

Managerial Incentives and the International Organization of Production*

by

Gene M. Grossman
Princeton University

and

Elhanan Helpman
Harvard University, Tel Aviv University, and CIAR

November 2002

Abstract

We develop a model in which the heterogeneous firms in an industry choose their modes of organization and the location of their subsidiaries or suppliers. We assume that the principals of a firm are constrained in the nature of the contracts they can write with suppliers or employees. Our main result concerns the sorting of firms with different productivity levels into different organizational forms. We use the model to examine the implications of falling trade costs for the relevant prevalence of outsourcing and foreign direct investment.

JEL Classification: L22, F23, D23

Keywords: outsourcing, direct foreign investment, theory of the firm, intra-firm trade

*We are grateful to Daron Acemoglu, Pol Antras, Oliver Hart and Jean Tirole for helpful discussions and to the National Science Foundation for financial support.

1 Introduction

The term “globalization” applies most aptly to a description of the modern manufacturing processes in many industries. Rather than specializing in the production of different goods from start to finish, countries increasingly contribute bits of value added to goods that end up being quite multinational in their origin. The process of vertical specialization lies behind the rapid growth in international trade of intermediate inputs, components, and specialized producer services, which has far outpaced in recent years the growth of world trade in final goods.¹

Vertical specialization takes two primary forms. Firms may procure specialized components or services from arms-length providers under contractual arrangements, or they may undertake the various production and assembly activities within the boundaries of a single firm by engaging in foreign direct investment (FDI). Borga and Zeile (2001) and Hanson et al. (2001, 2002) document and analyze the substantial rise in intra-firm trade in intermediate inputs that has taken place within multinational corporations. Contractual dealings are more difficult to isolate in the trade data, but the business press is replete with stories about foreign outsourcing.

In this paper, we develop a model that can be used to study the underlying causes of the growth of vertical specialization in trade and especially the form that such trade takes in different industries. We model the endogenous choice of organizational form by principals who are unable to monitor all of the actions undertaken by their agents. We consider an industry with many firms distinguished by their potential productivity. Each principal that enters the industry acquires the technology to produce a differentiated consumer good. But production requires the cooperation of a skilled agent who has the know-how to produce an essential component or service. The principal may hire the agent to manage a “parts division” or else contract with an entrepreneur to serve as independent supplier. The principal also faces the choice of whether to engage the agent as manager or supplier in its home country or to seek to import the intermediate inputs from a subsidiary or supplier located in a foreign land.

Our model incorporates several important trade-offs that a firm faces in its choice of location and organizational form. First, a principal who operates a vertically-integrated

¹A burgeoning empirical literature documents the rapid growth of world trade in intermediate inputs and the increasing extent of vertical specialization. See, for example, Campa and Goldberg (1997), Feenstra (1998), Hummels et al. (1998), Hummels et al. (2001), Feenstra and Hanson (2002), Borga and Zeile (2001), Yeats (2001), and Hanson et al. (2001, 2002).

firm may be better able to monitor her partner than one who deals at arms length. We capture this notion by assuming that a principal can observe a manager's efforts on some fraction of tasks, but she cannot monitor the efforts of an independent contractor at all. Moreover, the ability to monitor may vary with proximity. We assume that a principal is able to observe a manager's efforts in a larger fraction of tasks when the manager's division is located near to the firm's headquarters as compared to when it is located across national borders.

Second, the contracts that the principal can use to motivate her agent may differ in the alternative organizational forms. We do not attempt to derive the restrictions on contracting from first principals, but rather we imbue the alternative contractual relationships with realistic differences. In particular, we assume that a principal cannot ask an employee to post a bond which will be forfeited in case his efforts to serve the principal fail. Nor can the principal ask an employee to front the costs of inputs that will be put at risk in the production-sharing relationship. Rather, the principal structures a contract for the manager that pays him a non-negative wage provided that he performs satisfactorily on tasks that the principal can monitor and a bonus that he receives if the project succeeds. In an outsourcing relationship, the principal similarly cannot fine an entrepreneur for failing to deliver acceptable components (or, at least, the size of any bond that can be posted by a supplier is limited). The supply contract specifies a payment by the principal that will be paid no matter how the project turns out (for example, to defray the entrepreneur's expense for investing in the project and to compensate his efforts), and an amount that will be paid in return for delivery of acceptable components. An important difference between the organizational forms arises from the assumption that the principal bears the cost of labor, capital, and material inputs in an integrated firm, whereas the entrepreneur pays these costs at least initially when he operates a legally distinct entity.

Our model bears a familial relationship to previous research on the organization of the firm and optimal design of contracts for managers. This is a large literature, so we mention only two of many related papers. Holmström and Milgrom (1991) have modeled the choice of organizational form in a setting in which an agent must perform multiple tasks for the principal, some of which can be better observed than others. But their emphasis is on externalities in contract design; that is, on how the incentives provided for one task reflect the difficulty of measuring performance on others. They apply their reasoning to asset ownership, and show that "high-powered incentives" should be

more common when the agent owns the productive asset (outsourcing relationship) than when the principal owns the asset (employment relationship). In our model too the optimal contract for a potential supplier often provides higher-powered incentives than the optimal contract for a manager, but this has more to do with the restrictions we place on payments from the agent to the principal and on our assumptions about who initially bears the cost of labor, capital, and material inputs.

Like us, Horn et al. (1995) study the design of optimal incentive contracts for managers in a world of international trade. However, they do not consider the choice between vertical integration and arms-length dealing. Rather, they focus on whether international trade, by increasing the degree of competition in product markets, brings welfare gains that can be associated with increased effort by the manager and improved internal efficiency of the firm.

The remainder of our paper is organized as follows. In the next section, we consider the choice of organizational form by a principal with a given potential productivity. The principal can manufacture a fixed quantity of final goods if she can obtain the necessary intermediate inputs. The inputs must be produced by a skilled partner, who may manage a division of the principal's firm or head an independent supplier. The ability of the partner to deliver suitable inputs is not assured, but depends on the partner's efforts in a variety of tasks. If the principal hires the agent as an employee, she can monitor the agent's efforts in a fraction of these tasks. If the agent is hired as an independent contractor, no monitoring is possible. In either case, the principal designs an optimal contract subject to the constraints described above, and offers it to an agent with given outside opportunities on a take-it-or-leave-it basis.

Our main result in Section 2 — which allows for no choice of location — concerns the relationship between the principal's potential revenues and her preferred organizational form. We show that outsourcing is preferred by principals who have very high or reasonably low potential revenues, whereas vertical integration may offer the greatest expected profits to a principal whose potential revenues fall in an intermediate range.

In Section 3, we illustrate our result from Section 2 for a special case in which there is a piecewise linear relationship between effort on a task and its contribution to the probability of success of the venture. Then we introduce the locational dimension of the principal's decision problem and show that among firms that opt for vertical integration, those that elect to keep their parts division close to their headquarters have higher potential revenues than those that engage in FDI.

We close the model in Section 4 by specifying demand for the group of competing outputs and by allowing for endogenous entry into the industry at a given cost. Principals who pay the entry fee draw a productivity level from a known distribution of potential productivities. In equilibrium, each entrant has expected operating profits equal to the fixed entry cost. The distribution of productivities and the endogenous choice of organizational form together determine the market shares of suppliers, of foreign affiliates of multinational corporations and of integrated producers in the North.

In the succeeding sections, we study the determinants of these market shares. In Section 5, we show how improvements in the ability to monitor distant managers result in an increased market share for multinational corporations, and declines in the market shares of components produced by independent suppliers and by vertically integrated producers in the home country. Section 6 analyzes the effects of falling trade costs. There we show that trade liberalization or improvements in transportation may boost the prevalence of outsourcing or of FDI, depending on whether the industry is one in which most outsourcing is undertaken by firms that are highly productive or by firms that are the least productive among those active in the industry. A concluding section contains a summary of our results.

2 Organization of the Firm

In this section, we develop a theory of the firm based on the alternative means that a principal has to address the problems caused by imperfect observability of a manager's actions. The principal owns the technology for producing a particular product. But, to manufacture the good, she needs the cooperation of a skilled partner who can oversee the production of an essential component. The principal can hire a partner to work as a division manager, in which case she must provide the manager with the inputs needed to produce the components and structure a suitable incentive contract for him. Alternatively, she can turn to an arms-length supplier of components. Such suppliers are led by "entrepreneurs" with skills similar to those of the managers. For now, we ignore issues to do with the location of the potential parts division or supplier; later we shall allow for a choice between a local manager and one who operates in a foreign subsidiary, and between domestic and foreign outsourcing. We focus here on the decisions of a single principal, but in Section 4 we shall embed the individual's choice in a model of industry equilibrium in which firms characterized by different productivity

levels manufacture competing products.

We assume that the principal can only operate a firm of a given (maximum) size.² If the principal succeeds in acquiring suitable components either from a subsidiary or an external supplier, her output will be θ , where θ indexes the potential productivity of her firm. If she fails to acquire components, output is zero. Output generates revenue $R(\theta)$, with $R(0) = 0$ and $\lim_{\theta \rightarrow \infty} R(\theta) = \infty$. For now, we suppress the potential interactions with other firms in the industry.

The production of components requires “effort” on the part of the skilled partner — be he a manager or an entrepreneur heading a supplier firm — in a variety of tasks. Let $e(j)$ be the effort exerted by the manager or entrepreneur on task j . Then, with probability $\int_0^1 h[e(j)]dj$, the attempt to manufacture the requisite components “succeeds,” and the resulting components can be used by the principal to produce the final good. But with probability $1 - \int_0^1 h[e(j)]dj$ the project fails, and the plans to manufacture the final good must be aborted.³ We impose the following properties for the $h(\cdot)$ function:

Assumption 1 There exists a finite $E > 0$ such that (a) $0 \leq h(e) \leq 1$ for all $e \geq 0$; (b) $h'(e) > 0$ and $h''(e) \leq 0$ for all $e < E$; and (c) $h'(e) = 0$ for all $e > E$.

For simplicity, we assume that the principal cannot monitor at all the activities of an entrepreneur who operates a legally distinct firm. However, if the principal hires a division manager, it will be possible for her to observe the manager’s effort on a fraction δ of the tasks. All else equal, the ability to monitor effort on some tasks gives an advantage to in-house production relative to outsourcing. Later, we shall distinguish the fraction of tasks that can be monitored in a local plant from those that can be monitored in a foreign subsidiary. Presumably, monitoring is more difficult in a remote plant than in one that is geographically proximate to the firm’s headquarters.

The production of components requires additional inputs. Some of these may be fixed costs, independent of the scale of component production. Others may be variable costs. However, with a fixed scale of operation for the final producer, the number of components that can be processed is given, and the fixed and variable costs for the parts

²In the appendix, we show how the model can readily be extended to allow for variables scale of production.

³This is an extreme assumption that is used to simplify the algebra. The flavor of the analysis would be preserved if the manager’s efforts were to determine the productivity of the plant that manufactures components.

manufacturer need not be distinguished.⁴ We denote the cost of the inputs needed to produce the requisite quantity of components by c . These costs are paid initially by the principal in the case of in-house production and by the entrepreneur in a supplier relationship.

The manager or entrepreneur bears a private cost of effort of $e(j)$ on task j . The total utility cost of effort is $\int_0^1 e(j) dj$. Since marginal returns to effort on a single task are non-increasing and all tasks contribute similarly to the success of the venture, the optimal supply contract induces an equal effort from the entrepreneur on all tasks; call it e_o . Similarly, an optimal employment contract for a division manager generates the same level of effort e_m on all monitorable tasks, and the same level of effort e_n on non-monitorable tasks. All agents are risk neutral, and income and effort are separable in the entrepreneur's or manager's utility function. Therefore, an entrepreneur achieves expected utility of $I_o - e_o$, where I_o is the expected profits net of input costs that accrue to the supplier under an outsourcing contract. A manager enjoys an expected utility of $I_m - \delta e_m - (1 - \delta)e_n$, where I_m is the expected income that accrues to the manager under an employment contract. Skilled individuals have an outside option to achieve utility \bar{s} elsewhere in the economy. Thus, any outsourcing or employment contract must provide the entrepreneur or manager with at least this level of well-being.

We assume that the principal tenders a take-it-or-leave-it offer to the pool of skilled individuals, subject to some constraints. First, if the principal hires a skilled individual as a manager, the total compensation paid to the manager cannot be negative in any state of nature. In other words, a manager cannot be asked to post a bond as a condition of employment.⁵ Second, if the principal seeks a supplier in an outsourcing relationship, the supply contract cannot require a net payment from the supplier to the principal in any state of nature. Again, the entrepreneur cannot be asked to post a bond that is forfeited in case the attempt to produce components fails. We do not try to justify these restrictions on contracting from first principles, but rather take them to approximate realistic institutional and legal constraints that exist in many modern economies. At most, the principal can design a contract that entails a zero payment to the manager or entrepreneur in case of poor performance or an unlucky outcome; penalties or fines are not allowed.⁶

⁴In the appendix, where we allow for variable production of final goods, we distinguish the fixed costs k from the variable costs ℓ .

⁵See Katz (1986) for a discussion of the theoretical and practical difficulties that inhibit the use of performance bonds for employees.

⁶We do not actually need the limit on the smallest payment in case of an unsuccessful project

With these restrictions on the feasible contracts, we see a second difference between in-house production and outsourcing. When components are manufactured in a wholly-owned subsidiary, the principal pays the costs of the primary inputs. Then, if the project fails, the principal stands to lose this investment. In contrast, in an outsourcing relationship it is the entrepreneur who fronts the cost of the inputs, unless the principal chooses to include a fixed payment for this purpose in the contract offer. This means that an entrepreneur may have more at stake than a manager and it opens the possibility that higher-powered incentives can be offered under this arrangement. Also, in case the principal finds it optimal to design a contract that leaves (expected) rents to her skilled partner, the ability to shift input costs to a supplier but not to a manager may affect the relative attractiveness to the principal of the alternative organizational forms. When suppliers and managers cannot be asked to post bonds, the fact that the supplier pays the up-front cost of the inputs into parts production tends to favor outsourcing relative to in-house production from the perspective of the principal. This cuts against the advantage of in-house production that stems from the opportunity it affords the principal to monitor some of the manager's actions.

We proceed now to derive the (constrained) optimal contracts under each organizational form. An outsourcing contract is characterized by an amount s that the principal promises to pay the entrepreneur whether or not the project succeeds and an amount p_o that the principal will pay in case the supplier is able to provide the components that are needed for production of the final good. The payment of s allows for the possibility of (endogenous) cost sharing, while our restriction constrains the contract to have $s \geq 0$.

An employment contract is characterized by a level of effort e_m that the manager is expected to exert on all tasks that can be monitored by the principal, a wage payment w that the manager will receive irrespective of the outcome of the project provided that he has exerted at least the indicated level of effort on the monitorable tasks, and a bonus b that he will receive in case the project succeeds. Our restriction requires $w \geq 0$.

In deriving the optimal offers, we will make use of the optimal responses of an entrepreneur or manager to the incentives that are provided in his contract. An entrepreneur chooses e_o to maximize his expected utility, $s + p_o h(e_o) - e_o$. A manager

to be zero; a small enough finite negative number would suffice to yield qualitatively similar results. Note too that the exogenous restrictions on the negative payments would not be needed if we were to assume that managers and entrepreneurs are risk averse.

chooses \tilde{e}_m and e_n to maximize $w + b[\delta h(\tilde{e}_m) + (1 - \delta)h(e_n)] - \delta\tilde{e}_m - (1 - \delta)e_n$, subject to the constraint that the level of effort on the monitorable tasks \tilde{e}_m must be at least as great as the effort e_m specified in the contract. The reader may verify later that the manager has no incentive to exert “extra” effort on the monitorable tasks. So we set $\tilde{e}_m = e_m$ in this problem. Then note that the payment p_o that must be provided to an entrepreneur to induce an effort e on all tasks is the same as the bonus b that is needed to induce that same level of effort from a manager on the monitorable tasks. We will use the function $\phi(e)$ to denote the *smallest* size of the incentive needed to induce an effort level e on an unobservable task; then $p_o = \phi(e_o)$ and $b = \phi(e_n)$. We can now regard the principal as if she were choosing the efforts exerted by an entrepreneur or a manager subject to the constraint that these choices must be incentive compatible. Note that $\phi(0) = 0$ and $\phi(e) = 1/h'(e)$ for $0 < e < E$.⁷

2.1 Operating Profits under Outsourcing

Suppose the principal chooses to outsource the production of components. Such a principal must choose s and e_o to maximize

$$\Pi_o = h(e_o)R(\theta) - s - h(e_o)\phi(e_o) , \quad (1)$$

the difference between her expected revenues and her expected total payments to her parts supplier. The principal’s choices are constrained by the requirements that $s \geq 0$ and

$$s + h(e_o)\phi(e_o) - c - e_o \geq \bar{s} . \quad (2)$$

The latter is a *participation constraint*, ensuring that the entrepreneur’s expected utility (equal to his expected income less the cost of the primary inputs and the utility cost of his efforts) is no less than the utility he could attain elsewhere in the economy. Of course, the principal could always choose not to engage with any supplier or to produce any output, in which case her operating profits will be equal to zero.

⁷Consider the problem

$$\max_{e>0} zh'(e) - e ,$$

which has the first-order condition $zh'(e) \leq 1$ and the complementary-slackness condition $[zh'(e) - 1]e = 0$. Denote the solution by $e^*(z)$. The solution has $e^* = 0$ for low values of z whenever $h'(0)$ is finite. It is an increasing function of z whenever $0 < e^* < E$. The reward function $\phi(e)$ is zero for $e = 0$, it is the inverse of $e^*(\cdot)$ whenever $0 < e^*(\cdot) < E$, and it is equal to the smallest value of z for which $e^*(z) = E$ for $e = E$. Note that $\phi(e)$ is discontinuous at $e = 0$ whenever $h'(0)$ is finite.

Let us ignore the non-negativity constraint on s for a moment. Without this constraint, the principal's operating profits would be maximized by a choice of s that would make the participation constraint (2) hold as an equality and the choice of e_o that maximizes $h(e_o)R - e_o$. We denote the effort level that maximizes $h(e)R - e$ by $e^*(R)$. Note that $e^*(0) = 0$ and that $e^*(R)$ is a non-decreasing function for all $R < R_{oE}$, where R_{oE} is the lowest value of R for which $e^*(R) = E$.⁸ If, when $e_o = e^*(R)$, the value of s that makes (2) hold as an equality is positive, then the principal will set $e_o = e^*(R)$ and s equal to

$$s^*(R) = c + \bar{s} + e^*(R) - h[e^*(R)]\phi[e^*(R)].$$

Indeed, the non-negativity constraint does not bind for low values of R , since $s^*(0) = c + \bar{s} > 0$ and $s^*(\cdot)$ is continuous at $R = 0$.⁹ But $s^*(R)$ is a non-increasing function of R for $R < R_{oE}$.¹⁰ It follows that the non-negativity constraint on s may bind for a range of higher values of R . Since the analysis is more interesting when this is so, we adopt

Assumption 2 $h(E)/h'(E) - E > \bar{s} + c > h(0)/h'(0)$.

The first inequality in Assumption 2 ensures the existence of an R_s between 0 and R_{oE} with the property that $s^*(R) \geq 0$ for all $R \leq R_s$ and $s^*(R) < 0$ for all $R > R_s$.

Now consider again the problem facing the principal who has chosen to deal with an arms-length supplier. If the firm's productivity θ is such that $R(\theta) \leq R_s$, the principal chooses $e_o = e^*(R)$ and $s = s^*(R)$. At small values of $R(\theta)$ such that the optimal contract induces no effort from the entrepreneur, $\Pi_o(R)$ must be negative by the second inequality in Assumption 2.¹¹ For larger values of R that imply above-

⁸The difference $h(e)R - e$ may be maximized by $e^* = 0$ for a range of small values of R . When $0 < e^*(R) < E$, $e^*(\cdot)$ is the inverse of the function $\phi(\cdot)$. Using the first-order condition, it is easy to verify that

$$R_{oE} = \frac{1 + \phi'(E)h(E)}{h'(E)},$$

where the derivatives in this expression are interpreted as left-hand-side derivatives.

⁹The function $s^*(\cdot)$ may be discontinuous only at the single point where $R = 1/h'(0)$; i.e., the point where effort turns positive and the bonus may jump.

¹⁰ $ds^*/dR = -(h\phi')(de^*/dR)$, which is zero or negative for $R < R_{oE}$.

¹¹We do not need the second inequality in Assumption 2 for any of our substantive results. Without this assumption, there is the possibility that an outsourcing operation may be profitable with no effort from the entrepreneur; this just adds another (uninteresting) case to the analysis. Note that the second inequality in Assumption 2 will be satisfied if the probability of success of a venture is small when the entrepreneur invests no effort, or if the marginal productivity of the first bit of effort is large.

minimum effort, $e^*(\cdot)$ is the inverse of the bonus function $\phi(\cdot)$. But this means that $p_o = \phi[e^*(R)] = R$ and therefore that $\Pi_o(R) = -s^*(R) \leq 0$. It follows that outsourcing cannot be profitable for any $R \leq R_s$ under Assumptions 1 and 2.

Accordingly, we can limit our attention to firms with productivity levels such that $R(\theta) > R_s$. These firms set $s = 0$. Then the participation constraint for the entrepreneur implies $h(e_o)\phi(e_o) - e_o \geq c + \bar{s}$. The maximization of operating profits $\Pi_o(R)$ subject to this constraint requires

$$Rh'(e_o) = 1 + (1 - \lambda)\phi'(e_o)h(e_o)$$

and

$$\lambda[h(e_o)\phi(e_o) - e_o - c - \bar{s}] = 0,$$

where $\lambda \geq 0$ is the Lagrange multiplier on the participation constraint. It follows that for R between R_s and $R_{o\lambda}$, $e_o = e^*(R_s)$, where

$$R_{o\lambda} = \frac{1 + \phi'[e^*(R_s)]h[e^*(R_s)]}{h'[e^*(R_s)]} \quad (3)$$

is the largest value of revenues for which the participation constraint binds ($\lambda > 0$). In other words, in the range between R_s and $R_{o\lambda}$, the effort expected of the entrepreneur is constant at the (unique) level that leaves him with the same utility as he could attain elsewhere. For higher levels of revenue than $R_{o\lambda}$, it is optimal for the principal to leave some rents for the entrepreneur in order to induce a higher level of effort than $e^*(R_s)$. Note that $R_{o\lambda} < R_{oE}$, so there always exists a range of revenues for which the participation constraint does not bind.

Figure 1 depicts the operating profits for a firm that chooses to outsource its parts production, as a function of R (and implicitly of θ). For $R < R_s$, these profits are negative, as we have already noted. A firm with potential revenues in this range prefers to shut down than to contract with an outsourcing partner, considering the cost of such a contract and the prospects for success. At $R = R_s$, the revenues just cover the operating costs, i.e., $\Pi_o(R_s) = 0$. Then, for a range of revenues between R_s and $R_{o\lambda}$, effort is constant, and operating profits are a linear function of R given by

$$\Pi_o(R) = Rh[e^*(R_s)] - \phi[e^*(R_s)]h[e^*(R_s)].$$

For R between $R_{o\lambda}$ and R_{oE} , the participation constraint does not bind, effort increases

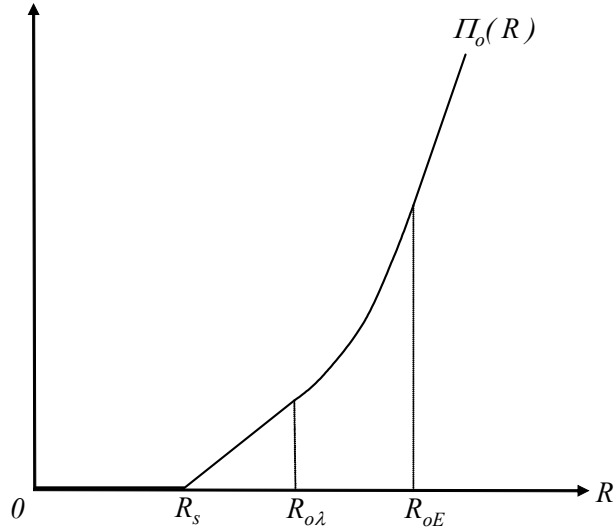


Figure 1: Expected Operating Profits from Outsourcing

with R , and operating profits are a convex function of R given by

$$\Pi_o(R) = \max_{e \geq 0} [R - \phi(e)]h(e).$$

Finally, the first inequality in Assumption 2 ensures that the principal will induce an effort of E for $R \geq R_{oE}$. For these highly productive firms, operating profits again are a linear function of R given by

$$\Pi_o(R) = [R - \phi(E)]h(E).$$

2.2 Operating Profits under Vertical Integration

Now suppose that the principal chooses instead to manufacture her own components by hiring a manager to oversee a parts division. The principal must choose a contract for the manager that specifies a wage, an expected level of effort on monitorable tasks, and a bonus for success. Equivalently, we can think of the manager as choosing w , e_m and e_n to maximize

$$\Pi_v(R) = [\delta h(e_m) + (1 - \delta)h(e_n)][R - \phi(e_n)] - w - c, \quad (4)$$

subject to

$$[\delta h(e_m) + (1 - \delta)h(e_n)]\phi(e_n) + w - \delta e_m - (1 - \delta)e_n \geq \bar{s}. \quad (5)$$

Here, the operating profits are the expected revenues net of the expected bonus payment, the cost of the inputs, and the wage of the manager. The constraint ensures that the welfare of the manager is at least as great as what he could attain by working elsewhere. The principal also is constrained to offer a non-negative wage rate ($w \geq 0$) and to choose non-negative levels of effort for both monitorable and non-monitorable tasks.

As with outsourcing, the principal will not be constrained by the requirement that wages be non-negative if the firm's productivity, and thus the potential revenues, are very low. For small values of $R(\theta)$, it is optimal for the principal to design the employment contract so that $e_m = e_n = e^*(R)$ and $w = w^*(R)$, where

$$w^*(R) = \bar{s} + e^*(R) - h[e^*(R)]\phi[e^*(R)].$$

As long as the implied wage is not negative, this contract achieves the first-best for the principal, with an efficient level of effort by the manager on all tasks and a compensation package that leaves him with no rents.

We can define a revenue level R_w that is analogous to R_s ; R_w is the largest level of revenues at which the principal achieves the first-best outcome. Note that $R_w \leq R_s$, because $w^*(R) = s^*(R) - c$ and so if $w^*(R)$ is non-negative, $s^*(R)$ certainly is non-negative. The principal can attain the first-best outcome with outsourcing for a wider range of revenue levels than with in-house production, because the outsourcing payment compensates the manager not only for his efforts and the opportunity cost of his time, but also for the cost of the inputs that he uses to manufacture components.

It follows immediately that outsourcing is more profitable than in-house production for potential revenues near R_s . With outsourcing, the principal induces an efficient level of effort at R_s and pays the entrepreneur just enough to satisfy the participation constraint. In an integrated firm, the wage at R_s that would leave a manager who exerts the efficient level of effort on all tasks indifferent between working for the principal and pursuing an outside option is negative. So the principal must either pay rents to the manager or sacrifice efficiency on at least some tasks. Note that $\Pi_o(R_s) = 0$, so $\Pi_v(R_s) < 0$.

When potential revenues are sufficiently greater than R_s , the principal can achieve

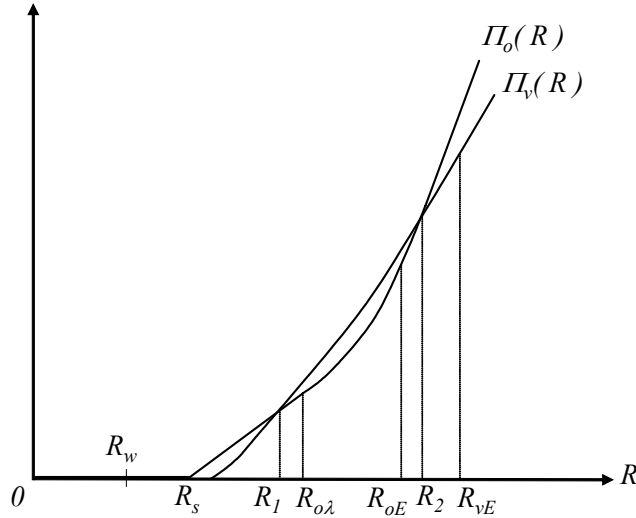


Figure 2: Outsourcing versus In-House Production

positive expected profits with in-house production. The optimal contract is one that maximizes (4) subject to (5) and $w \geq 0$. If potential revenues are very high, the principal will opt to have the manager exert maximal effort on all tasks. In the event, the manager will capture positive rents. For more moderate levels of revenues, the principal will tailor the contract so as to induce $e_m > e^*(R) > e_n$, and will accept an inefficient allocation of the manager's effort across tasks. The participation constraint binds whenever $e_m < E$, for otherwise the principal could demand a higher level of effort on the monitorable tasks at no cost to herself.

We are now ready to compare the expected operating profits under the alternative organizational forms. Figure 2 depicts this comparison for different levels of revenue and thus, implicitly, for different productivity levels of the final-good producer. For revenue levels below R_s , outsourcing and in-house production both generate expected operating losses. A principal that anticipates such a low level of potential revenues prefers to shut down than to engage an agent as either a supplier or a manager. A more productive principal, whose potential revenues slightly exceed R_s , can earn slightly positive expected profits by contracting with an external supplier, but would lose most of the input costs c by employing a manager. Such a principal, and indeed any one with potential revenues between R_s and R_1 , prefers outsourcing to in-house production. Outsourcing also yields higher expected profits to firms with very high

levels of productivity. Consider, for example, $R \geq R_{vE}$, where R_{vE} is the smallest level of revenue at which it is optimal for the principal to induce maximal effort on all tasks of the manager.¹² It is easy to show that $R_{vE} > R_{oE}$, i.e., when the principal would induce maximal effort on all tasks in an integrated firm, the optimal outsourcing contract would also induce maximal effort by the entrepreneur. With maximal effort and in-house production, the expected operating profits for the principal are

$$\Pi_v(R) = [R - \phi(E)]h(E) - c.$$

This is less than what the principal could achieve by outsourcing; the difference is simply the cost of the inputs needed to manufacture the components. In each case, when effort is maximal, the skilled partner captures some rents. But the rents are smaller with outsourcing than with an integrated production unit, because the input costs can be shifted to the partner when the components are produced by a separate firm but not when they are produced in house.

For intermediate levels of productivity, the principal may prefer vertical integration to arms-length dealing. Figure 2 shows this to be the case, as it must be when δ is close to one. Consider the potential operating profits under the alternative organizational forms when the firm's productivity is such that $R = R_{oE}$. Under outsourcing, the principal induces the entrepreneur to provide effort at level E , and operating profits are given by $\Pi_o(R_{oE}) = [R_{oE} - \phi(E)]h(E)$. If the principal were to hire a manager instead, she could insist on an effort level of E on the tasks that can be monitored while allowing zero effort on tasks that cannot be monitored. Such an effort profile would require no bonus payment from the principal (since $\phi(0) = 0$), but would require that the manager be paid a wage of at least $\bar{s} + \delta E$ to match his outside option. By choosing a contract with $e_m = E$, $w = \bar{s} + \delta E$ and $b = 0$, the principal would achieve an expected operating profit of $R_{oE}[\delta E + (1 - \delta)h(0)] - \delta E - c - \bar{s}$. Of course, this contract might not be optimal, which means that $\Pi_v(R_{oE})$ is at least as large as this amount. It follows that

¹²It is optimal for the principal to induce maximal effort on the tasks that cannot be monitored when $(1 - \delta)[Rh'(E) - 1] - h(E)\phi'(E) \geq 0$, or when

$$R \geq R_{vE} = \frac{1 - \delta + h(E)\phi'(E)}{(1 - \delta)h'(E)}.$$

$$\Pi_v(R_{oE}) - \Pi_o(R_{oE}) \geq h(E)\phi(E) - \delta E - c - \bar{s} - (1 - \delta)[h(E) - h(0)].$$

But Assumption 2 ensures that the right-hand side of this inequality is positive for δ close to one. In short, if the manager can be monitored in a large enough fraction of tasks, then the benefits of vertical integration outweigh the costs for some intermediate range of productivity levels.

To summarize, we have shown that a principal prefers to buy components from an independent parts supplier when her productivity is sufficiently high or rather low. Between these extremes there may be a range of productivity levels for which vertical integration is the preferred mode of organization. The advantage of outsourcing at high levels of productivity comes from the opportunity it affords a principal to reduce the rents that must be granted to the agent in a situation in which it is optimal for her to tolerate some rent sharing in order to induce maximal effort. Outsourcing cuts into the agent's rents, because the cost of the inputs used to produce the parts can be shifted to the supplier.¹³ For low levels of productivity, outsourcing is advantageous for a different reason. Here, the efficient level of effort is relatively low and the agent captures no rents. But for any level of effort, the contingent payment under outsourcing that leaves the entrepreneur without rents is larger than the bonus payment to a manager that similarly drives him to his reservation level of utility. This is because a contingent payment p_o must compensate an entrepreneur for the cost of the inputs as well as the disutility of his effort, while a bonus payment b repays only the effort. It follows that a principal can induce greater effort from an entrepreneur than from a manager for a given expected outlay. Put differently, an entrepreneur who bears the cost of the inputs has more at stake in a project than a manager who does not. When the principal brings the former to utility level \bar{s} , the resulting incentives have higher power than those that would bring a manager to the same level of expected utility.

The possible advantage of vertical integration for an intermediate range of productivity levels stems from the opportunity it affords the principal to monitor some of the manager's actions. On tasks that can be monitored, the principal can induce a high level of effort without having to leave rents to the manager. She can do so simply by paying a wage that compensates the manager for his effort on these tasks, and de-

¹³We show in the appendix that when the principal can vary her scale of operation, she leaves no rents to the entrepreneur even when potential revenues are high. However, the comparison between outsourcing and in-house production at different levels of potential revenue is qualitatively the same as what we describe in the main text.

manding that the effort be made. If enough tasks can be monitored, the principal can achieve as high a probability of success with integration as with outsourcing without having to share rents.

3 Choice of Location

In the first part of this section, we develop a special case of the model in Section 2 in which the contribution of effort on a task to the probability that the project will succeed is a piecewise linear function of the effort level. By using this simple, special case we are able to address the principal's location problem most clearly. In the latter part of the section, we allow for outsourcing in two regions distinguished by costs and also allow an integrated firm to operate a local subsidiary which affords relatively better opportunities for monitoring the manager or a remote subsidiary which affords relatively worse such opportunities.

3.1 A Special Case: Piecewise Linearity

In this section, we take the $h(\cdot)$ function to be piecewise linear, with $E = 1$ and $h(e) = 1$ for $e \geq 1$. We define h_0 , h_1 , and e_1 such that $h(0) = h_0$ and $h(e_1) = h_1$, as shown in Figure 3. The marginal productivity of effort is constant for efforts between 0 and e_1 , and also constant for efforts between e_1 and $E = 1$. Assumptions 1 and 2 place certain restrictions on the parameter values, which we write as

Assumption 1' $(h_1 - h_0)/e_1 > (1 - h_1)/(1 - e_1)$

and

Assumption 2' $(h_1 - e_1)/(1 - h_1) > \bar{s} + c > e_1 h_0 / (h_1 - h_0)$.

When the $h(\cdot)$ function is piecewise linear, the bonus function $\phi(\cdot)$ is discontinuous at both $e = 0$ and $e = e_1$. To induce an effort of e_1 from the entrepreneur on any task, or from the manager on a task that cannot be monitored, the principal must reward a success with a contingent payment of $e_1/(h_1 - h_0)$. To induce the higher effort of $e = 1$, the payment must be $(1 - e_1)/(1 - h_1)$.¹⁴

The potential operating profits from outsourcing are depicted by the solid, kinked curve in Figure 4. At very low levels of productivity such that $R(\theta) < R_a = (\bar{s} + c + e_1)/h_1$, the principal prefers to leave the market than to engage a supplier or a manager

¹⁴Recall that $\phi(e)$ is defined as the smallest bonus that induces the effort e .

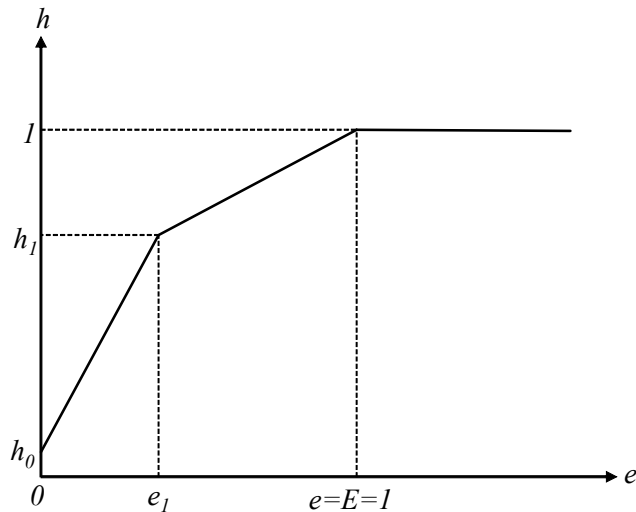


Figure 3: Piecewise Linear Productivity of Effort

and thereby suffer negative expected profits. For operating profits to be positive, potential revenues must be at least as great as R_a . At this revenue level, the principal offers an incentive payment of $e_1/(h_1 - h_0)$, thereby inducing the entrepreneur to exert an effort of e_1 on all tasks. The principal makes a fixed payment to the entrepreneur of $s = \bar{s} + c + e_1 - e_1 h_1/(h_1 - h_0)$ to ensure that the entrepreneur is willing to accept the contract.¹⁵ The project succeeds with probability h_1 and the principal earns expected operating profits of $\Pi_o(R_a) = 0$.

For a range of revenue levels above R_a , it remains optimal for the principal to write a contract that induces effort e_1 by the entrepreneur. In this range, expected operating profits are given by $\Pi_o(R) = R h_1 - e_1 - \bar{s} - c$. But when productivity is such that $R(\theta) \geq R_c = [(1 - e_1)/(1 - h_1)^2] - [(\bar{s} + c + e_1)/(1 - h_1)]$, the principal prefers to induce the maximal effort level of $E = 1$ by offering a contingent payment of $(1 - e_1)/(1 - h_1)$. Here, the non-negativity constraint for s binds. The principal sets $s = 0$ and achieves expected operating profits of $\Pi_o(R) = R - (1 - e_1)/(1 - h_1)$.

If the principal opts for in-house production of components, it is never optimal for her to induce the manager to work harder on the tasks that cannot be monitored than on those that are observable. We focus therefore on strategies that involve $e_m \geq e_n$. There are now several possibilities to consider. We can rule out $e_m = e_n = 0$,

¹⁵Note that $s > 0$ in light of the second inequality in Assumption 2'.

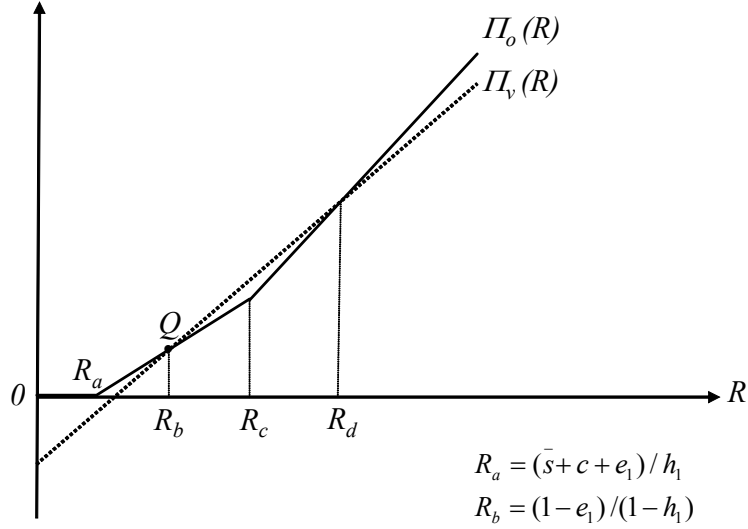


Figure 4: Outsourcing vs. In-House Production: A Special Case

because with these levels of effort, expected operating profits are always negative under Assumptions 1' and 2'. Also, by choosing $e_m = e_1$ and $e_n = 0$, the principal earns lower expected profits in an integrated operation than she does by outsourcing for all revenue levels at which expected profits under this integration strategy are positive.¹⁶ We can therefore disregard this strategy as well.

Another possible strategy has $e_m = e_n = e_1$. This requires a bonus offer of $b = e_1/(h_1 - h_0)$ and may or may not require a positive wage payment in order that the contract satisfies the participation constraint. If $\bar{s} > e_1 h_0 / (h_1 - h_0)$, then $w > 0$ when $e_m = e_n = e_1$, and the principal attains expected operating profits of $Rh_1 - e_1 - \bar{s} - c$. For revenue levels between R_a and R_c , these profits are exactly the same as what she can earn by outsourcing. Recall from Section 2 that generically there exists a range of low productivity levels at which the principal prefers outsourcing to in-house production. Since this preference is only a weak one when $\bar{s} > e_1 h_0 / (h_1 - h_0)$, we instead adopt

Assumption 3' $\bar{s} < e_1 h_0 / (h_1 - h_0)$.

Under Assumption 3', the manager of a parts division captures rents when $e_m = e_n = e_1$, $b = e_1 / (h_1 - h_0)$, and $w = 0$. The associated operating profits for the principal are

¹⁶With $e_m = e_1$ and $e_n = 0$, the principal sets $b = 0$ and $w = \bar{s} + \delta e_1$. Operating profits are $\Pi_v(R) = R[\delta h_1 + (1 - \delta)h_0] - \delta e_1 - \bar{s} - c$, which is positive when $R > (\bar{s} + c + \delta e_1) / [\delta h_1 + (1 - \delta)h_0]$. But for such values of R , the principal can earn profits of $Rh_1 - e_1 - \bar{s} - c$ from outsourcing by inducing effort of $e = e_1$. Assumptions 2' ensures that these profits are higher.

$Rh_1 - (h_1e_1)/(h_1 - h_0) - c$, which is less than what she could earn by outsourcing.

In-house production also yields lower expected profits than outsourcing when the principal would set $e_m = e_n = 1$. To induce an effort of $e = 1$ on tasks that cannot be monitored, the principal must offer a bonus payment of $b = (1 - e_1)/(1 - h_1)$. But then the non-negativity constraint on the wage binds, and the principal must leave rents to the manager. The resulting profits are $R - [(1 - e_1)/(1 - h_1)] - c$, which is less than what the principal could earn by outsourcing and inducing an effort level of $e = 1$ on all tasks.

The possible attractiveness of in-house production to the principal comes when the optimal contract entails $e_m = 1$ and $b = e_1/(h_1 - h_0)$, so that $e_n = e_1$. This can happen for an intermediate range of revenue levels, which is in line with our more general findings in Section 2. When $e_m = 1$ and $e_n = e_1$, the non-negativity constraint on the manager's wage may or may not bind. In either case, in-house production may yield higher operating profits to the principal than outsourcing for some values of R . However, to avoid a taxonomy, we shall focus on the case in which the principal pays a positive wage $w > 0$ in order to satisfy the manager's participation constraint. This case arises when

$$\bar{s} > [\delta + (1 - \delta)h_1] \left(\frac{e_1}{h_1 - h_0} \right) - \delta - (1 - \delta)e_1. \quad (6)$$

Then vertical integration with $e_m = 1$ and $e_n = e_1$ yields expected operating profits to the principal of $\Pi_v(R) = R[\delta + (1 - \delta)h_1] - [\delta + (1 - \delta)e_1] - \bar{s} - c$. These profits are depicted by the dotted line in Figure 4.¹⁷

We see in the figure that in-house production dominates outsourcing from the principal's perspective for R between R_b and R_d . It is easy to calculate that $R_b = (1 - e_1)/(1 - h_1)$, which is to the right of R_a and independent of δ . It follows that when inequality (6) is satisfied, there always exists a range of revenue levels for which the

¹⁷An alternative strategy for in-house production is for the principal to require $e_m = 1$, pay a wage $w = \bar{s} + \delta$, and offer no bonus. Facing such a contract, the manager would make no effort on tasks that are not monitored and would be indifferent between participating in the contract and not. The alternative strategy yields expected profits for the principal of $R[\delta + (1 - \delta)h_0] - \delta - \bar{s} - c$, which — for all $R \geq R_a$ — is strictly less than the profits of $R[\delta + (1 - \delta)h_1] - [\delta + (1 - \delta)e_1] - \bar{s} - c$ that she earns by pursuing the strategy described in the text. However, when inequality (6) is violated, the strategy of accepting $e_n = 0$ may be the best option available to the principal for some revenue levels. Then the figure would look somewhat different from what we have drawn in Figure 4, but the main conclusion about the choice of organizational form for firms with different productivity levels would remain the same.

principal prefers in-house production to outsourcing. This range is larger the greater is δ , because an increase in δ causes the dotted line to rotate in a counterclockwise direction around point Q .

Figure 4 reproduces the general pattern that we identified in Section 2. Firms with productivity such that $R(\theta) < R_a$ cannot achieve positive expected operating profits by either mode of organization, and so they exit the industry. Firms with potential revenues between R_a and R_b and those with potential revenues greater than R_d prefer to outsource; all others prefer to produce their components in house.

3.2 Choice of Location for Parts Production

At last, we are ready to address the principal’s problem about where to locate the parts production. There are two regions, North and South, with associated variables represented by subscripts N and S , respectively. The principal has her headquarters in the North. If she chooses to outsource the production of components, she may choose a supplier in the North or in the South. We assume that input costs are lower in the South ($c_S < c_N$) and that the outside options for those with the skills needed to head a production unit are no better there ($\bar{s}_S \leq \bar{s}_N$).¹⁸ The principal might also create a subsidiary to produce components in either the North or the South, with such a plant managed by a local employee. We assume that the principal is able to monitor the manager’s efforts in a greater fraction of tasks when the production unit is located in the North than when it is located in the distant South ($\delta_N > \delta_S$).

Note that our model gives no advantage to outsourcing in the North to compensate for the higher costs there. We might, for example, have allowed the principal to monitor some (small) fraction of an entrepreneur’s actions, with greater opportunities for this in the North than in the South. Then outsourcing in the North might have become viable for some productivity levels. Or we might have allowed for differences in the “thickness” of the markets for components or in the completeness of enforceable contracts in the alternative legal environments. The role that these latter differences play in the location of outsourcing activity was the focus of Grossman and Helpman (2002). Here we prefer to keep matters simple, and so we accept that outsourcing in the North is a dominated option in this setting.

However, foreign direct investment (FDI) and in-house production of parts in the North both may be viable options for some parameter values and some productivity

¹⁸Actually, our qualitative results require only that $c_S + \bar{s}_S < c_N + \bar{s}_N$.

levels. The former affords an opportunity for monitoring some of the manager's efforts without giving up the advantage of low costs. The latter provides still better opportunities for monitoring, but costs are higher. Our next task is to compare the three options of outsourcing (in the South), home production and FDI for different levels of productivity. We can readily show that vertical integration in location j can only compete with outsourcing when the principal would choose $e_{mj} = 1$ for the manager's effort on tasks that can be monitored in location j and $e_{nj} = e_1$ for tasks that cannot be monitored. So we limit our attention to these strategies.

First note that for low levels of R at which foreign outsourcing is only marginally profitable, outsourcing is the preferred mode of organization for the principal. The comparison between foreign outsourcing and FDI is the same as we conducted in Section 3.1; the former dominates at low levels of productivity and revenues, because the principal can structure higher-powered incentives for an entrepreneur who must front the cost of inputs than she can for a manager who must be paid a non-negative wage no matter what the outcome of the project. The principal in turn prefers FDI to home in-house production when inequality (6) is satisfied for (at least) all revenue levels $R \leq R_b = (1 - e_1)/(1 - h_1)$. To see this point, recall that an increase in δ rotates the line representing profits under vertical integration around the point Q in Figure 4. This means that the fraction of monitorable tasks has no bearing on operating profits under vertical integration when $R = R_b$. But the lower costs in the South give FDI a clear advantage over in-house production in the North at $R = R_b$; in fact, $\Pi_{vS}(R_b) - \Pi_{vN}(R_b) = (\bar{s}_N + c_N) - (\bar{s}_S + c_S) > 0$.

Figure 5 shows the maximal expected operating profits for a firm with potential revenues R . Implicitly, this figure reveals the optimal choice of organization and location for each value of R . For $R < R_a = (\bar{s}_S + c_S + e_1)/h_1$, expected operating profits must be negative for each of the three organizational forms, and so exit is the best available strategy. For $R = R_a$, outsourcing in the South achieves zero expected profits. Next comes a range of revenue levels between R_a and R_b for which outsourcing is profitable and the best of the three alternatives. As we just described, outsourcing offers low costs to the principal and affords the opportunity for her to structure high-powered incentives for the entrepreneur without sharing rents. When $R = R_b$, the expected profits for FDI match those for outsourcing and exceed those available to a firm that produces its own components in the North. In the range between R_b and R_v , the principal chooses FDI over outsourcing, because she values the ability to monitor the

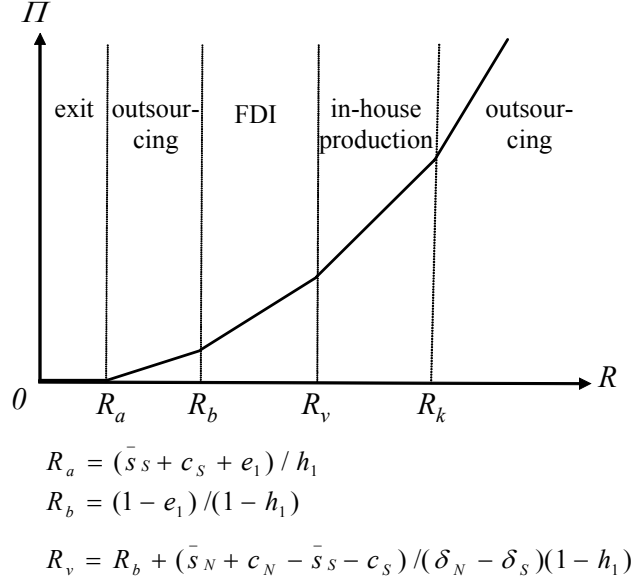


Figure 5: Expected Profits for Piecewise Linear Productivity of Effort

manager on a fraction δ_S of the tasks. But in this range, the benefit of lower costs still outweighs the cost of less monitoring in the comparison between FDI and in-house production at home.

The figure shows $R_v = R_b + [(\bar{s}_N + c_N) - (\bar{s}_S + c_S)] / [(\delta_N - \delta_S)(1 - h_1)]$ to be the revenue level at which FDI and in-house production in the North yield equal expected profits. At $R = R_v$, the cost savings that favor FDI are matched by the benefits from closer monitoring of the manager's efforts. Since potential revenues are reasonably high in the range above R_v , the principal places great importance on achieving a high probability that the project will succeed. She is willing to pay more for inputs and to compensate the manager more handsomely in order to mitigate the damage caused by the manager's moral hazard.

Finally, highly productive principals prefer foreign outsourcing to in-house production in the North and to FDI. Specifically, when θ is such that $R(\theta) > R_k$, where

$$R_k = \frac{\frac{1-e_1}{1-h_1} - [\delta_N + (1-\delta_N)e_1 - (\bar{s}_N + c_N)]}{1 - [\delta_N + (1-\delta_N)h_1]},$$

the principal achieves higher expected operating profits by outsourcing than by opening

a local subsidiary.¹⁹ The reason is that when very large revenues are at stake the principal will do whatever is necessary to ensure maximal effort by the entrepreneur or manager and a high probability of success. To induce an effort of $e = 1$ on all tasks, the principal must share rents with the entrepreneur or manager. But the principal foregoes fewer rents with outsourcing than with (any form of) in-house production, because the input bill that is initially paid by the entrepreneur serves as a tax on his take of rents.

To summarize, we have shown that the firms in an industry have different incentives to open foreign subsidiaries and to engage in foreign outsourcing depending on their productivity levels. Our model suggests that the least and most productive firms will turn to external suppliers for the component needs, while the firms that operate foreign subsidiaries will be less productive than those that manufacture their own components in a plant nearer to their headquarters.

4 Industry Equilibrium

In this section, we embody our model of a firm's choice of organizational mode and location in a setting of industry equilibrium. We assume that principals can enter the industry by bearing a fixed entry cost of f . Those that pay this cost draw a productivity level from a known distribution $G(\theta)$, just as in Melitz (2002) and Helpman, Melitz and Yeaple (2002). The firms then choose their organizational form (including location) and design an optimal purchase or employment contract in the light of their decision. Firms that are successful in acquiring components manufacture differentiated products that compete for consumers' spending.

We assume a world populated by many consumers, each with the utility function $u = y_0 + y^\eta/\eta$, where y_0 is consumption of a homogenous good and y is an index of consumption of the varieties of the differentiated product.²⁰ The elasticity of demand for the group of differentiated products (with respect to an ideal price index) is $1/(1 - \eta)$, where $\eta < 1$. We aggregate consumption of individual varieties using the familiar, CES

¹⁹There is no guarantee that $R_k > R_v$ for all parameter values. If δ_N is not very much larger than δ_S or if $c_N + \bar{s}_N$ is very much greater than $c_S + \bar{s}_S$, then there will be no values of R for which the principal prefers in-house production in the North to both FDI and foreign outsourcing.

²⁰This measure of utility does not include the separable component of utility loss associated with the effort that an individual exerts on her job. Note that we could accommodate many industries with differentiated products by assuming $u = y_0 + \sum_l \omega_l y_l^{\eta_l} / \eta_l$; then our analysis here would apply to any industry l .

preference function

$$y = \left[\int y(j)^\alpha dj \right]^{1/\alpha}, \quad 0 < \alpha < 1,$$

where $y(j)$ is consumption of variety j . With this specification, $1/(1-\alpha)$ is the elasticity of demand for variety j and a higher value of α implies greater substitution across varieties and a higher demand elasticity for each brand. We assume $\alpha > \eta$, so that the varieties substitute more closely for one another than does the group of differentiated products substitute for the numeraire good.²¹

We normalize the measure of consumers to equal one. Then the aggregate world demand for variety j is described implicitly by the first-order condition

$$y^{\eta-\alpha} y(j)^{\alpha-1} = p(j),$$

where $p(j)$ is the price of brand j . A firm's revenue from selling brand j is $y^{\eta-\alpha} y(j)^\alpha$.

If a firm with productivity θ is successful in obtaining components, it can produce θ units of output and accrue revenue of

$$R(\theta) = y^{\eta-\alpha} \theta^\alpha. \quad (7)$$

Clearly, a firm's potential revenue is the product of two components; a component that depends on aggregate industry characteristics, summarized by y , and an idiosyncratic component, θ . The cumulative distribution function of θ , $G(\theta)$, induces a cumulative distribution on R .

Equation (7) can be used to map each revenue cutoff level in Figure 5 into a productivity cutoff level. For example, since R_v is the revenue level at which a principal is just indifferent between in-house production in the North and FDI, $\theta_v = R_v^{1/\alpha} y^{(\alpha-\eta)/\alpha}$ is the productivity level that generates such indifference. And similarly for the revenue levels R_a , R_b , and R_k , and the corresponding productivity levels θ_a , θ_b , and θ_k . Since $\alpha > \eta$, all of these productivity cutoff levels are increasing in the index of industry consumption.

We can now calculate the expected operating profits facing a principal prior to entry, i.e., before she learns her productivity level. Let $\Pi_{ij}(R)$ denote the operating profits for a firm with revenue R that operates with mode of organization i ($i = o$ or

²¹The discussion in this section does not require specific functional forms for the utility function. We use these forms for convenience only. An exponential function $y(j)^\alpha$ is, however, needed in the next section to derive closed-form solutions for the market shares.

v) and locates parts production in country j ($j = S$ or N). If a principal draws a productivity level below $\theta_a = R_a^{1/\alpha} y^{(\alpha-\eta)/\alpha}$, she will forego the opportunity to engage a supplier or hire a manager and earn zero operating profits. If her productivity level falls between θ_a and θ_b , she will choose to buy components from a supplier in the South (as we know from our earlier discussion) and earn operating profits of $\Pi_{oS} (y^{\eta-\alpha} \theta^\alpha)$. If productivity falls between θ_b and θ_v , FDI will be the chosen mode of organization, with expected profits of $\Pi_{vS} (y^{\eta-\alpha} \theta^\alpha)$. In-house production in the North is indicated for productivity levels between θ_v and θ_k , with resulting expected profits of $\Pi_{vN} (y^{\eta-\alpha} \theta^\alpha)$. Finally, for high levels of productivity above θ_k , the principal opts for outsourcing in the South and earns expected profits of $\Pi_{oS} (y^{\eta-\alpha} \theta^\alpha)$. In equilibrium, the expected operating profits for a principal prior to entry match the entry cost, f . We write the free-entry condition as

$$\begin{aligned} & \int_{R_a^{1/\alpha} y^{(\alpha-\eta)/\alpha}}^{R_b^{1/\alpha} y^{(\alpha-\eta)/\alpha}} \Pi_{oS} (y^{\eta-\alpha} \theta^\alpha) dG(\theta) + \int_{R_b^{1/\alpha} y^{(\alpha-\eta)/\alpha}}^{R_v^{1/\alpha} y^{(\alpha-\eta)/\alpha}} \Pi_{vS} (y^{\eta-\alpha} \theta^\alpha) dG(\theta) \\ & + \int_{R_v^{1/\alpha} y^{(\alpha-\eta)/\alpha}}^{R_k^{1/\alpha} y^{(\alpha-\eta)/\alpha}} \Pi_{vN} (y^{\eta-\alpha} \theta^\alpha) dG(\theta) + \int_{R_k^{1/\alpha} y^{(\alpha-\eta)/\alpha}}^{\infty} \Pi_{oS} (y^{\eta-\alpha} \theta^\alpha) dG(\theta) = f . \end{aligned} \quad (8)$$

Expected operating profits for a potential entrant are a strictly decreasing function of y , which means that there is a unique index of industry consumption that delivers expected profits equal to the entry cost.

We are interested in how falling trade costs (and other changes in the production environment) affect the relative prevalence of the different modes of organization. For this, we need to define measures of relative prevalence. We could measure this in terms of the numbers of components produced by different sorts of entities, by the output of final goods that embody components produced in different entities, or by the revenues collected by firms of the different types. Fortunately, all of these measures yield similar answers to the questions of interest, so we can focus on just one. We shall measure relative prevalence in the industry by the shares of components that are manufactured by arms-length suppliers, by foreign subsidiaries, and by in-house parts divisions located in the North.

We let X denote the total output of components manufactured by producers of all types and let n denote the number of principals that enter the industry. A fraction

$G(\theta_b) - G(\theta_a)$ of the entrants draw productivity levels between θ_a and θ_b , which means that they engage in outsourcing. Of these, a fraction h_1 is successful in acquiring components, because principals with productivity levels in the indicated range induce their outsourcing partners to exert effort of e_1 . It follows by the law of large numbers that $n[G(\theta_b) - G(\theta_a)]h_1$ units of components are produced by the supplier firms of principals with productivities in this range. Southern subsidiaries of multinational corporations produce a total of $n[G(\theta_v) - G(\theta_b)][\delta_S + (1 - \delta_S)h_1]$ units of components, because a fraction $G(\theta_v) - G(\theta_b)$ of principals draw productivity levels that make hiring a Southern manager the optimal strategy, and these managers are induced to exert maximal effort on tasks that can be monitored and intermediate effort of e_1 on tasks that are not observable. By similar reasoning, $n[G(\theta_k) - G(\theta_v)][\delta_N + (1 - \delta_N)h_1]$ is the number of units of components produced in-house by firms with a parts division located in the North. Finally, a fraction $1 - G(\theta_k)$ of entrants draw productivity levels above θ_k . These firms engage in outsourcing and all succeed in acquiring the needed components by inducing their partners to exert maximal effort. The resulting number of components is $n[1 - G(\theta_k)]$. The total output of components is the sum of these numbers, or

$$\begin{aligned} X = & n[G(\theta_b) - G(\theta_a)]h_1 + n[G(\theta_v) - G(\theta_b)][\delta_S + (1 - \delta_S)h_1] \\ & + n[G(\theta_k) - G(\theta_v)][\delta_N + (1 - \delta_N)h_1] + n[1 - G(\theta_k)]. \end{aligned} \quad (9)$$

We can now readily compute the shares of components produced under the different modes of organization. Let σ_o represent the share of components produced by outsourcing partners, σ_{vS} represent the share produced by firms that engage in FDI, and σ_{vN} represent the share produced in-house in the North. Then

$$\sigma_o = \frac{n}{X} \{ [G(\theta_b) - G(\theta_a)]h_1 + [1 - G(\theta_k)] \} , \quad (10)$$

$$\sigma_{vS} = \frac{n}{X} \{ [G(\theta_v) - G(\theta_b)][\delta_S + (1 - \delta_S)h_1] \} , \quad (11)$$

and

$$\sigma_{vN} = \frac{n}{X} \{ [G(\theta_k) - G(\theta_v)][\delta_N + (1 - \delta_N)h_1] \} . \quad (12)$$

Notice that the market shares do not depend on the number of principals that enter,

because n/X is independent of n .²²

5 Improved Monitoring in the South

In this section, we show how the model can be used to investigate the effects of changes in the production environment on the relative prevalence of different modes of organization. We consider an increase in δ_S , which is the fraction of tasks undertaken by a division manager in the South that can be monitored by the principal in the North. Such a gain in monitoring possibilities may result from improvements in communications technology or perhaps from changes in the legal system.

In order to derive specific results, we need to make additional assumptions about the distribution of productivity levels among potential entrants. There is evidence to suggest that this distribution is well approximated by a Pareto distribution, so rather than derive sufficient conditions for $G(\cdot)$ under which certain results may hold, we choose to work with this simple functional form.²³ We henceforth assume that $G(\theta) = 1 - \theta^{-\beta}$, for $\theta \geq 1$ and $\beta > 1$. With this distribution, the minimum productivity level for a potential entrant is one and there is no limit on the maximum productivity level.

As we have noted previously, an increase in δ_S rotates the dotted line in Figure 4 in a counterclockwise direction around point Q . It therefore has no effect on R_b , the level of potential revenue at which outsourcing and FDI yield the same expected profits. It does, however, cause R_v to rise, which means that the principal prefers FDI to in-house production in the North for a wider range of potential revenues. As δ_S increases, there is no change in R_a or R_k , as the ability to monitor managers in the South has no bearing on the profitability of outsourcing in the South or on the relative profitability of outsourcing versus in-house production in the North.

For given y , an improvement in a principal's ability to monitor managers in the South must increase the expected operating profits of potential entrants, the term on

²²The equilibrium number of entrants n can be solved using the definition of y , the equilibrium output of differentiated products of the various types, and the free-entry condition (8), which determines a unique equilibrium level of y .

²³See, for example, Axtell (2001), who provides evidence that the Pareto distribution fits well the distribution of sales by firm in the United States. Helpman, Melitz and Yeaple (2002) have shown how a Pareto distribution of firm sizes will emerge from a Pareto distribution of productivity levels. They also verify that Axtell's aggregate results hold for all but a handful of the 52 industries in their data set.

the left-hand side of (8). This means that additional principals will enter the industry (n rises), causing y to grow as well. However, the induced growth in the number of entrants and the consumption index have no direct effect on the market shares.

The growth in y will also cause an upward shift in all of the cutoff productivity levels, i.e., in θ_a , θ_b , θ_v , and θ_k . This is because, with greater industry competition as measured by the index of industry output, a given firm must have higher productivity itself to achieve the same level of revenues as before. However, all of the cutoff productivity levels appear similarly in the numerators and denominators of the expressions for market share; each such term is raised to the power of $-\beta$. It follows that the aggregate output of a component type (the numerators in the market share expressions) and the aggregate output of all components (the denominators in the market share expressions) grow by the same proportion as y increases and thus the aggregate consumption index has no indirect effect on the market shares as well.

We are left with two effects of an increase in the fraction of tasks that can be monitored in the South. An increase in δ_S generates an increase in the probability that any given parts division in the South will be successful in producing components, because managers of Southern subsidiaries devote more effort to tasks that are monitored than to those that are not. And an increase in δ_S causes a wider range of principals to locate their parts divisions in the South, as R_v rises. The former effect tends to increase X/n , the average number of components produced per entrant, while the latter effect tends to reduce it (since parts divisions in the South succeed less often than those in the North). On net, however, the former effect dominates with a Pareto distribution of productivity levels, and so the average output of components per entrant rises.²⁴

Now we are ready to discuss the shifts in the market shares. Consider first the market share of components produced by arms-length suppliers, σ_o . There is no effect of a change in δ_S on the cutoff productivity levels at which principals choose to outsource their components beyond what results from the change in y , which we know does not affect the market shares. So the term in curly brackets in (10) does not change for given y . But we have just noted that X/n rises for given y , due to the dominant effect of the improved productivity of Southern parts divisions. It follows that the relative prevalence of outsourcing falls.

²⁴To justify this claim, we totally differentiate $I/[ny^{\beta(\eta-\alpha)/\alpha}]$ with respect to δ_S , taking account of the fact that $R_v = R_b + (\bar{s}_N + c_N - \bar{s}_S - c_S) / (\delta_N - \delta_S)(1 - h_1)$. The sign of this derivative is the same as the sign of $q^{-\beta/\alpha} - 1 - (1 - q)\beta/\alpha$, where $q = R_b/R_v < 1$. Since $q^{-\beta/\alpha}$ is convex in q , this derivative must be positive.

The fraction of components emanating from integrated firms with a parts division in the North falls by an even greater percentage than σ_o . From (12) we see that σ_{vN} falls, because the term in curly brackets falls for given y as R_v rises and also because X/n grows. The growing productivity of subsidiaries in the South means that a subsidiary with given productivity will generate a higher average output of components and it also spells a shift away from in-house production in the North in favor of FDI. For both reasons, the market share of Northern parts divisions falls.

Finally, the three market shares sum to one. Since σ_o and σ_{vN} both fall, it must be that σ_{vS} rises, i.e., the output of components by Southern subsidiaries grows by more than the aggregate output of components. Thus, an increased ability to monitor managers in the South makes in-house production in the South a more attractive option relative to both of the alternative modes of organization.

Our results in this section are broadly consistent with some recent empirical evidence on the determinants of the form of foreign investment and the extent of vertical specialization in multinational firms. For example, Lin and Png (2002) examine the form of FDI undertaken by 148 Taiwanese firms that made investments in China between 1987 and 1991. They consider the firms' decision whether to structure the FDI as a joint venture or as a wholly-owned subsidiary, taking distance from Hong Kong as a proxy for the principal's ability to monitor the local agent. A joint venture is more like outsourcing in our model, inasmuch as the local entrepreneur has a greater stake in the project than does the local manager of a wholly-owned subsidiary. Lin and Png find that joint ventures are more likely to be chosen as distance from Hong Kong increases, which is in keeping with our finding that the range of (low-productivity) firms that choose outsourcing over FDI expands as δ_S declines. In related work, Hanson et al. (2002) examine the determinants of the extent of foreign affiliate processing of inputs imported from a parent U.S. firm. After controlling for trade barriers and transportation costs between the parent and subsidiary, they find that the extent of such vertical specialization is decreasing in the distance between parent and subsidiary, is greater for subsidiaries located in English speaking countries than those that are not, and is higher in Mexico and Canada than would be predicted based on distance (and the other variables) alone. All of these findings can be viewed as consistent with our prediction that an increase in δ_S (proxied by distance, common language, and adjacency) raises the relative profitability of FDI compared to other modes of organization.²⁵

²⁵Hanson et al. (2002) also find that the cost share of intermediate inputs imported by a foreign

6 Falling Trade Costs

In this section, we study how falling trade costs affect the international organization of production. Both trade liberalization and declining transportation costs have contributed in recent years to the globalization of economic activity. We are interested in whether and under what circumstances a decline in trade costs will favor one mode of organization over another.

To examine this issue, we first must extend our model to include trade costs. We assume that transporting a component from South to North entails a per unit cost of τ . This cost may reflect a shipping charge, or it may result from an import tariff imposed by the government of the North. The same cost applies whether the component is traded within the firm (as when the part is produced in a foreign subsidiary) or at arms-length (as when a firm arranges for delivery of parts from a Southern supplier). In case of foreign outsourcing, the contract term p_o now refers to the amount that is paid to the foreign entrepreneur in the event that he delivers suitable components to the principal's assembly plant in the North.²⁶

The presence of trade costs modifies the relationship between expected operating profits and potential revenues. Consider first a principal with low productivity for whom foreign outsourcing will be the preferred mode of organization if she chooses to operate at all.²⁷ Such a principal designs a contract under which the expected utility of the foreign entrepreneur is just equal to \bar{s}_S , the utility he could achieve by pursuing his outside option. This means that the incidence of the trade costs falls on the principal. Since we have normalized the number of components she can process to equal one, τ is the total trade cost she will bear in the event that the foreign entrepreneur is able to deliver the components. Thus, expected operating profits from outsourcing are $\Pi_{oS}(R) = (R - \tau)h_1 - e_1 - (\bar{s}_S + c_S)$ for $R \leq R_b$. The minimum potential revenues necessary for such a principal to enter into any contract with a supplier are $R_a = \tau + (e_1 + \bar{s}_S + c_S)/h_1$.

At R_b , the principal is indifferent between outsourcing and engaging a manager to

affiliate from its U.S. parent for further processing is decreasing in the wage of unskilled labor in the host country. This too is consistent with the comparative static properties of our model, i.e., that the market share of FDI falls with c_S .

²⁶Alternatively, the principal might choose to make the account payable upon delivery of components in the foreign country. It can be shown, however, that at all productivity levels, a principal prefers (at least weakly) to set a c.i.f. price rather than a f.o.b. price.

²⁷We assume that the trade costs are not so large as to make domestic outsourcing a more attractive option than foreign outsourcing.

head a subsidiary in the South. In the event of FDI, the principal bears any trade costs that arise. Expected operating profits in this mode of organization and with the optimal employment contract for the manager are $\Pi_{vS}(R) = (R - \tau)[\delta_S + (1 - \delta_S)h_1] - [\delta_S + (1 - \delta_S)]e_1 - (\bar{s}_S + c_S)$, for R between R_b and R_v . From $\Pi_{oS}(R_b) = \Pi_{vS}(R_b)$ we find that $R_b = \tau + (1 - e_1)/(1 - h_1)$ in the presence of trade costs.

Costs that arise from importing components do not affect the potential profitability of in-house production in the North for a given R . Thus $\Pi_{vN}(R) = R[\delta_N + (1 - \delta_N)h_1] - [\delta_N + (1 - \delta_N)]e_1 - (\bar{s}_N + c_N)$ for R between R_v and R_k , as before. We calculate the new value of potential revenues that makes the principal indifferent between in-house production in the South and in the North, and find

$$R_v = \frac{1 - e_1}{1 - h_1} + \frac{(\bar{s}_N + c_N) - (\bar{s}_S + c_S) - \tau[\delta_S + (1 - \delta_S)h_1]}{(\delta_N - \delta_S)(1 - h_1)}.$$

Finally, when potential revenues are sufficiently great, a principal will find it optimal to outsource the production of components. Moreover, she will opt to provide sufficient incentives for the foreign entrepreneur to exert maximal effort, thereby creating rents for the entrepreneur. The trade costs, like the cost of the inputs used to produce the components, are borne initially by the foreign entrepreneur, since the contingent payment p_o is made only when the goods are delivered to the principal in the North. When the entrepreneur captures rents, the incidence of the trade costs falls on him. This means that the principal achieves the same profits from outsourcing as when trade costs are zero; namely $\Pi_{oS}(R) = R - (1 - e_1)/(1 - h_1)$ when $R \geq R_k$. The potential revenues R_k at which in-house production in the North and foreign outsourcing yield equal expected profits also is the same as before.

We consider now the effects of a fall in trade costs τ . In Figure 6, the dotted line indicates that when τ declines, operating profits rise for firms that outsource in the South and for those that undertake FDI. The point at which principals elect to engage a supplier rather than exit the market shifts from R_a to R'_a , because principals are able to make positive operating profits for a wider range of revenue levels when trade costs are lower. The level of potential revenues at which principals are indifferent between FDI and outsourcing also falls (from R_b to R'_b). This reflects the fact that a fall in trade costs boosts $\Pi_{vS}(\cdot)$ by more than it does $\Pi_{oS}(\cdot)$ at a given level of potential revenues, inasmuch as a multinational achieves a higher probability of successfully producing components than does an arms-length supplier; thus, the expected cost savings is

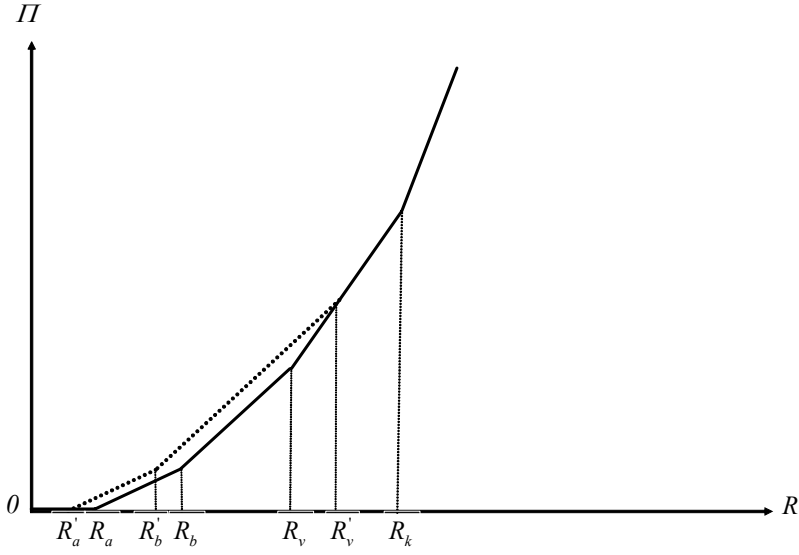


Figure 6: Fall in Trade Costs

greater for the former than the latter.²⁸ The fall in trade costs makes the South a relatively more attractive location for in-house production of components, so R_v rises. Note again that R_k is unaffected.

It follows readily from this discussion that the market share of components produced by vertically integrated firms in the North falls.²⁹ It is hardly surprising that a reduction in trade costs increases the market share of imported components. But it remains to be seen whether this reflects an expansion in the share of components purchased from supplier firms, the share produced by foreign subsidiaries of multinational

²⁸The probability of success for a multinational firm operating in the South is $\delta_S + (1 - \delta_S)h_1$, which exceeds h_1 , the probability of success for an independent supplier.

²⁹Using (9), (12), $G(\theta) = 1 - \theta^{-\beta}$ and $\theta_i = R_i^{1/\alpha} y^{(\alpha-\eta)/\alpha}$ for $i = a, b, v$, and k , we can write the market share of components produced in-house in the North as

$$\sigma_{vN} = \frac{\left(R_k^{-\beta/\alpha} - R_v^{-\beta/\alpha} \right) [\delta_N + (1 - \delta_N)h_1]}{\Omega}$$

where

$$\begin{aligned} \Omega = & \left(R_a^{-\beta/\alpha} - R_b^{-\beta/\alpha} \right) h_1 + \left(R_b^{-\beta/\alpha} - R_v^{-\beta/\alpha} \right) [\delta_S + (1 - \delta_S)h_1] \\ & + \left(R_k^{-\beta/\alpha} - R_v^{-\beta/\alpha} \right) [\delta_N + (1 - \delta_N)h_1] + R_k^{-\beta/\alpha} . \end{aligned}$$

Then, the fact that $dR_a/d\tau = dR_b/d\tau > 0$, $dR_v/d\tau < 0$, and $dR_k/d\tau = 0$ implies $d\sigma_{vN}/d\tau > 0$.

corporations, or both. We will now show that the answer to this question depends sensitively on the characteristics of the firms that engage in outsourcing, and in particular on whether these are predominantly low-productivity firms that outsource so as to generate higher-powered incentives for their partners or predominantly high-productivity firms that outsource to conserve on the rents that they pay to their partners.

To make this point, we consider two extreme cases. Suppose first that $h_1 = (e_1 + \bar{s}_S + c_S)/(1 + \bar{s}_S + c_S)$, so that $R_a = R_b$. Then, there are no firms with low productivity that outsource the production of components; all outsourcing takes place among the firms with the very highest levels of productivity.³⁰ The fall in import costs confers no direct benefit to these firms, because such trade costs are borne in any case by the foreign entrepreneurs. It follows that there is no direct spur to outsourcing in this case, and thus no change in the term in curly brackets in (10). But $X/ny^{\beta(\eta-\alpha)/\alpha}$ grows as τ shrinks, because in-house production of components in the South expands by more than in-house production in the North contracts.³¹ Therefore, the relative prevalence of outsourcing falls. Even the absolute volume of outsourcing might fall under these circumstances, because the firms that outsource garner no benefit from the fall in trade costs but face greater competition from multinationals and from new entrants to the industry.

In this case where σ_{vN} and σ_o both decline, the market share of components produced by foreign subsidiaries of multinational corporations must rise. These firms benefit directly from the fall in trade costs, which expands the range of potential revenue levels for which FDI takes place at both the upper and lower ends. The rise in output (given n and y) outpaces the rise in total output of components, because output by integrated firms in the North contracts.

Now suppose that δ_S does not differ much from δ_N and that both are very close to one. We also assume for the purposes of this example that $\bar{s}_N + c_N > \bar{s}_S + c_S + \tau[\delta_S + (1 - \delta_S)h_1]$. In this case, both R_v and R_k are very large, which means that virtually all outsourcing takes place among firms with relatively low levels of productivity. Moreover, in this situation, virtually all components are produced in the South.³²

³⁰The relative weight of low productivity firms also declines when the distribution $G(\theta)$ shifts to the right. For example, the Pareto distribution can be extended to $1 - (\gamma/\theta)^\beta$, where $\gamma > 0$. In this case the support is $\theta \geq \gamma$. An upward shift in γ shifts the distribution to the right.

³¹We find that $d[I/ny^{\beta(\eta-\alpha)/\alpha}]/d\tau = (\beta/\alpha)(R_v^{-\beta/\alpha-1} - R_b^{-\beta/\alpha-1}) < 0$.

³²Now the weight of the high productivity firms falls when the distribution $G(\theta)$ shifts to the left. With $G(\theta) = 1 - \theta^{-\beta}$, this happens when β rises. A higher β reduces the weight on the tail of the distribution.

A little algebra suffices to show that as trade costs fall under these different conditions, the relative prevalence of outsourcing must increase.³³ But since firms that outsource and those that engage in FDI share virtually the entire market when R_v and R_k are large, the fraction of components produced by multinationals must fall. As we have noted, a fall in trade costs causes R_b to fall, which means that the volume of components produced by foreign subsidiaries rises. But the growth in FDI is more than matched in percentage terms by the growth in outsourcing (associated with the decline in R_a), which occurs because some principals with low productivity now find it profitable to engage a supplier rather than to exit the industry. And the number of low productivity principals is larger than the number of high productivity principals when $G(\theta)$ is a Pareto distribution.

We conclude that the effects of falling trade costs on the mode of organization and the nature of international trade will vary across industries. Trade liberalization tends to favor multinational activity in industries in which outsourcing is conducted predominantly by high productivity firms that are seeking to minimize the rents they must share with the head of a parts division in order to induce high levels of effort. In contrast, trade liberalization spurs arms-length trade with suppliers when most outsourcing is undertaken by relatively low productivity firms that are seeking to boost the power of the incentives they can provide to their component producers. In the former case, the fall in trade costs does little to increase the profitability of outsourcing, but firms that engage in FDI realize an immediate cost savings. In the latter case, the principal benefits under either mode of organization involving trade, but outsourcing receives a greater boost because it can expand at the extensive margin.

7 Conclusions

We have developed a model in which the heterogeneous firms in an industry choose their modes of organization and the location of their subsidiaries or suppliers. We assume that the principals of a firm are institutionally or legally constrained in the nature of the contracts they can write with suppliers and employees. In particular, a supplier cannot be asked to post a (large) bond that will be forfeited in the event that the firm's efforts to produce suitable components prove unsuccessful. Similarly, a manager

³³By differentiating the definition of σ_o , we find that the sign of $d\sigma_o/d\tau$ is the same as the sign of $R_b(R_b^{-\beta/\alpha-1} - R_a^{-\beta/\alpha-1}) + (R_a^{-\beta/\alpha} - R_b^{-\beta/\alpha}) = R_a^{-\beta/\alpha}(1 - R_b/R_a)$, which is negative, because $R_b > R_a$.

cannot be asked to pay a fine if his division performs poorly, nor can he be asked to pay personally for the inputs used by his division. In such an institutional setting, the contracts that principals can write with their suppliers and division managers may not induce efficient levels of effort on all relevant tasks.

We identified two reasons why a principal may benefit from engaging an external supplier to manufacture components in a setting like this. First, the principal can confront an agent with higher-powered incentives when the agent has more at stake. A supplier can be made to front the cost of the inputs needed to manufacture components and so can be given a greater stake in the project than a manager. Second, when a principal finds it desirable to induce a very high level of effort from her agent, the cost to the principal of providing the necessary incentives is less for an outside supplier than for an employee. Again, the input costs play a key role in this. In the main text, we showed that principals who are capacity constrained must leave rents to their agents when they induce the highest level of effort. But the rents are smaller for an entrepreneur than for a manager, because the principal can pass along input costs to the former but not the latter. In the appendix, we relaxed the assumption of a fixed firm size and found a similar mechanism at work. With variable scale, the rents that might potentially go to an entrepreneur can be fully taxed away by a principal in a supplier relationship. The principal who wants high effort offers a large bonus, but also demands a large delivery of components, thereby minimizing the net cost of the incentive contract.

Against the benefits of outsourcing, there is an advantage to in-house production that stems from the greater opportunity it affords the principal to monitor the actions of her agent. We assume that the ability to monitor an agent declines with distance; a vertically integrated firm is able to observe a division manager's actions on more tasks when the division is located near the headquarters than when it is located in a different country. Thus, FDI suffers the disadvantage of lesser monitoring compared to in-house production near the headquarters, but the possible advantage of lower costs.

Our main result concerns the sorting of firms in an industry into different organizational forms. The least productive firms that are active in equilibrium choose to subcontract the production of components to suppliers in the South. For these firms, the ability to offer higher-powered incentives with outsourcing weighs most heavily. Firms with intermediate levels of productivity opt for vertical integration, with the less productive of these undertaking foreign investment in the South and the more

productive operating a parts division in the North. FDI does not appeal as much to the more productive firms, because the ability to monitor a manager's efforts becomes more valuable as potential revenues rise. Finally, outsourcing is the preferred option for the most productive firms in an industry, because the principals of these firms who want to induce a high level of effort are able to pass along input costs to a supplier but not to a manager.

We used our model to examine the implications of falling trade costs for the relative prevalence of the different organizational modes. An important observation is that trade liberalization may promote mostly FDI or mostly outsourcing, depending on the characteristics of an industry. In particular, the market share of imports from suppliers will expand as trade costs fall if most of the outsourcing is undertaken by low productivity firms in which the principals are motivated by a desire to give their agents a greater stake in the venture. But the market share of imports from suppliers will contract as trade costs fall if most outsourcing is undertaken by high productivity firms in which the principals are motivated by a desire to minimize the rents captured by their agents. The equilibrium sorting of firms by productivity level plays an important part in delivering these conclusions.

References

- [1] Axtell, Robert L. (2001), “Zipf Distribution of U.S. Firm Sizes,” *Science*, 293, 1818-1820.
- [2] Borga, Maria and Zeile, William J. (2002), “Foreign Outsourcing and the Intrafirm Trade of U.S. Multinational Corporations,” U.S. Bureau of Economic Analysis, manuscript.
- [3] Campa, José and Goldberg, Linda (1997), “The Evolving External Orientation of Manufacturing: A Profile of Four Countries,” *Federal Reserve Bank of New York Economic Policy Review*, 4, 53-81.
- [4] Feenstra, Robert C. (1998), “Integration of Trade and Disintegration of Production in the Global Economy,” *Journal of Economic Perspectives*, 12, 31-50.
- [5] Feenstra, Robert C. and Hanson, Gordon H. (2002), “Global Production Sharing and Rising Inequality: A Survey of Trade and Wages,” in J. Harrigan (ed.), *Handbook of International Trade* (London: Basil Blackwell Publishers), forthcoming.
- [6] Grossman Gene M. and Helpman, Elhanan (2002), “Outsourcing in the Global Economy,” NBER Working Paper No. 8728.
- [7] Hanson, Gordon H., Mataloni, Raymond J. and Slaughter, Matthew J. (2001), “Expansion Strategies of U.S. Multinational Corporations,” *Brookings Trade Forum 2001*, 245-294.
- [8] Hanson, Gordon H., Mataloni, Raymond J. and Slaughter, Matthew J. (2002), “Vertical Specialization in Multinational Firms,” University of California, San Diego, manuscript.
- [9] Helpman, Elhanan, Melitz, Marc and Yeaple, Stephen (2002). “Exports vs. FDI,” manuscript, Harvard University.
- [10] Holmström, Bengt and Milgrom, Paul (1991), “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics and Organization* 7, 24-52.

- [11] Horn, Henrik, Lang, Harald, and Lundgren, Stefan (1995). “Managerial Effort, Incentives, X-inefficiency and International Trade”, *European Economic Review*, 39, 117-138.
- [12] Hummels, David, Rapoport, Dana, and Yi, Kei-Mu (1998), “Vertical Specialization and the Changing Nature of World Trade,” *Federal Reserve Bank of New York Economic Policy Review*, 79-99.
- [13] Hummels, David, Ishii, Jun, and Yi, Kei-Mu (2001), “The Nature and Growth of Vertical Specialization in World Trade,” *Journal of International Economics*, 54, 75-96.
- [14] Katz, Lawrence F. (1986), “Efficiency Wage Theories: A Partial Evaluation,” *NBER Macroeconomics Annual*, 1, 235-275.
- [15] Lin, Chu-Chia S. and Png, Ivan (2002), “Monitoring Costs and the Mode of International Investment,” National University of Singapore, manuscript.
- [16] Melitz, Marc (2002) “The Impact of Trade on Intra-Industry Reallocations on Aggregate Industry Productivity,” NBER Working Paper No. 8881.
- [17] Yeats, Alexander J. (2001), “Just How Big is Global Production Sharing?” in S.W. Arndt and H. Kierzkowski, eds., *Fragmentation: New Production Patterns in the World Economy*, Oxford: Oxford University Press.

8 Appendix: Variable Scale of Production

In the main text, we assumed that the principal can only operate a firm of a given size. In this appendix, we allow the principal to choose her scale of production. We take the output of components in a successful venture to be proportional to the inputs of labor and materials, and normalize the unit variable cost of a component to equal one. The production of components may also entail a fixed cost, which we denote by k .

Once again, principals are distinguished by their potential productivity. A principal with productivity θ can produce $\theta\ell$ units of output from ℓ components and collect revenues of $R(\theta\ell)$. A principal who attempts to produce components in house will choose ℓ_o directly, and will bear the fixed and variable costs of this operation. The probability that such a venture succeeds depends on the efforts of a manager in a continuum of tasks. If the principal chooses instead to contract with an independent supplier, she will specify the size of her order, ℓ_o . We assume that she can verify whether the entrepreneur purchases the inputs that are needed to make the production of ℓ_o units of the component possible. But the principal cannot observe the efforts exerted by the entrepreneur on any of the tasks that determine the probability that usable components will result from the project.

As in Section 2, we neglect the locational component of the principal's decision to focus instead on the choice of organizational mode. Consider first a principal with productivity θ who opts to outsource the production of components. Such a principal chooses a (non-negative) guaranteed payment s , an order size ℓ_o , and a payment for delivery p_o to maximize expected profits. Recognizing again that the principal can induce effort e_o with a payment $p_o = \phi(e_o)$, the principal's problem is to maximize

$$\Pi_o = h(e_o)R(\theta\ell_o) - s - h(e_o)\phi(e_o) \tag{A1}$$

subject to $s \geq 0$ and

$$s + h(e_o)\phi(e_o) - \ell_o - k - e_o \geq \bar{s}. \tag{A2}$$

We assume (in place of Assumption 2) that $h(E)/h'(E) - E > \bar{s} + k$; i.e., if the principal elects to induce the highest level of effort E with the incentive payment $\phi(E) = 1/h'(E)$, the entrepreneur would be willing to purchase some positive quantity of the inputs that are necessary for the production of components.

Observe first that the participation constraint (A2) always binds in this problem. Were it not to bind, the principal could expand her order of components (i.e., increase

ℓ_o) and raise expected profits at no cost to herself. Note too that if the constraint $s \geq 0$ does not bind, the principal can choose the effort level e_o and the order size ℓ_o that maximizes joint welfare for herself and the entrepreneur. But since the entrepreneur would achieve only his reservation level of utility \bar{s} , this outcome would be first-best for the principal. We denote by $e^* = \varsigma_e(\theta)$ the entrepreneur's effort level in the principal's first-best, and by $\ell^* = \varsigma_\ell(\theta)$ the corresponding order size. For an exponential revenue function with exponent between zero and one — such as arises for the CES utility function described in Section 4 — both e^* and ℓ^* are increasing functions of θ as long as $e^* < E$. Moreover, there exists a finite productivity level θ_E such that $\varsigma_e(\theta) < E$ and $\varsigma_\ell(\theta) < \ell_E \equiv \varsigma_\ell(\theta_E)$ for all $\theta < \theta_E$ and $\varsigma_e(\theta) = E$ and $\varsigma_\ell(\theta) = \ell_E$ for all $\theta \geq \theta_E$.

Next consider a principal with productivity θ who chooses to manufacture components in house. Such a principal will choose a wage w , an effort level e_m on observable tasks, an effort level e_n (and associated bonus payment) on non-observable tasks, and a scale of operation ℓ_v to maximize

$$\Pi_v = [\delta h(e_m) + (1 - \delta)h(e_n)] [R(\theta \ell_v) - \phi(e_n)] - w - \ell_v - k$$

subject to $w \geq 0$ and

$$w + [\delta h(e_m) + (1 - \delta)h(e_n)] \phi(e_n) - \delta e_m - (1 - \delta)e_n \geq \bar{s}.$$

We compare now the maximum expected profits under the alternative organizational forms for different values of θ . For the moment, we disregard the possibility that for certain very low levels of productivity, the principal may opt to exit the industry rather than to engage either a manager or a supplier.

For low values of θ , the constraint $s \geq 0$ does not bind in the problem to maximize Π_o , nor does the constraint $w \geq 0$ bind in the problem to maximize Π_v . For such levels of productivity, the principal achieves her first-best with either organizational form. In doing so, the guaranteed payment for outsourcing is $s^*(\theta)$ and the base wage of a manager is $w^*(\theta)$. We note that $s^*(\theta) > w^*(\theta)$, because the former includes compensation for the input costs k and ℓ_o , whereas the latter does not.³⁴

With an exponential revenue function, both $s^*(\theta)$ and $w^*(\theta)$ decline as θ rises. But

³⁴More specifically,

$$s^*(\theta) = \bar{s} + k + \varsigma_\ell(\theta) + \varsigma_e(\theta) - h[\varsigma_e(\theta)]\phi[\varsigma_e(\theta)]$$

and

$$w^*(\theta) = \bar{s} + \varsigma_e(\theta) - h[\varsigma_e(\theta)]\phi[\varsigma_e(\theta)].$$

w^* declines faster than s^* , because $s^*(\theta) = w^*(\theta) + k + \ell_o(\theta)$, and ℓ_o increases with θ . It follows that the non-negativity constraint on w will bind for some low levels of productivity at which $s^*(\theta) > 0$. For these values of θ , the principal can achieve her first-best level of operating profits by contracting with an outside supplier but not by hiring a manager to produce components in house. Outsourcing is the preferred mode of production for productivity levels in this range.³⁵

Next we show that a principal with very high productivity also prefers to outsource. For high levels of productivity, the constraint that $s \geq 0$ may or may not bind. If it does not bind, then outsourcing gives the principal her first-best level of expected profits while in-house production does not. Clearly, outsourcing is preferred in this case. Suppose, instead, that the principal cannot achieve her first-best outcome by outsourcing for values for θ above some θ_s . As θ grows large, the optimal contract for an in-house manager induces effort of E on all tasks. This leaves the manager with positive rents under our assumption that $h(E)/h'(E) - E > \bar{s} + k \geq \bar{s}$. With $e_m = e_n = E$, the principal chooses a scale of production that satisfies $h(E)\theta R'(\theta\ell_v) = 1$. If the principal were instead to offer an external supplier a contract with $e_o = E$, $\ell_o = \ell_v(\theta)$, and $s = \ell_v(\theta)$, the entrepreneur would willingly accept it.³⁶ But such a contract, which is not necessarily the optimal offer to an external supplier, yields higher expected profits than in-house production by the amount k .

Finally, we show that if δ is close enough to one and if there exists a θ_s such that the principal cannot achieve her first-best outcome by outsourcing for $\theta \geq \theta_s$, then there also exists an intermediate range of productivity levels at which a principal prefers in-house production to outsourcing. To see this, suppose that a principal could monitor all of the activities of a manager of an in-house division. Such a principal could achieve the first-best level of expected profits by specifying $e_m = \varsigma_e(\theta)$ and $\ell_v = \varsigma_\ell(\theta)$ and by setting a bonus payment that just leaves the manager indifferent between accepting the contract and not. With perfect monitoring, in-house production clearly would be preferred to outsourcing in situations where the latter organizational form does not achieve the first best. But expected profits for an integrated firm are a continuous

³⁵Note that, as before, $\Pi_o(\theta) = 0$ for the smallest value of θ such that $s^*(\theta) = 0$. Therefore, the principal will not engage a supplier or manager for lower levels of productivity than this.

³⁶Under this contract,

$$s + h(E)\phi(E) - E - \ell_v = h(E)/h'(E) - E,$$

which exceeds $\bar{s} + k$ by assumption.

function of δ . It follows that if δ is close enough to one and θ is such that outsourcing does not achieve the first best, then in-house production will be the preferred mode of organization for the principal.

We conclude that the sorting of firms by productivity level is qualitatively the same with a variable scale of production as when the scale is fixed. The least productive firms opt to exit the industry. Those with low productivity above a certain threshold can make positive expected profits by outsourcing but not by in-house production. Thus, there is a range of low productivity levels for which outsourcing is the preferred mode of organization for the principal. Outsourcing also is preferred by the principal when productivity is sufficiently high. For an intermediate range of productivity levels, the principal will opt for in-house production if δ is sufficiently large.