proportion of each sector's intermediate inputs that are not traded on organized exchanges and are thus more susceptible to potential contracting problems. Both panels suggest a negative correlation between the share of intrafirm trade and contractibility, which is broadly in line with what one would expect from transaction-cost models of firm boundaries, but can also be rationalized within the property-rights theory, provided that the lack of contractibility emanates from investments carried out by headquarter services (and possibly embodied in the good transacted across borders). Other findings in the literature, however, appear to be harder to rationalize in transaction-cost models, while they continue to be consistent with the property-rights approach. For instance, both Nunn and Treffer (2008) and Bernard et al. (2010) find that Nunn's measure of contractibility has a positive effect on the share of intrafirm trade in regressions that incorporate the cross-country variation in the data, suggesting that the correlation in the right panel in Figure 12 might simply reflect that the US imports highly contractible goods from countries that export to the United States mostly at arm's-length (perhaps for other reasons). In any case, future work should be directed at better identifying different sources of variation in contractibility so as to better discriminate among alternative models of multinational firm boundaries.

Throughout this section, I have focused on discussing empirical tests based on the cross-sectoral implications of the property-rights theory. The cross-national nature of intrafirm trade data naturally also permits an analysis of some of the cross-country implications of the framework. For instance, the Antràs (2003) framework predicts that the share of intrafirm imports in total imports should be increasing in the aggregate capital-labor ratio of the exporting country,<sup>39</sup> while the framework in Antràs and Helpman (2008) might suggest that the quality of a country's contracting institutions might well have a positive impact on the share of intrafirm trade, a counterintuitive result from the point of view of transaction-cost theories. Consistently with the findings of Antràs (2003), Nunn and Treffer (2008), and Bernard et al. (2010), the two panels in Figure 13 confirm the existence of these correlations in the data. Nevertheless, the standard concerns associated with cross-country regressions (omitted variable biases, endogeneity, etc.) apply here as well, so one should be cautious in interpreting these correlations as formal

39. The Antràs and Helpman (2004) framework would also generate the same prediction if wages in the exporting country are positively affected by that country's aggregate capital-labor ratio.

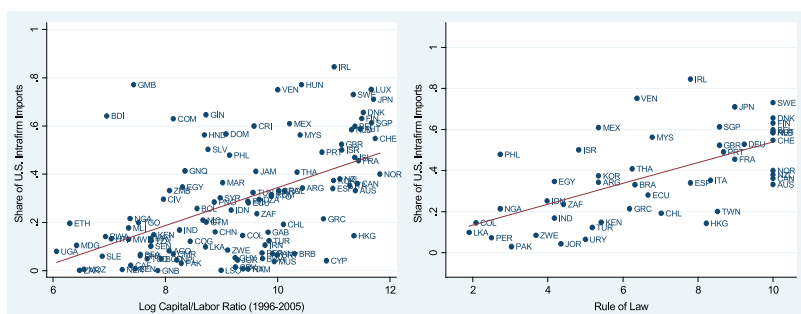


Figure 13. Cross-Country Determinants of the Share of US Intrafirm Imports.

validations of the theory. Nunn and Treffer (2008) and Bernard et al. (2010) also experiment with the addition of controls that interact country and industry variables, but it is not always easy to map these results to the simple models that have been developed so far.<sup>40</sup>

Due to data availability, the bulk of work using product-level data to test the property-rights theory has employed US intrafirm import data. Feenstra and Hanson (2005) and Fernandes and Tang (2010) are two notable exceptions that instead use Chinese data (see also Feenstra 2011, for an overview). In particular, product-level export data from the Customs General Administration of the People's Republic of China contain detailed information on whether the exporter is a foreign-owned plant or not. It is not clear that foreign-owned plants will necessarily export their output to affiliated parties (thus generating intrafirm trade), but regardless the data are suitable for an analysis of the determinants of foreign ownership of suppliers in China. Both sets of authors also make use of the fact that the data distinguish between different types of customs regimes (pure-assembly or import-and-assembly), depending on whether the plant in China is in charge of importing inputs or that responsibility falls to a foreign producer). This motivates the development of rich variants of the property-rights theory, which appear to be able to successfully account for the patterns observed in the data.

#### 4.2 Tests with Firm-Level Data

The property-rights theory is a theory of firm boundaries and thus firm-level data would appear to be the ideal laboratory to use in testing it. An obvious limitation, however, is that this type of data are not readily available. Recent studies have, however, unveiled the existence of a few firm-level data sets that can help shed some light on the empirical

40. For instance, Bernard et al. (2010) find that the share of intrafirm trade is increasing in the interaction of physical capital abundance and physical capital intensity, whereas in Antràs (2003), conditional on capital intensity, the incentive to integrate suppliers is independent of factor prices (and thus of physical capital abundance).

relevance of the property-rights theory. In this section, I will focus on describing four of these firm-level data sets (those that I am aware of), while discussing both their main advantages and limitations, and outlining some of the results that have been obtained when exploiting these data sets.

The first paper to use firm-level data to assess the validity of the property-rights theory of multinational firm boundaries is Tomiura (2007), who uses data from the Basic Survey of Commercial and Manufacturing Structure and Activity in Japan. The survey covers 118,300 Japanese manufacturing firms and according to the author, “is regarded as an accurate overall representation of the whole of manufacturing in Japan,” though unfortunately it was carried out only in one year, 1998. The survey contains various data on the operations of firms (sales, employment, capital expenditures, exports, foreign direct investment) and crucially also asks firms whether they “contract out manufacturing or processing tasks to other firms overseas.” Hence, the survey can be used to explicitly distinguish firms that are engaged in foreign outsourcing versus those that are engaged in foreign direct investment. A key limitation of the data is that they do not appear to contain information on the volumes (i.e., the intensive margin) of foreign insourcing and outsourcing. Tomiura (2007) uses the data set to show that, consistently with the sorting predicted by the Antràs and Helpman (2004) model, firms that are engaged in FDI are distinctively more productive than firms that are engaged in foreign outsourcing, which in turn are more productive than domestic firms. No other predictions from the model are tested. An interesting feature of the data is that most firms are neither “pure FDI” firms nor “pure outsourcing” firms, which suggests that current models are too stylized to capture the rich organizational decisions of firms.

A second line of papers, most notably Defever and Toubal (2007), Corcos et al. (forthcoming) and Carluccio and Fally (forthcoming) have used French data from the EIIG (Échanges Internationaux Intra-Groupe), a survey conducted in 1999 by the SESSI (Service des Études Statistiques Industrielles), which documents the sourcing mode (through independent suppliers or through affiliates) of each firm’s yearly imports of intermediate inputs by origin country and by four-digit HS product codes in 1999. The survey includes all French firms that trade more than one million euros and that are owned by manufacturing groups that control at least 50% of the equity capital of an affiliate based outside France. Though not all firms responded to the survey, the 4305 responding firms represent more than 80% of total exports and imports of French multinationals in 1999.<sup>41</sup> A key limitation of the EIIG data set is the potential for sample selection biases arising from including only firms with at least

41. This data source can in turn be matched with another SESSI database, the EAE (Enquête Annuelle Entreprise), which provides balance sheet data on manufacturing firms with at least 20 employees.

one affiliate outside France. Corcos et al. (forthcoming) acknowledge this problem and complement the data set with data coming from the French Customs Office, documenting the universe of yearly imports and exports flow in 1999 at the firm, origin country and product level, hence allowing them to offer a more representative picture of the foreign outsourcing operations of French firms. The goals and scope of the papers using the EEIG dataset are somewhat different, but they all find supportive evidence of a positive correlation between headquarter intensity and the relative importance of intrafirm trade, with the measures of headquarter intensity in Corcos et al. (forthcoming) being a *firm-level* measure (namely, capital intensity, skill intensity and the ratio of value added over sales of the importing firm). Another conclusion that emerges from these studies is that selection bias identified by Corcos et al. (forthcoming) appears to be important in the sense that the productivity advantage of FDI firms over foreign outsourcers predicted by the theory (see Figure 1) arises only when including firms that only engage in outsourcing into the analysis.

More recently, the property-rights theory of the multinational firm has been tested using firm-level data from Spain provided by the Fundación SEPI. In particular, Kohler and Smolka (2009) make use of the ESEE (Encuesta sobre Estrategias Empresariales), which surveys approximately 2000 Spanish firms with at least 10 employees on a yearly basis since 1990 and provides information on their income and balance sheet statistics, and also on a variety of more specific organizational variables. A notable characteristic of the ESEE is its representativeness, which is ensured by the careful statistical criteria used in the initial year of the sample and the special attention that has been given to account for entry and exit of firms of different sizes in subsequent years. For the purposes of testing the property-rights theory, a particularly relevant feature of the data is that they allow one to compute the overall spending on intermediate inputs by firms and their breakup into (i) domestic purchases from independent suppliers, (ii) domestic purchases from affiliated parties, (iii) imports from foreign independent suppliers, and (iv) imports from foreign affiliates. Hence, one can easily map some of the variables of the survey into the key equilibrium variables in existing models, such as the Antràs and Helpman (2004, 2008) frameworks with intraindustry heterogeneity. An important disadvantage of this Spanish data set is that it only distinguishes between domestic and foreign input purchases, with the latter not being disaggregated by country of origin. Kohler and Smolka (2009) use the ESEE data set and find evidence supportive of the productivity sorting of firms into organizational forms predicted by the Antràs and Helpman (2004, 2008) models (see Figure 1). In particular, conditional on the location of sourcing (domestic or foreign), integrating firms appear to be more productive than nonintegrating ones, and they also provide evidence that firms outsourcing abroad appear to be slightly more productive than firms integrating in Spain. Furthermore, Kohler



and Smolka (2009) find that higher levels of firm-level capital intensity and skill intensity are associated with higher intrafirm shares in both domestic and foreign intermediate input purchases, and also that the effect of productivity on the choice between integration and outsourcing, both at home and abroad, is observed only at high capital intensity levels, again consistently with the predictions in Antràs and Helpman (2004, 2008).

A final international firm-level data set that has been used to shed light on the property-rights theory is Dun & Bradstreet's WorldBase database, which contains public and private plant-level observations in more than 200 countries and territories. The data set does not contain an awful lot of operational data related to these plants, but as pointed out by Alfaro et al. (2010), it does offer a comprehensive picture of firm boundaries across borders. In particular, it contains detailed information on the location, ownership (e.g., its domestic or global parent) and industry classification for hundreds of thousands of plants worldwide. Alfaro et al. (2010) use these data to document a positive association between higher tariffs on final products (as measured by MFN tariffs at the four-digit SIC industry level for all WTO members) and an index of domestic vertical integration constructed with the ownership information in the Dun & Bradstreet's WorldBase database and input–output tables. The empirical exercise exploits both cross-section and time-series variation in trade policy, as well as a more significant trade liberalization episode, namely China's entry into the WTO in 2001. The authors take these results as empirically validating the model of organizational design in Legros and Newman (2009) and Conconi et al. (forthcoming), but they are also consistent with the findings in Antràs and Helpman (2004, 2008).

#### 4.3 Road Ahead

The findings of the studies I have described in this section are interesting and offer broad support for some of the key predictions of the property-rights theory of multinational firm boundaries. Nevertheless, the evidence is far from conclusive because, for the most part, the tests that have been performed up to now have relatively low statistical power. In my view, successful testing of the theory will need to follow one of the two following approaches.

A first possibility is to better exploit the large variation in the relative prevalence of integration retrievable from intrafirm trade data, and most notably, from US intrafirm import data. In particular, the cross-industry studies that I have described above are interesting and informative but they cannot convincingly identify a causal effect of headquarter intensity (even when appropriately measured) on the share of intrafirm trade. A potential avenue for future research is to use narrower slices of the data, perhaps (1) focusing on the patterns in a single industry, but exploiting exogenous changes in sector characteristics driven by technological or demand-driven shocks, in the spirit of Baker and Hubbard (2003), or perhaps (2) performing analyses exploiting within-country variation

stemming from changes in the institutional characteristics of countries, such as observable changes in the quality of institutions or in restrictions on foreign ownership in those countries.

A second, and perhaps even more fruitful area of future research, entails a more structural use of the available firm-level data sets. At present, little work has been devoted to structurally estimating the models I have discussed in this survey. This is partly due to the stylized nature of some of these frameworks, and partly due to the under-utilization of this type of empirical techniques in the international trade field. Yet, as I have argued in Section 3.4, future theoretical developments are likely to provide much richer property-rights frameworks that will be more amenable for structural work. Part of the appeal of using this set of techniques is that this might permit an evaluation of the quantitative importance of multinational firm boundaries for firm-level performance, a question that so far has not been sufficiently explored in the literature.<sup>42</sup>

## 5. Concluding Remarks

In this article, I have surveyed the influence of Grossman and Hart's (1986) seminal paper for the study of the international organization of production. I have discussed the implementation of the theory in open-economy environments and its implications for the structure of international trade flows and multinational activity. I have also reviewed empirical work suggestive of the empirical relevance of the property-rights theory. Along the way, I have developed novel theoretical results and have also outlined some of the key limitations of existing contributions.

I want to conclude by briefly discussing the broader impact of GH in the field of international trade. First, as emphasized earlier in this article, GH has been an inspiration for the development of various complementary theories of the organization of production, some of which have also been put to work both theoretically and empirically in open-economy environments (see, for instance, the references described at the end of Section 3.3). These papers may not be the direct *offsprings* of GH, but they clearly carry some of GH's traits.

Perhaps more significantly, soon after the development of property-rights theories of multinational firm boundaries, the literature acknowledged that incomplete contracting of the type introduced by GH could shape not only the ownership structure of firms but also their geographical location. In other words, contracting institutions are not only important for understanding vertical integration decisions, but they also constitute a source of comparative advantage. The literature on trade and institutions has exploded in the last few years, with work studying both theoretically as well as empirically the effects of contracting, financial and labor market

42. The work of Guadalupe et al. (forthcoming) is a first promising step in that direction.

institutions on trade patterns, multinational activity, and more broadly the impact of these institutions on the workings of general equilibrium models.<sup>43</sup> The main ideas behind these papers can all be traced back in some form or other to the seminal work of GH.

Finally, by formalizing the idea of *power* in market and nonmarket economic transactions, the ideas in GH have been shown to have new and interesting implications for how the process of globalization affects different agents in society (depending on their endowments of assets or information), and also for how the effects of trade policies are transmitted across countries, with implications for the efficacy of the rules that currently govern negotiations at the World Trade Organization (see, among others, [Antràs and Costinot 2011](#); Antràs and Staiger forthcoming).

## Appendix A

**Derivation of Equation (8)** Due to  $R(h, m)$  being homogeneous of degree  $\alpha$ , we can write

$$hR_h + mR_m = \alpha R$$

$$hR_{hh} + mR_{hm} = (\alpha - 1)R_h$$

$$hR_{hm} + mR_{mm} = (\alpha - 1)R_m,$$

while the formula for the elasticity of substitution between  $h$  and  $m$  is given by

$$\sigma_{h,m} = \frac{-R_h R_m (hR_h + mR_m)}{hm(R_{hh}(R_m)^2 + R_{mm}(R_h)^2 - 2R_{hm}R_h R_m)}.$$

Denoting by

$$\eta_{R,h} = \frac{hR_h}{R}; \quad \eta_{R,m} = \frac{mR_m}{R},$$

we can use the formulas above (together with  $\eta_{R,h} + \eta_{R,m} = \alpha$ ) to obtain:

$$\frac{R_{hh}}{R} = - \frac{(\sigma_{h,m}(1 - \alpha)\eta_{R,h} + \eta_{R,m})\eta_{R,h}}{\alpha h^2 \sigma_{h,m}}$$

$$\frac{R_{mm}}{R} = - \frac{(\sigma_{h,m}(1 - \alpha)\eta_{R,m} + \eta_{R,h})\eta_{R,m}}{\alpha m^2 \sigma_{h,m}}$$

$$\frac{R_{hm}}{R} = \frac{(1 - \sigma_{h,m}(1 - \alpha))\eta_{R,h}\eta_{R,m}}{\alpha hm \sigma_{h,m}}.$$

43. See, among many others, [Acemoglu et al. \(2007\)](#); [Levchenko \(2007\)](#); [Nunn \(2007\)](#) on contracting institutions, [Antràs and Caballero \(2007\)](#); [Antràs et al. \(2009\)](#); and [Manova \(2010\)](#) on financial institutions, and [Helpman et al. \(2010\)](#) on labor–market institutions.

Now plugging these expressions into (3) using (4) and (5) delivers

$$\frac{\beta^*}{1 - \beta^*} = \frac{\eta_{R,h} \sigma_{h,m}(1 - \alpha)\eta_{R,m} + \eta_{R,h} - \alpha\beta^*}{\eta_{R,m} \alpha\beta^* - \eta_{R,h} + \sigma_{h,m}\eta_{R,h}(1 - \alpha)},$$

which in turn simplifies to (8), or

$$\frac{\beta^*}{1 - \beta^*} = \sqrt{\frac{\eta_{R,h} \alpha(1 - \eta_{R,m}) + (1 - \alpha)(\sigma_{h,m} - 1)\eta_{R,m}}{\eta_{R,m} \alpha(1 - \eta_{R,h}) + (1 - \alpha)(\sigma_{h,m} - 1)\eta_{R,h}}}.$$

Straightforward differentiation confirms that  $\beta^*$  is increasing in  $\eta_{R,h}$  and decreasing in  $\eta_{R,m}$  for any  $\sigma_{h,m} > 0$  (remember that  $\eta_{R,h} > 0, \eta_{R,m} > 0$  and  $\eta_{R,h} + \eta_{R,m} = \alpha$ ).

*Proof of Proposition 3.* Following the algebra in Antràs and Helpman (2008), it can easily be verified that when  $F$  obtains a share of revenue  $\beta_k$  in the ex post bargaining, the equilibrium level of revenues is given by

$$R = A^{1/(1-\eta_h-\eta_m)} \left( \frac{\beta_k \eta_h}{c_h} \right)^{\eta_h/(1-\eta_m-\eta_h)} \left( \frac{(1-\beta_k)\eta_m}{c_m} \right)^{(1-\eta_h)/(1-\eta_m-\eta_h)}$$

while input choices satisfy

$$c_h h = \beta_k \eta_h R$$

$$c_m m = (1 - \beta_k) \eta_m R.$$

Using equations (1) and (9), the relative profits obtained by  $F$  under integration versus outsourcing are then given by

$$\frac{\pi_V}{\pi_O} = \frac{(1 + \delta)(1 - \eta_h) + \phi(1 - \delta)(1 - \eta_m)}{(1 - \eta_h) + \phi(1 - \eta_m)} ((1 + \delta)^{\eta_h} (1 - \delta)^{\eta_m})^{1/(1-\eta_h-\eta_m)}.$$

To characterize the effects of the different parameters on this ratio of profits, it is simplest to work with the following monotonic transformation of  $\pi_V/\pi_O$ :

$$\begin{aligned} \Theta(\eta_h, \eta_m, \phi, \delta) &= (1 - \eta_h - \eta_m) \ln \left( \frac{\pi_V}{\pi_O} \right) \\ &= (1 - \eta_h - \eta_m) \ln \left( \frac{(1 + \delta)(1 - \eta_h) + \phi(1 - \delta)(1 - \eta_m)}{(1 - \eta_h) + \phi(1 - \eta_m)} \right) \\ &\quad + \eta_h \ln(1 + \delta) + \eta_m \ln((1 - \delta)). \end{aligned}$$

Simple differentiation shows that  $\Theta(\eta_h, \eta_m, \phi)$  is a decreasing function of  $\phi$ , and thus the relative profitability of integration is higher, the tighter are financial constraints (i.e., the lower is  $\phi$ ). We next show that  $\Theta(\eta_h, \eta_m, \phi)$  is also increasing in  $\eta_h$  and decreasing in  $\eta_m$ .

We begin by noting that

$$\frac{\partial^2 \Theta(\eta_h, \eta_m, \phi, \delta)}{\partial(\eta_h)^2} = 4(1-\eta_m)\delta\phi \frac{\eta_m(1+\delta)(1-\eta_h) + \phi(1-\eta_m)(\eta_m(1-\phi) + \phi(1-\delta) + \eta_m\delta\phi + 1 - \eta_h)}{((1+\delta)(1-\eta_h) + \phi(1-\delta)(1-\eta_m))^2((1-\eta_h) + \phi(1-\eta_m))^2} > 0$$

and

$$\frac{\partial^2 \Theta(\eta_h, \eta_m, \phi, \delta)}{\partial(\eta_m)^2} = -4(1-\eta_h)\delta\phi \frac{\phi^2\eta_h(1-\eta_m)(1-\delta) + \phi(1-\eta_h)(\eta_h + 1 - \eta_m) + (1-\eta_h)^2(1+\delta)}{((1+\delta)(1-\eta_h) + \phi(1-\delta)(1-\eta_m))^2((1-\eta_h) + \phi(1-\eta_m))^2} < 0,$$

so it suffices to show that  $\partial\Theta(\eta_h, \eta_m, \phi)/\partial\eta_h \geq 0$  when  $\eta_h=0$  and  $\partial\Theta(\eta_h, \eta_m, \phi)/\partial\eta_m \leq 0$  when  $\eta_m=0$ .

One can verify, however, that

$$\left. \frac{\partial\Theta(\eta_h, \eta_m, \phi, \delta)}{\partial\eta_h} \right|_{\eta_h=0} = \ln(1+\delta) - \ln\left(\frac{1+\delta+\phi(1-\delta)(1-\eta_m)}{1+\phi(1-\eta_m)}\right) - \frac{2\delta\phi(1-\eta_m)^2}{(1+\phi(1-\eta_m))(1+\delta+\phi(1-\delta)(1-\eta_m))} \geq 0,$$

where the sign follows from the right-hand side being increasing in  $\delta$  and equalling 0 when  $\delta=0$ . Similarly, we have

$$\left. \frac{\partial\Theta(\eta_h, \eta_m, \phi, \delta)}{\partial\eta_m} \right|_{\eta_m=0} = \ln(1-\delta) - \ln\left(\frac{\phi(1-\delta) + (1+\delta)(1-\eta_h)}{1-\eta_h+\phi}\right) + \frac{2\delta\phi(1-\eta_h)^2}{(1-\eta_h+\phi)(1-\eta_h+\phi+\delta-\delta\phi-\delta\eta_h)} \leq 0,$$

where the sign follows now from the right-hand side being decreasing in  $\delta$  (which can be verified via differentiation) and equalling 0 when  $\delta=0$ .

Next note that for sufficiently high  $\eta_h$ , we must have  $\Theta(\eta_h, \eta_m, \phi, \delta) > 0$ . This can be verified analytically but it suffices to point out that this is true when  $\phi=1$  (see Proposition 2) and the function  $\Theta(\cdot)$  is decreasing in  $\phi$ . Conversely, for sufficiently low  $\eta_h$ , i.e.,  $\eta_h \rightarrow 0$ , we have

$$\Theta(\eta_h, \eta_m, \phi, \delta) \rightarrow (1-\eta_m)\ln\left(\frac{(1+\delta) + \phi(1-\delta)(1-\eta_m)}{1+\phi(1-\eta_m)}\right) + \eta_m\ln((1-\delta)),$$

which is negative for sufficiently high  $\eta_m$ , but may be positive for sufficiently low  $\eta_m$  (this is easily verified by setting  $\phi=0$ , for instance). In sum, given that  $\pi_{V/\pi_O}$  is increasing in  $\eta_h$ , we can conclude that there exists a headquarter-intensity cutoff  $\hat{\eta}_h \in [0, 1)$  such that for all  $\eta_h > \hat{\eta}_h$ , integration necessarily dominates outsourcing. When  $\hat{\eta}_h > 0$ , outsourcing dominates integration for all  $\eta_h < \hat{\eta}_h$  but note that it is possible that  $\hat{\eta}_h = 0$  and thus integration dominates outsourcing for all  $\eta_h \in (0, 1)$ . Finally,

given the comparative statics discussed above, we can use the implicit function theorem to conclude that the cutoff  $\hat{\eta}_h$  is lower the larger is  $\phi$  and the smaller is  $\eta_m$ .

*Proof of Proposition 4.* We first derive the formulas for the ex post payoffs of the firm and the suppliers. To compute the Shapley value for supplier  $j$ , first note that the firm is an *essential* player in the bargaining game and thus a supplier's marginal contribution is equal to zero when a coalition does not include the firm. When it does include the firm and a measure  $n$  of suppliers, the marginal contribution of supplier  $j$  is equal to  $m(j, n) = \partial R(h, \mathbf{m}, \mathbf{n}) / \partial n$ , where  $R(h, \mathbf{m}, \mathbf{n}) = Ah^{\eta_h} \left( \int_0^n m(i)^\rho di \right)^{\frac{\eta_m}{\rho}}$ . This produces

$$m(j, n) = \frac{\eta_m}{\rho} Ah^{\eta_h} m(-i)^{\eta_m} \left( \frac{m(i)}{m(-i)} \right)^\rho n^{\frac{\eta_m - \rho}{\rho}},$$

where  $m(-i)$  represents the (symmetric) investments of all suppliers other than  $i$  and where in the second equality we have used the fact that supplier investments will be symmetric in equilibrium.

The Shapley value of supplier  $j$  is the average of her marginal contributions to coalitions that consist of players ordered below her in all feasible orderings. A supplier that has a measure  $n$  of players ordered below her has a marginal contribution of  $m(j, n)$  if the firm is ordered below her (probability  $n$ ) and 0 otherwise (probability  $1 - n$ ). Averaging over all possible orderings of the players and using the above formula for  $m(j, n)$  we obtain:

$$s_O(h, m(-j), m(j)) = \int_0^1 nm(j, n)dn = \frac{\eta_m}{\eta_m + \rho} Ah^{\eta_h} m(-i)^{\eta_m} \left( \frac{m(i)}{m(-i)} \right)^\rho,$$

which corresponds to equation (12). In the case of vertical integration, the marginal contribution of a supplier is reduced by a factor  $1 - \delta$  and so is her final payoff, thus resulting in equation (13).

Denote by  $\gamma_k$  the share of revenue accruing to the firm under the (symmetric) ownership structure  $k \in \{V, O\}$ , so  $\gamma_O = \rho / (\eta_m + \rho)$  and  $\gamma_V = (\eta_m \delta + \rho) / (\eta_m + \rho)$ . From equations (12) and (13) and imposing symmetry, we have that equilibrium input choices satisfy

$$\begin{aligned} \gamma_k \eta_h A(h_k)^{\eta_h - 1} (m_k)^{\eta_m - 1} &= c_h \\ (1 - \gamma_k) \rho A(h_k)^{\eta_h} (m_k)^{\eta_m - 1} &= c_m. \end{aligned}$$

Combining these equations, we have that joint profits are given by

$$\begin{aligned} \pi_k &= (1 - \gamma_k \eta_h - (1 - \gamma_k) \rho) A^{1/(1 - \eta_m - \eta_h)} \left( \frac{\gamma_k \eta_h}{c_h} \right)^{\eta_h / (1 - \eta_m - \eta_h)} \\ &\quad \times \left( \frac{(1 - \gamma_k) \rho}{c_m} \right)^{\eta_m / (1 - \eta_m - \eta_h)}. \end{aligned}$$

Computing the ratio of profits under integration and outsourcing, taking logarithms and multiplying by  $1 - \eta_m - \eta_h > 0$ , we can write

$$\begin{aligned}\Theta(\eta_h, \rho) &= (1 - \eta_h - \eta_m) \ln\left(\frac{\pi_V}{\pi_O}\right) \\ &= (1 - \eta_h - \eta_m) \ln\left(\frac{(\eta_m + \rho) - (\eta_m \delta + \rho)\eta_h - \eta_m(1 - \delta)\rho}{\eta_m(1 - \rho) + \rho(1 - \eta_h)}\right) \\ &\quad + \eta_h \ln\left(\frac{(\eta_m \delta + \rho)}{\rho}\right) + \eta_m \ln(1 - \delta)\end{aligned}$$

Straightforward differentiation delivers:

$$\frac{\partial^2(\Theta(\eta_h, \rho, \delta))}{\partial \eta_h \partial \delta} = \frac{(\eta_m \delta + \delta \rho + 1 - \delta)(\eta_m + \rho)(1 - \rho)(\eta_m)^2}{((\eta_m + \rho) - (\eta_m \delta + \rho)\eta_h - \eta_m(1 - \delta)\rho)^2(\rho + \eta_m \delta)} > 0,$$

while it is also easily verified that

$$\left. \frac{\partial \Theta(\eta_h, \rho)}{\partial \eta_h} \right|_{\delta=0} = 0.$$

Which implies that  $\frac{\partial \Theta(\eta_h, \eta_m, \rho)}{\partial \eta_h} \geq 0$  and the relative profitability of integration is increasing in  $\eta_h$ .

We next focus on the effect of  $\rho$  on the integration decision. Note first that

$$\begin{aligned}\lim_{\rho \rightarrow 0} \Theta(\eta_h, \rho) &= (1 - \eta_h - \eta_m) \ln(1 - \delta \eta_h) + \eta_h \lim_{\rho \rightarrow 0} \left( \ln\left(1 + \frac{\eta_m \delta}{\rho}\right) \right) \\ &\quad + \eta_m \ln(1 - \delta) = +\infty,\end{aligned}$$

and

$$\Theta(\eta_h, 1) = (1 - \eta_m) \ln(1 + \eta_m \delta) + \eta_m \ln(1 - \delta) < 0,$$

where the last expression is negative because it is decreasing in  $\delta$  and equals 0 at  $\delta = 0$ . Hence, integration dominates outsourcing for sufficiently low  $\rho$ , and the converse is true for a high enough  $\rho$ . To demonstrate the existence of a unique threshold as stated in Proposition 4, note that

$$\begin{aligned}\frac{\partial(\Theta(\eta_h, \rho))}{\partial \rho} &= \Delta \left[ \rho^2 + \rho \frac{(\delta \eta_m(1 - \eta_h) - \eta_h(1 - \delta))(1 - \eta_h - \eta_m) - \eta_h(1 - \eta_m(1 - \delta) - \eta_h)}{(1 - \eta_m - \eta_h)(1 - \delta \eta_h)} \right. \\ &\quad \left. - \frac{\eta_m \eta_h}{(1 - \eta_m - \eta_h)} \right],\end{aligned}$$

with

$$\Delta = \frac{(\eta_m)^2 \delta (1 - \eta_m - \eta_h)(1 - \delta \eta_h)}{\rho(\eta_m \delta + \rho)(\eta_m + \rho - \eta_h(\rho + \eta_m \delta) - \eta_m \rho(1 - \delta))(\eta_m(1 - \rho) + \rho(1 - \eta_h))} > 0.$$



Note that the term in brackets in (18) constitutes a quadratic equation in  $\rho$  of the form  $\rho^2 + b\rho + c$ , with  $c < 0$ . This implies, however, that it can only take a value of 0 at most once for  $\rho > 0$ . Together with the limiting values  $\lim_{\rho \rightarrow 0} \Theta(\eta_h, \rho) = +\infty$  and  $\Theta(\eta_h, 1) < 0$ , we can thus conclude that  $\Theta(\eta_h, \hat{\rho}) = 0$  for a unique value  $\hat{\rho} \in (0, 1)$ , as stated in the Proposition.

## References

- Acemoglu, Daron, Philippe Aghion, Rachel Griffith, and Fabrizio Zilibotti. 2010. "Vertical Integration and Technology: Theory and Evidence," 8 *Journal of the European Economic Association* 1–45.
- Acemoglu, Daron, Pol Antràs, and Elhanan Helpman. 2007. "Contracts and Technology Adoption," 97 *American Economic Review* 916–43.
- Aghion, Philippe, and Jean Tirole. 1994. "The Management of Innovation," 109 *Quarterly Journal of Economics* 1185–209.
- . 1997. "Formal and Real Authority in Organizations," 105 *Journal of Political Economy* 1–29.
- Alfaro, Laura, Paola Conconi, Harald Fadinger, and Andrew F. Newman. 2010. "Trade Policy and Firm Boundaries," NBER Working Paper No. 16118, National Bureau of Economic Research.
- Antràs, Pol. 2003. "Firms, Contracts, and Trade Structure," 118 *Quarterly Journal of Economics* 1375–418.
- . 2005. "Incomplete Contracts and the Product Cycle," 95 *American Economic Review* 1054–73.
- Antràs, Pol, and Ricardo J. Caballero. 2007. "Trade and Capital Flows: A Financial Frictions Perspective," 117 *Journal of Political Economy* 701–44.
- Antràs, Pol, and Davin Chor. 2011. "Organizing the Global Value Chain," Working paper, Harvard University.
- Antràs, Pol, and Arnaud Costinot. 2011. "Intermediated Trade," 126 *Quarterly Journal of Economics* 1319–74.
- Antràs, Pol, Mihir Desai, and C. Fritz Foley. 2009. "Multinational Firms, FDI Flows and Imperfect Capital Markets," 124 *Quarterly Journal of Economics* 1171–219.
- Antràs, Pol, and Elhanan Helpman. 2004. "Global Sourcing," 112 *Journal of Political Economy* 552–80.
- . 2008. "Contractual Frictions and Global Sourcing," in E. Helpman, D. Marin, and T. Verdier, eds., *The Organization of Firms in a Global Economy*. 9–54. Cambridge, MA: Harvard University Press.
- Antràs, Pol, and Esteban Rossi-Hansberg. 2009. "Organizations and Trade," 1 *Annual Review of Economics* 43–64.
- Antràs, Pol, and Robert W. Staiger. Forthcoming. "Offshoring and the Role of Trade Agreements," *American Economic Review*.
- Aumann, Robert J., and Lloyd S. Shapley. 1974. *Values of Non-Atomic Games*. Princeton, NJ: Princeton University Press.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm," 117 *Quarterly Journal of Economics* 39–84.
- Baker, George P., and Thomas N. Hubbard. 2003. "Make Versus Buy In Trucking: Asset Ownership, Job Design, And Information," 93 *American Economic Review* 551–72.
- Basco, Sergi. 2010. "Financial Development and the Product Cycle," Working paper, Universidad Carlos III.
- Bernard, Andrew B., J. Bradford Jensen, Stephen Redding, and Peter K. Schott. 2010. "Intra-Firm Trade and Product Contractibility," 100 *American Economic Review* 444–48.
- Campa, Jose M., and Goldberg Linda S. 1997. "The Evolving External Orientation of Manufacturing: A Profile of Four Countries," 3 *Federal Reserve Bank of New York Economic Policy Review* 53–81.

- Carluccio, Juan, and Thibault Fally. Forthcoming. "Global Sourcing under Imperfect Capital Markets," *The Review of Economics and Statistics*.
- Casson, Mark. 1979. *Alternatives to the Multinational Enterprise*. London: Palgrave Macmillan.
- Caves, Richard E. 1996. *Multinational Enterprise and Economic Analysis*. 2nd ed. Cambridge, UK: Cambridge University Press.
- Chen, Yongmin, and Robert C. Feenstra. 2008. "Buyer Investment, Export Variety and Intrafirm Trade," 52 *European Economic Review* 1313–37.
- Chen, Yongmin, Ig Horstmann, and James Markusen. 2008. "Physical Capital, Knowledge Capital and the Choice between FDI and Outsourcing," 45 *Canadian Journal of Economics* 1–15.
- Coase, Ronald H. 1937. "The Nature of the Firm," 4 *Economica* 386–405.
- Conconi, Paola, Patrick Legros, and Andrew F. Newman. Forthcoming. "Trade Liberalization and Organizational Change," 86 *Journal of International Economics* 197–208.
- Corcos, Gregory. 2006. "Globalisation, Vertical Linkages, and Relational Contracts," Working paper, Paris School of Economics.
- Corcos, Gregory, Delphine M. Irac, Giordano Mion, and Thierry Verdier. Forthcoming. "The Determinants of Intra-Firm Trade," *The Review of Economics and Statistics*.
- Defever, Fabrice, and Farid Toubal. 2007. "Productivity and the Sourcing Modes of Multinational Firms: Evidence from French Firm-Level Data," CEP Discussion Paper No. 842. Centre for Economic Performance. London School of Economics and Political Science.
- Diez, Federico J. 2010. "The Asymmetric Effects of Tariffs on Intra-Firm Trade and Offshoring Decisions," Working Paper no. 10-4. Federal Reserve Bank of Boston.
- Dixit, Avinash K., and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity," 67 *American Economic Review* 297–308.
- Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2003. "Courts," 118 *Quarterly Journal of Economics* 453–517.
- Dunning, John H. 1981. *International Production and the Multinational Enterprise*. London: Allen and Unwin.
- Ethier, Wilfred J. 1986. "The Multinational Firm," 101 *Quarterly Journal of Economics* 805–33.
- Ethier, Wilfred J., and James R. Markusen. 1996. "Multinational Firms, Technology Diffusion and Trade," 41 *Journal of International Economics* 1–28.
- Feenstra, Robert C. 2011. "Offshoring to China: The Local and Global Impacts of Processing Trade," *Lectures presented at the University of International Business and Economics*. Beijing; April 18–20, 2011.
- Feenstra, Robert C., and Gordon H. Hanson. 1996. "Globalization, Outsourcing, and Wage Inequality," 86 *American Economic Review* 240–45.
- . 2005. "Ownership and Control in Outsourcing to China: Estimating the Property-Rights Theory of the Firm," 120 *Quarterly Journal of Economics* 729–61.
- Fernandes, Ana P., and Heiwai Tang. 2010. "Determinants of Vertical Integration in Export Processing: Theory and Evidence from China," Working paper, Tufts University.
- Grossman, Sanford J., and Oliver D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," 94 *Journal of Political Economy* 691–719.
- Grossman, Gene M., and E. Helpman. 2002. "Integration vs. Outsourcing in Industry Equilibrium," 117 *Quarterly Journal of Economics* 85–120.
- Grossman, Gene M., and Elhanan Helpman. 2003. "Outsourcing versus FDI in Industry Equilibrium," 1 *Journal of the European Economic Association* 317–27.
- . 2004. "Managerial Incentives and the International Organization of Production," 63 *Journal of International Economics* 237–62.
- Guadalupe, Maria, Olga Kuzmina, and Catherine Thomas. Forthcoming. "Innovation and Foreign Ownership," *American Economic Review*.

- Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure*. New York, NY: Oxford University Press.
- Hart, Oliver, and John Moore. 1990. "Property Rights and the Nature of the Firm," 98 *Journal of Political Economy* 1119–58.
- Hart, Oliver, and Bengt R. Holmström. 2010. "A Theory of Firm Scope," 125 *Quarterly Journal of Economics* 483–513.
- Helpman, Elhanan. 2006. "Trade, FDI and the Organization of Firms," 44 *Journal of Economic Literature* 589–630.
- Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding. 2010. "Inequality and Unemployment in a Global Economy," 78 *Econometrica* 1239–83.
- Helpman, Elhanan, and Paul R. Krugman. 1985. Chapter 12: Single-Product Firms, *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy* Cambridge, MA: MIT Press.
- Hummels, David, Jun Ishii, and Kei-Mu Yi. 2001. "The Nature and Growth of Vertical Specialization in World Trade," 54 *Journal of International Economics* 75–96.
- Hymer, Stephen H. 1960. *The International Operations of National Firms: A Study of Direct Foreign Investment*, Ph.D dissertation, Massachusetts Institute of Technology, Department of Economics. Published posthumously 1976, Cambridge, MA: MIT Press.
- Kohler, Wilhelm, and Marcel Smolka. 2009. "Global Sourcing Decisions and Firm Productivity: Evidence from Spain," CESifo Working Paper No. 2903, Munich: CESifo Group.
- Legros, Patrick, and Andrew F. Newman. 2008. "Competing for Ownership," 6 *Journal of the European Economic Association* 1279–308.
- . 2009. "A Price Theory of Vertical and Lateral Integration," CEPR Discussion Paper No. 7211, London: Centre for Economic Policy Research.
- Levchenko, Andrei. 2007. "Institutional Quality and International Trade," 74 *Review of Economic Studies* 791–819.
- McLaren, John. 2000. "Globalization and Vertical Structure," 90 *American Economic Review* 1239–54.
- Manova, Kalina. 2010. "Credit Constraints, Heterogeneous Firms and International Trade," Working paper, Stanford University.
- Marin, Dalia, and Thierry Verdier. 2009. "Power in the Multinational Corporation in Industry Equilibrium," 38 *Economic Theory* 437–64.
- Markusen, James R. 1995. "The Boundaries of Multinational Enterprises and the Theory of International Trade," 9 *Journal of Economic Perspectives* 169–89.
- Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," 71 *Econometrica* 1695–725.
- Midler, Paul. 2009. *Poorly Made in China*. Hoboken, NJ: John Wiley & Sons, Inc.
- Nunn, Nathan. 2007. "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade," 122 *Quarterly Journal of Economics* 569–600.
- Nunn, Nathan, and Daniel Treffer. 2008. "The Boundaries of the Multinational Firm: An Empirical Analysis," in E. Helpman, D. Marin, and T. Verdier, eds., *The Organization of Firms in a Global Economy*. Cambridge, MA: Harvard University Press.
- . 2011. "Incomplete Contracts and the Boundaries of the Multinational Firm," Working paper, Harvard University.
- Ponzetto, Giacomo. 2009. "Intellectual Property Rights and Efficient Firm Organization," Working paper, CREI-Universitat Pompeu Fabra.
- Puga, Diego, and Daniel Treffer. 2010. "Wake Up and Smell the Ginseng: The Rise of Incremental Innovation in Low-Wage Countries," 91 *Journal of Development Economics* 64–76.
- Qiu, Larry D., and Barbara Spencer. 2002. "Keiretsu and Relationship-Specific Investment: Implications for Market-Opening Trade Policy," 58 *Journal of International Economics* 49–79.

- Rajan, Raghuram G., and Luigi Zingales. 2001. "The Firm as a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms," 116 *Quarterly Journal of Economics* 805–52.
- Rodrik, Dani. 2000. "How Far Will International Economic Integration Go?," 14 *Journal of Economic Perspectives* 177–86.
- Romalis, John. 2004. "Factor Proportions and the Structure of Commodity Trade," 94 *American Economic Review* 67–97.
- Rugman, Alan M. 1981. *Inside the Multinationals: The Economics of Internal Markets*. New York, NY: Columbia University Press.
- Schwarz, Christian, and Jens Suedekum. 2011. "Global Sourcing of Complex Production Processes," CESifo Working Paper No. 3559, Munich: CESifo Group.
- Spencer, Barbara. 2005. "International Outsourcing and Incomplete Contracts," 38 *Canadian Journal of Economics* 1107–35.
- Tomiura, Eiichi. 2007. "Foreign Outsourcing, Exporting, and FDI: A Productivity Comparison at the Firm Level," 72 *Journal of International Economics* 113–27.
- van Biesebroeck, Johannes, and Lijun Zhang. 2011. "Global Sourcing of a Complex Good," Working paper, K.U. Leuven.
- Whinston, Michael D. 2003. "On the Transaction Cost Determinants of Vertical Integration," 19 *Journal of Law, Economics, and Organization* 1–23.
- Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York, NY: The Free Press.
- . 1985. *The Economic Institutions of Capitalism*. New York, NY: The Free Press.
- World Trade Organization. 2008. *World Trade Report 2008: Trade in a Globalizing World*. Geneva: World Trade Organization.
- Yeaple, Stephen R. 2006. "Offshoring, Foreign Direct Investment, and the Structure of U.S. Trade," 4 *Journal of the European Economic Association* 602–11.
- Yeats, Alexander J. 2001. "Just How Big Is Global Production Sharing?," in S. Arndt, and H. Kierzkowski, eds., *Fragmentation: New Production Patterns in the World Economy*. New York, NY: Oxford University Press.
- Zeile, William J. 1997. "U.S Intrafirm Trade in Goods," 78 *Survey of Current Business* 23–38.