The Boundary of the Firm in a Model of Trade Within a Hierarchy*

Nadav Levy†

September 2003

Abstract

In this paper I present a theory of the boundary of the firm that accounts for some important characteristics of real-world multidivisional firms: Operative decisions are in the hands of middle managers who are rewarded with incentive contracts based on the performance of their units; Managers' decisions are subject to approval and intervention by the top management of the firm; and managers are better informed regarding the affairs of their divisions than their superiors in the firm's hierarchy. In this setup, the integration of a producer of an intermediate input and its buyer as separate divisions within a single firm is unambiguously desirable, as long as the choice of trading partners can be credibly delegated to the divisions' managers. I show that this is satisfied not only under the assumption of full commitment by the general office of the firm, but also interestingly, if it has no commitment power at all. At the time of trade, the uninformed general office prefers to delegate the choice of trading partners to the divisions whose decision is ex-post optimal. An explanation of the boundaries of the firm emerges only if we assume that the general office retains some limited commitment power. The general office may then mandate internal trade in order to encourage the divisions to specialize towards one another before the trade. In the context examined, I show that the general office faces a 'time-consistency' problem. It tends to mandate internal trades in more instances than would have been optimal with full commitment, adversely affecting the levels of investment taken by the divisions' managers. Whenever such inconsistency arises, it may be optimal to have the trade conducted between independent, non-integrated parties.

*I am grateful to Michael Whinston for guidance and encouragement and to Rob Porter and Asher Wolinsky for helpful discussions and comments. I also benefited from the comments of Nir Jaimovich, Randal Watson and Oded Bizan. Financial support from the Center for the Study of Industrial Organization at Northwestern University is gratefully acknowledged.

†Department of Economics, University at Albany, SUNY *email:* nadav@albany.edu

1 Introduction

Economists have been intrigued with the question of the boundary of the firm ever since Coase's (1937) seminal paper. Following the work of Williamson (1975, 1985), Klein, Crawford and Alchian (1978) and others, the discussion has focused on the difficulties of market transactions when the contracting environment is incomplete. Integrating successive production stages within the confines of a single firm was perceived to alleviate some of these costs, reducing opportunistic behavior and minimizing ex-post holdups. These benefits are to be weighed against the increased governance costs of managing a larger organization.

Drawing on these informal ideas, modern theories of the firm emphasize the critical role that the change in the control of firms' assets due to integration has on the decisions taken by stakeholders within firms. In the prominent Property Rights Theory of the Firm (Grossman and Hart (1986), Hart and Moore (1990)), the joint surplus from trade is divided between the participants by *ex-post* negotiations. By transferring ownership of all physical assets essential for the trade into the hands of the acquiring firm's owner, integration changes the bargaining outcomes. Consequently the parties willingness to make *ex-ante* trade-specific investments is affected in opposite directions. The costs and benefits of integration are thus clearly identified, and an explanation of firms' boundaries emerges, alluding to such factors as the marginal importance of the parties' investments and the degree of complementarity between the assets. Holmstrom and Milgrom (1991), building on Williamson's ideas, present a model of integration in which the main determinant of the boundary are measurement costs. Here again integration matters as it fundamentally changes the incentives of managers who are no longer owners. In this case the incentives are for activities that are hard to measure such as asset maintenance and proper utilization.

While the work described above has been fundamental in shaping our thinking on these issues, these models are more suited to study the implications of the acquisition of an entrepreneurial small firm than to analyze a merger between two large public firms into a multidivisional entity.[1] The reason is that these models do not distinguish between the ownership and management of firms which are taken to be in the hands of the same individual. In most modern enterprises on the contrary, the control of many of the operative decisions is in the hands of managers who act as agents of a diffuse set of owners and hold only a negligible stake of the firm. Managers have only limited bargaining power and their compensation is primarily determined by monetary incentives specified contractually, and is not directly tied to the set of assets owned by the firm. As a result the forces identified in the models

[1] See also Bolton and Scharfstein (1998) for a discussion of related issues.

above do not directly apply in these environments. The fundamental change brought by integration is to place the two trading stages under a unified management structure, where they are controlled and managed through the firm's hierarchy.[2] As described by Chandler (1962), many large firms came to be organized in a multidivisional structure, distinguished by the delegation of most operating decisions to the managers of semi-autonomous divisions whose decisions are supervised and coordinated by a general office. For this reason, the direction of integration, that plays an important role in the Property Rights model, is no longer as relevant, since either forward or backward integration would result in a similar final structure.

In this paper, I use a model of internal trade within an hierarchal structure to develop a theory of the firm's boundary for such environments. A main difference between integration and non-integration in this setting, is in the objective of the common principal, the general office of the firm. Controlling both trading units, the general office may be inclined to change managerial compensation and to intervene in the decisions taken by division managers. Specifically, the general office may limit the divisions' sourcing autonomy and mandate trade within the firm. As was documented by Eccles (1985), firms vary substantially in this respect. In his study of transfer price practices of large multidivisional firms, mandated transactions were quite common, but in other firms, divisions were given broad autonomy to choose external partners, even when viable internal alternatives existed. General offices have even sometimes employed different policies with respect to different inputs traded between divisions in the firm.

Within this framework, I identify the general office's ability to credibly delegate the choice of trading partners to the divisions as central to an explanation of firms' boundaries. Under the incomplete contracts paradigm, maintaining the divisions' option to trade for inputs outside the firm is valuable as it prompts general investments (such as in quality) by the divisions' managers. Integration is always desirable as long as divisional sourcing autonomy can be sustained, yet the general office may be inclined to limit it if the divisions decisions do not comply with the corporation best interest. In section 2, I consider a simple model where in the absence of commitment to divisional sourcing autonomy, and if internal trade is always efficient, the incentives of divisions' managers to invest are severely impaired under integration. Surprisingly however, if external trade may be efficient, even only in extreme circumstances, and provided that the values of different trades are observed only by divisions' managers, the general office finds it optimal to let the divisions choose their

[2] On this, see the discussion in Williamson (2000).

trading partners. This setting does not offer a theory of vertical integration as integration always outperforms nonintegration even if ex-ante commitment is not feasible.

In order to explain the boundary I turn next to consider a richer environment. I employ and build on a framework by Holmstrom and Tirole (1991) who study the use of alternative transfer pricing policies within an integrated firm. In their model, division managers make general investments as but also decide on how much to specialize their divisions' production towards each other. If divisions are allowed to trade externally, investments are high, but the divisions do not specialize sufficiently, each trying to opportunistically increase its share of the gains from trade.

Section 3 presents a general model which extends Holmstrom and Tirole's framework to allow for external trade to be efficient at times, and posits that the ranking of different trading relations is known only to the divisions' managers, not to the general office. These plausible assumptions were demonstrated to be important in the context of the simpler model of section 2. In section 4, I show that if full commitment to a divisional sourcing autonomy is feasible, integration always dominates nonintegration. The advantage of integration, as identified by Holmstrom and Tirole, is in the "coordination" of incentives between the managers. Under common ownership, the effect of an increase in a division's investment on the profits of the entire firm is internalized and incentives are adjusted accordingly. In section 5, I consider the case of no commitment. I analyze a particular model of bargaining between the divisions and demonstrate that the intuition of the simpler model generalizes to this setting, and integration always outperforms nonintegration. I conclude that limited, but less than full commitment power by the general office is necessary to explain firms' boundaries in this framework.

In section 6 , I then study a particular form of limited commitment, assuming that the general office is unable to commit ex-ante (before investments are taken) but only in an interim stage, before the decisions determining the exact specifications of the trade are made. I show that the general office faces a "*time inconsistency*" problem: since in the interim stage it no longer takes into account the effect on investments that are already sunk, it tends to mandate internal trade in too many instances compared to what is optimal ex-ante. The divisions anticipate this equilibrium behavior, and lower their investments as a result. Nonintegration may be an optimal organizational form in situations where the integrated firm suffers from such inconsistency. I also show that the asymmetry of information between the general office and the divisions regarding the value of external trade opportunities implies that any choice of transfer pricing policy within an integrated firm results in some trading inefficiency. In section 7, I consider how various changes in the environment, including the relative attractiveness of external opportunities and the degree of cooperativeness of investments affect the optimal organizational form. In section 8, I investigate whether the findings of the previous section still hold if the general office can use elaborate mechanisms in order to extract the divisions' information on the value of different trades. I show that the essence of the results is retained and that the possibility of using such mechanisms can adversely affect the incentives of the divisions to invest, and is therefore potentially harmful. Section 9 concludes.

2 Prologue

How does trade between two units within the hierarchial structure of an integrated firm differ from that between two independent firms? I argue that within the confines of an integrated firm, the units can no longer threaten to opt out of an efficient internal trading relationship, as the general office of the firm may not allow it. This limits the role that information from external markets can play in monitoring the performance of the units, and has an adverse effect on the investments made by them, lowering the overall efficiency of the integrated organization. A safeguard against such intervention is the fact that internal trade may not always be efficient, and the units are typically better informed than the general office with respect to the values of different trades. This tends to discourage intervention and restore incentives to invest.

In this section, I demonstrate these ideas using a simple model in the spirit of Hart (1995). The purpose here is mainly expository, and the analysis is therefore not intended to be rigorous. Consider a trading relation between two units, a buying unit B and a selling unit S. The units can trade a single unit of input that is to be used by B. Each unit is headed by an employee-manager. Prior to the trading period, each manager can take an investment that would increase the value of trade. Let $v(b)$ denote the value to the buying unit B given an investment of b, where $v' > 0, v'' \leq 0$. Let $c(s)$ denote the cost to the selling unit S given an investment s, where $c' < 0, c'' \geq 0$. Investment cost is privately borne by the managers. Both B and S have alternative trading opportunities, the best of which yield $\omega_B(b,\theta)$ and $\omega_S(s,\theta)$ respectively. The values of those trades increase with investment as well, subject to decreasing returns. Hence $\frac{\partial \omega_B}{\partial b} > 0$, $\frac{\partial^2 \omega_B}{\partial b^2} \leq 0$ and $\frac{\partial \omega_s}{\partial s} > 0$, $\frac{\partial^2 \omega_s}{\partial s^2} \leq 0$. External trading opportunities' values also depend on a realization of a random variable θ which is observed by both managers before the investment decision. Managers receive monetary compensation based on their unit's profit $\alpha_i + \beta_i \pi_i$, $i \in \{B, S\}$ where π_i

is unit i's profit. Their utilities are $U_B = \alpha_B + \beta_B \pi_B - b$ and $U_S = \alpha_S + \beta_S \pi_S - s$, where $\pi_B = v(b) - t$ if B and S trade with each other, t is the transfer price, and $\pi_B = \omega_B(b)$ if the parties trade elsewhere. Similarly $\pi_S = t - c(s)$ if they trade with each other and $\pi_S = \omega_S(s)$ if trading elsewhere.

We compare two alternative forms of organizing the trading relation. Under nonintegration, the two units are independent firms, whereas in an integrated relation the units are divisions within a single firm. Unlike in Hart's model, the ownership of assets plays no role here, as managers are employees both under integration and non-integration. However, I maintain Hart's assumption that describing contractually the exact nature of the input that would be required by B is prohibitive costly, and therefore no contract is signed before investments are made. At the time of trade, the units bargain over the transfer price. External trade opportunities affect the division of the gains from trade between the units in the bargaining, and hence the investment decisions taken prior to the trade.

Consider first the case where internal trade is always efficient. Assume

$$v(b) - c(s) \geq \omega_B(b,\theta) + \omega_S(s,\theta), \forall (b,s,\theta).$$

When B and S are two independent firms, assume, as in Hart, that the transfer price splits the surplus over the external trade payoffs. The profits to the units are then

$$\begin{aligned} \pi_B &= \omega_B(b,\theta) + \frac{1}{2}\left[v(b) - c(s) - (\omega_B(b,\theta) + \omega_S(s,\theta))\right], \\ \pi_S &= \omega_S(s,\theta) + \frac{1}{2}\left[v(b) - c(s) - (\omega_B(b,\theta) + \omega_S(s,\theta))\right]. \end{aligned}$$

The optimal investments $\left(b^{NI}, s^{NI}\right)$ then satisfy

$$\begin{aligned} b^{NI} &= \arg\max_b \beta_B \frac{v(b) + \omega_B(b,\theta)}{2} - b, \\ s^{NI} &= \arg\max_s \beta_S \frac{-c(s) + \omega_S(s,\theta)}{2} - s. \end{aligned}$$

Now consider a similar trading relation within the confines of an integrated firm. Divisions are supervised by the general office of the firm, who is allocated all the control rights and has the last say over decisions. Assume that the general office observes neither the investments made by the divisions, nor θ. The general office hence does not know the exact values of respective trades, and the division managers are allowed to bargain over the transfer price. Given our assumption that internal trade is always efficient however, the general office does know the ranking of different trades.

Trade is always efficient and therefore internal. If the general office could have committed itself to let the divisions bargain freely and trade externally if they choose to do so, then investments levels would be as in a non-integrated relation, provided that managerial incentives, β_B, β_S are the same as under nonintegration. In fact, as will be shown below for the more general model, the general office sets stronger incentives for the divisions than those chosen by independent owners. As a result integration would outperform nonintegration.

Things are very different if the general office cannot credibly commit ahead of time not to intervene. Because internal trade is commonly known to be always efficient, the general office would not allow the divisions to opt out of jointly profitable internal trade. A threat by one of the units to opt out and trade externally is no longer credible. Even though opting out may only occur off the equilibrium path, the possibility of trading externally has a positive effect on the incentives of the divisions to invest. With no credible option to trade externally, optimal investments solve

$$\begin{aligned} b^I &= \arg\max_b \beta_B \frac{v(b)}{2} - b, \\ s^I &= \arg\max_s \beta_S \frac{-c(s)}{2} - s. \end{aligned}$$

Compared with the investments derived under nonintegration, we note that holding incentives fixed across organizational forms, both investments are lower under integration. This contrasts with the result in Hart's model, where under integration, one unit's investment is higher, and the other is lower than their levels under nonintegration.

In this setting, the boundaries of firms can be explained by a trade-off between two conflicting effects. On one hand, under integration higher–powered incentives are set for the units, but the level of investment chosen is lower for any given level of incentives. The relative efficiency of integration and nonintegration then depends on the magnitudes of the two effects. This is however somewhat misleading. As we now show, this result only holds under the assumption that external trade is never efficient. It is then in the interest of the general office to intervene whenever one of the parties choose to opt out. This assumption does not seem very plausible, as it is reasonable to believe that external trading options may be superior at least in some extreme situations. If external trade is sometimes efficient, the general office cannot rank the respective trades and it will not necessarily prevent opting out. Hence we assume instead that either internal or external trade can be efficient at times:

$$\forall (b, s), \exists \theta, s.t\ \omega_B(b, \theta) + \omega_S(s, \theta) \geq v(b) - c(s).$$

I now sketch the intuition why in this case the general office would not intervene, even if it cannot commit itself to do so: Bargaining between the divisions always results in an efficient trade.[3] In equilibrium, the divisions opt out if and only if $\omega_B(b,\theta)+\omega_S(s,\theta) > v(b)-c(s)$. Whenever this is satisfied, a transfer price t exists such that $v(b) - t \geq \omega_B(b,\theta)$ and $t - c(s) \geq \omega_S(s,\theta)$, and no such price exists otherwise. The general office observes neither θ nor the investments, and its information set therefore does not distinguish when internal trade is optimal. Given that in equilibrium the trading decision is efficient, it is optimal for the general office not to overrule the divisions' decisions to opt out. But then the external trade opportunities are again viable as outside options, as the divisions can credibly threaten to opt out. Investments are once again responsive to the returns due to external trading options, and equal those achieved if full commitment is feasible. An immediate implication is therefore that integration necessarily dominates nonintegration in this setting.

This simple framework demonstrates that a potential drawback of internalizing trade within the hierarchial structure of an integrated firm is that the possibility of discretionary intervention by the general office would have an adverse effect on the investment decisions taken by the subordinate managers. It is also shown that asymmetric information within the integrated firm's hierarchy regarding the ranking of different trades helps to discourage such intervention, and restore the investment incentives of the managers. These two themes are central in the analysis below. The simple framework is limited for explaining variation in the vertical structure of firms, since as long as external trade is sometimes optimal, integration dominates nonintegration. I now turn to consider a more complex model in which giving the units autonomy over the trading decision bears a real cost. In this context, the problem of the general office's inability to commit not to intervene reappears. In section 5 below, I consider a similar setup to that of this section in the context of the general model, and provide a more rigorous treatment.

3 The Model

The model is based on the framework of Holmstrom and Tirole (1991) (henceforth HT).[4] I study the organization of trade between a selling unit $i = 1$ and a buying unit $i = 2$. The

[3] It is important to stress that the division managers bargain over the transfer price and cannot use their resources to contract secretly between themselves. If secret side payments between the managers are possible, then unless $\beta_B = \beta_S$, the trading decision need not be efficient. Suppose that $\beta_B > \beta_S$ and that $v - c = 13$, $\omega_B = 10$ and $\omega_S = 4$. In that case external trade is efficient. However by transferring all profits from internal trade to B using a transfer price of $t = c$ and then having B secretly pay S, the divisions would do better by trading internally if $\beta_B(v-c) \geq \beta_B\omega_B + \beta_S\omega_S$ or $13\beta_B \geq 10\beta_B + 4\beta_S$, that is provided that $\beta_B \geq \frac{4}{3}\beta_S$.

[4] The model corresponds to their example 2. They also allow for monetary investment in quality and cost-reducing effort, which I do not consider here.

units may trade a single unit of intermediate product that would be used by the buying unit. Each unit is headed by an employee manager, whose incentives are determined by contract. There are two possible forms of organization: Under a nonintegrated relationship the units are independent and separately owned firms. The integrated firm is organized in a multidivisional form and each one of the units functions as a separate division. Divisions' managers respond to a general office, who has formal authority over all decisions.[5] Each division is capable of trading with external customers and suppliers without the assistance of other units being required.

Decision Variables and the Values of Alternative Trades

Each manager controls the level of two variables: a *quality-enhancing* effort s_i which increases the values of both the division's internal and external trade opportunities, and the *market orientation* of its operation $m_i \in [\underline{m}, \overline{m}]$ which determines their relative values. The cost of quality, $c(s_i)$, is borne by the manager, measured in monetary terms and strictly convex. A higher level of m_1 increases the value of the intermediate input to external costumers but at a cost of a decrease in its specialization to unit 2's needs. A higher level of m_2 raises the value to unit 2 of using substitute inputs from external suppliers but at the same time lowers the value of using unit 1's intermediate good. Thus m_i bears an opportunity cost, lowering the value of internal trade, but no direct cost to the manager. The choices made by the managers are not observed by the owners of the units. Following HT, I simplify the exposition by taking the units to be symmetric in cost and benefits. Denote by $\Phi(s_1, m_1) + \Phi(s_2, m_2)$ the value of trade between the two units, where $\Phi(s_i, m_i)$ is the value added by unit i if the units trade together (I will refer to this as internal trade, though in the nonintegrated case it should be thought of as mutual trade). Denote by $\Gamma(s_i, m_i, \theta)$ the value of division i's best external trading alternative. The value of external trade is uncertain, depending on the realization of a state variable $\theta \in [\underline{\theta}, \overline{\theta}]$. θ is distributed according to a continuous distribution with density $f(\theta)$.

Due to the incompleteness of contracts, the intermediate good cannot be described in a contract before investment decisions are made. The state variable θ and the values of the possible trades are observed later by both units' managers, but are not observed, nor verifiable by any third party, including the general office of the integrated firm.

[5]We take the general office's objective to be maximization of the owners' surplus. We hence ignore a second tier of agency relations between the general office and the firms' owners (stockholders). Bolton and Scharfstein (1998) emphasize the importance of considering explicitly this two-tier agency relation.

I make the following functional form assumptions on Φ and Γ. Those assumptions are similar to the ones made by HT, extended to the case of uncertain external trade.

Assumption 1

1. $\Phi(s,m) = s * x(m,k)$ *where* $x_m < 0$, $x_{mm} \leq 0$, $x_k \leq 0$, $x_{mk} \leq 0$, $x(\underline{m},k) = 1$, $\forall k$ *and* $x(m,\underline{k}) = 1$, $\forall m$.

2. $\Gamma(s,m,\theta) = s * \gamma(m,\theta)$ *where* $\gamma_m > 0$, $\gamma_{mm} \leq 0$, $\gamma_\theta > 0$, $\gamma_{m\theta} \geq 0$.

$x(\cdot,k)$ is a family of functions parameterized by $k \in \left[\underline{k},\overline{k}\right]$. k measures the degree to which the value of internal value depreciates with an increase in the market orientation of each unit. Both the opportunity cost and the marginal cost of market orientation increase with k. For $k = \underline{k}$, market orientation is "costless".

Assumption 2 $\Gamma\left(s,\overline{m},\overline{\theta}\right) \geq \Phi(s,\underline{m})$ *and* $\Phi(s,\overline{m}) \geq \Gamma(s,\overline{m},\underline{\theta})$. *External trade can be more or less profitable than internal trade.*

In an integrated firm, the general office can choose one of two transfer pricing policies: *exchange autonomy* or *mandated internal trade.*[6] Under the first policy the decision whether to trade with the internal partner or externally is delegated to the division managers. Under the latter the divisions are obliged to trade with one another or not at all, but are free to bargain over the transfer price and whether to trade.

Incentives and Preferences

Unit $i \in \{1,2\}$'s recorded profits π_i are a noisy measure of the actual profits. I assume that

$$\pi_i = \overline{\pi}_i + \varepsilon_i,$$

where $\varepsilon_i \sim N\left(0,\sigma^2\right)$ is additive random noise, and $\overline{\pi}_i$ actual profits. Unit managers are compensated with a linear wage contract, based on the unit's profits π_i,[7]

$$w_i = a_i + b_i\pi_i,$$

[6] This terminology follows Eccles (1985). Eccles distinguishes between mandated cost-based methods, where the transfer price is based either on actual costs or standard costs (budgeted costs), and mandated price-based ones. This distinction does not exist here as we assume that such variables are not verifiable to the general office, and transfer price is therefore negotiated even if trade is mandated.

[7] There are some good reasons why less restrictive compensation schemes, including profit sharing arrangements, may not be practical. See Holmstrom and Tirole (1991) for a detailed discussion.

where a_i is a base salary and b_i a bonus (or a piece rate) over profits. Managers are risk-averse, with mean-variance preferences:[8]

$$U_i = Ew_i - rVar(w_i) - c(s_i).$$

Each manager's reservation utility is $\underline{U}$. I assume that the risk due to θ is unsystematic and can therefore be fully diversified by the manager, but that the risk associated with ε_i is not. Therefore only the uncertainty due to ε_i enters the variance term above, which equals $b_i^2\sigma^2$.[9] Finally we assume that the general office of an integrated firm maximizes the return to the firm's owners.

Timing

The sequence of events is depicted in Figure 1:

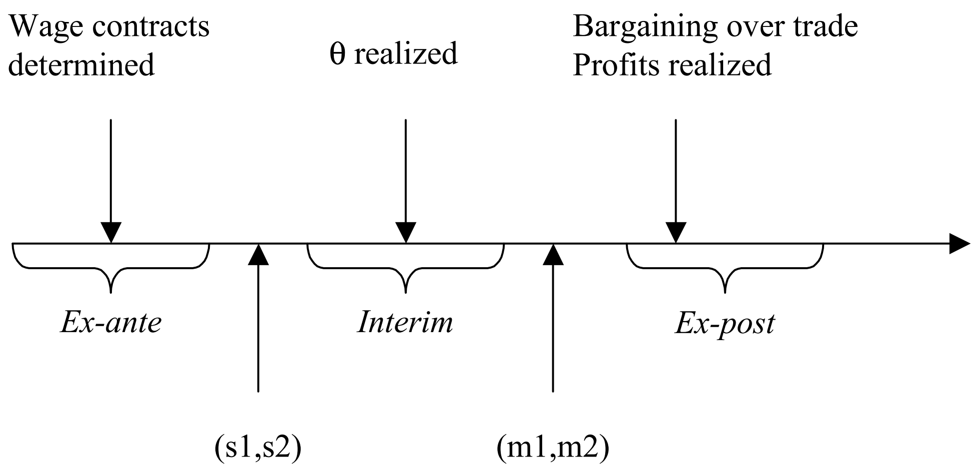

Figure 1: Sequence of Events

Contractual arrangements with the managers are made at the beginning of the employment relation. One can think of the investments s_i as representing a non-contractible investment in divisional-specific know-how over a long period of product development. The market

[8]If managers preferences are represented by a CARA utility, given the linearity of wages and the additive normal noise, utility can be represented in this mean-variance form. See Holmstrom and Milgrom (1987).

[9]Addressing the risk due to external trading value directly would unnecessarily complicate the model without any clear payoff. Risk aversion establishes a cost of incentives, so that the level of incentives may vary across organizational forms. Alternatively, one can assume risk neutrality, and introduce a cost of incentives through other means, for example by having the performance measure depend on random variables that are unobserved at the time the contract is signed but revealed only to the agent before the choice of effort level, as in Baker (1992).

orientation decision m_i is taken later, closer to the date of trade, and after the realization of θ. I identify three time periods: the *ex-ante* stage, when incentive schemes for the units managers are determined; the *interim* stage, after the choice of s_i and the realization of θ but before the decision on m_i; and the *ex-post* stage, following the decision on m_i and throughout the bargaining stage. I have not yet indicated the time at which the general office in an integrated firm chooses a transfer pricing policy. The timing of this decision turns out to be critical. I will discuss it in detail beginning in section 4.

Bargaining over the Transfer Price

At the ex-post stage, the units bargain over the transfer price t to be paid from unit 2 to unit 1. If they trade together, the gain from trade $\Phi(s_1, m_1) + \Phi(s_2, m_2)$ net of the transfer price is recorded as profit for unit 2, and the transfer price is recorded as unit 1's profit. Bargaining yields an efficient trading rule given $(s_1, m_{1,} s_2, m_2, \theta)$. Under trade between the units, the transfer price splits the surplus over the disagreement profits equally. In a non-integrated relationship and in an integrated firm under a policy of exchange autonomy, units are free to trade externally and the disagreement payoffs are $\Gamma(s_i, m_i, \theta)$ for $i \in \{1, 2\}$. The units trade with one another if and only if $\Phi(s_1, m_1)+\Phi(s_2, m_2) \geq \Gamma(s_1, m_{1,}\theta)+\Gamma(s_2, m_{2,}\theta)$. If the divisions of an integrated firm are mandated to trade internally, the disagreement payoffs equal the value of no-trade.

Under non-integration (NI) and integration+exchange autonomy (E) the transfer price is therefore determined as follows:

$$\Phi(s_1, m_1) + \Phi(s_2, m_2) - t - \Gamma(s_2, m_{2,}\theta) = t - \Gamma(s_1, m_{1,}\theta).$$

Unit $i \in \{1, 2\}$'s actual profit is then

$$\overline{\pi}_i(s_i, m_i, s_j, m_j, \theta) = \begin{cases} \frac{\Gamma(s_i,m_i,\theta)+\Phi(s_i,m_i)+\Phi(s_j,m_j)-\Gamma(s_j,m_j,\theta)}{2} & \text{if internal trade is efficient} \\ \Gamma(s_i, m_i, \theta) & \text{if external trade is efficient} \end{cases} \tag{1}$$

For the integration+mandated internal trade (M) case, the transfer price t equals

$$t = \frac{\Phi(s_1, m_1) + \Phi(s_2, m_2)}{2},$$

and unit $i \in \{1, 2\}$'s actual profit is

$$\overline{\pi}_i(s_i, m_i, s_j, m_j, \theta) = \frac{\Phi(s_i, m_i) + \Phi(s_j, m_j)}{2}. \tag{2}$$

I now proceed to the analysis of the game and consider three alternative behavioral assumptions regarding the use of transfer pricing policy in an integrated firm: First, I assume that the transfer pricing policy can be publicly announced and committed to ex-ante, before any investment decisions by the divisions. Following the announcement the general office refrains from any intervention, and the divisions comply with the announced policy. Second, I assume that no commitment is possible, and the general office may intervene in the ex-post bargaining. Finally, I consider a case where commitment is not possible ex-ante but is possible interim, after the quality-enhancing investments are sunk but before the decisions over market orientation are taken.

4 Ex-ante Commitment to a Transfer Pricing Policy

Suppose that the general office of an integrated firm can make a public commitment to a transfer pricing policy at the ex-ante stage, before any investment is made by the managers. This is the case considered in HT. Under this assumption, the integrated organization always weakly dominates the non-integration relation. The integrated firm can always commit to an exchange autonomy policy, and replicate the non-integrative structure. Furthermore, as HT establish, and as is shown in section 6 below for this model, division managers within an integrated firm receive more powered incentives, as the firm internalizes the full impact of an increase in a division's quality investment on profits. This implies that under exchange autonomy, an integrated firm would do better than independent firms. The integrated firm may do even better than this by committing to a mandated internal trade policy, spurring the divisions to specialize more closely towards one another whenever trading internally.

There are several reasons to doubt that such commitment is possible in organizations. As documented in Eccles (1985), firms tend to change transfer pricing policies over time, as external circumstances change and due to internal politics. The fact that the top management of organization changes over time may imply that the current management may not feel bound by a policy adopted by previous management. Also, large multidivisional corporations tend to employ several transfer pricing policies for different trades within the firm, making an early commitment less plausible.

5 No Commitment to a Transfer Pricing Policy

Suppose now that the general office of an integrated firm is unable to commit to a transfer pricing policy throughout the ex-ante and interim stages. Any early announcement of a transfer pricing policy can be ignored at no cost when the time to trade arrives. The general office can therefore intervene in the bargaining between the two divisions, and dictate both the identity of the trading partners and the transfer price for internal trade. In order to analyze the implications for the comparison between organizational forms, a model of the bargaining between the divisions subject to the possibility of discretionary intervention by the general office is required. Consider first the following one-stage bargaining procedure, without the possibility of intervention:

- With probability $\frac{1}{2}$, each one of the units is chosen to offer a division of internal gains from trade $\Phi_i + \Phi_j$, where $\Phi_k = \Phi(s_k, m_k)$, $k \in \{1, 2\}$. We denote the proposer by i, and the offer by $z = (z_i, z_j) \in \mathbb{R}_+^2$, such that $z_i + z_j = \Phi_i + \Phi_j$.
- Unit j then responds to i's offer. If it accepts, then the units trade with each other and the gains of trade are divided according to z. If it rejects the offer, the units trade externally and each unit $k \in \{1, 2\}$ gets $\Gamma_k = \Gamma(s_k, m_k, \theta)$. Denote j's response by $r(z) \in \{Y, N\}$.

Unit $k \in \{1, 2\}$'s preferences are given by $U_i = b_k \pi_k$. Such bargaining procedure corresponds to the non-integrated relation. One can verify the following result:

Lemma 1 *The bargaining game without possibility of intervention has a unique family of subgame perfect equilibria:* $z^* = (\Phi_i + \Phi_j - \Gamma_j, \Gamma_j)$ *if internal trade is efficient and any offer with* $z_i^* \geq \Gamma_i$ *otherwise, and* $r_j^*(z) = Y$ *if and only if* $z_j \geq \Gamma_j$*. The expected profits for unit* $k \in \{1, 2\}$ *in any equilibrium are*

$$E[\pi_k] = \begin{cases} \Gamma_k & \text{if external trade is efficient} \\ \frac{1}{2}\Gamma_k + \frac{1}{2}(\Phi_k + \Phi_{-k} - \Gamma_{-k}) = \frac{\Phi_k + \Phi_{-k} + \Gamma_k - \Gamma_{-k}}{2} & \text{if internal trade is efficient} \end{cases}$$

Thus under nonintegration, the bargaining results in an efficient trade, and a split (in expectation) of the surplus over the external trade payoffs when trade is internal.

Next consider an integrated setting, where we allow for an intervention by the general office of the integrated firm, G. We modify the bargaining procedure above as follows:

- If unit j rejects unit i's offer, the general office can intervene and force unit j to accept it. If the general office does not intervene then the units trade externally. Denote by $g(z, r(z)) \in \{0, 1\}$ the general office decision where $g = 0$ implies no-intervention.

The general office preferences are given by $U_G = (1 - b_1)\pi_1 + (1 - b_2)\pi_2$. We restrict attention to the symmetric case $b_1 = b_2 = b$ and hence $U_G = (1 - b)(\pi_1 + \pi_2)$. The general office has not observed θ, and it holds a prior $f(\theta)$ over its distribution. It also has not observed the previous choices of (s_1, s_2, m_1, m_2).

Suppose first, as is assumed in HT, that internal trade is always efficient. That is, $\forall (s_1, s_2, m_1, m_2, \theta)$, $\Phi_1 + \Phi_2 > \Gamma_1 + \Gamma_2$. We can then show the following:

Lemma 2 *If internal trade is always efficient, then in every weak perfect baysian equilibrium (PBE) of the bargaining game with intervention, $g(z, N) = 1$ and $z = (\Phi_1 + \Phi_2, 0)$. The general office always mandates internal trade if no agreement is reached by the divisions. The proposer claims all the gains from trade to itself. The expected profits of unit $k \in \{1, 2\}$ are*

$$E[\pi_k] = \frac{1}{2} * 0 + \frac{1}{2}(\Phi_k + \Phi_{-k}) = \frac{\Phi_k + \Phi_{-k}}{2}$$

Proof. As internal trade is always efficient, it maximizes joint profits and therefore given G's preferences above, internal trade is always mandated by the general office if unit j rejects the offer. Given that behavior, it is optimal for unit i to ask for the entire surplus for itself. ■

The bargaining splits the surplus over the no-trade payoffs. The values of external trading opportunities have no effect on the divisions' profits. Although the general office does not actually observe the exact values of different trades, it is still able to rank them, and as a result, it intervenes whenever a non-efficient external trade would take place. Though this happens only off the equilibrium path, the result is that external trade opportunities can no longer play a "monitoring" role, and therefore have no effect on the investment decisions of the divisions. The divisions' behavior would be identical to that in the case where internal trade is mandated ex-ante.

Now suppose as in assumption 2, that either internal or external trade may be efficient. As the next lemma shows, the bargaining game then has an equilibrium which yields an efficient trade and a split of the surplus over external trade payoffs when internal trade is efficient.

Lemma 3 *Under assumption 2, the following is a weak PBE of the game, in which trade is efficient: $z^* = (\Phi_i + \Phi_j - \Gamma_j, \Gamma_j)$ if $\Phi_i + \Phi_j > \Gamma_i + \Gamma_j$ and any offer with $z_i^* \geq \Gamma_i$ otherwise. $r^*(z) = Y$ if and only if $z_j \geq \Gamma_j$, and $g(z, N) = 0$ for every z.*

Proof. Given the equilibrium behavior of the units, the general office believes that with probability one external trade is efficient whenever j opts out, and therefore finds it optimal not to intervene. As the general office never intervenes, the equilibrium behavior of the units is optimal. ■

In the equilibrium described above, the general office does not intervene in the bargaining procedure. Whenever the divisions choose to trade externally, the general office believes that external trade is efficient and refrains from intervening. The expected profits to the units are as in the case of bargaining without intervention.[10]

It is useful to interpret this result using the concepts of "real" and "formal" authority (Aghion and Tirole (1997)). The general office of an integrated firm retains formal authority over all of the decisions regarding the trade. However, when the divisions are better informed than the general office on the ranking of different trades, the general office can credibly transfer real authority over the choice of trading partners to the units.

Even though we analyze a special bargaining procedure, the insights seem to be shared with other bargaining procedures, provided that the choice of trading externally is irreversible and final. An alternating offers game with external trade options playing the role of "outside options" (as in Osborne and Rubinstein (1990)) is likely to give a similar result.[11]

To conclude, when commitment to a transfer pricing policy is infeasible, and under assumption 2, the divisions would be given autonomy to decide on their trading partners. Holding incentives fixed at the levels of nonintegrated relation, the divisions investments (s_1, s_2), interim decisions (m_1, m_2) and their profits under integration would equal those under nonintegrated relation. By the same reasoning of the last section, integration then dominates non-integration, as incentives can be coordinated. Because internal trade is never mandated, the joint surplus under no-commitment can be lower than under ex-ante commitment.

6 Interim Commitment to a Transfer Pricing Policy

The two polar assumptions on the commitment "technology" considered above imply that integration always weakly dominates nonintegration. As such they are a limited basis for a

[10]The bargaining game has several other equilibria in which the general office intervenes, either on or off the equilibria path. For example, if we assume that $E_f[\Gamma_i + \Gamma_j] < \Phi_i + \Phi_j$, and that $\Phi_i + \Phi_j > \Gamma_i$ for all $(s_i, m_i, s_j, m_j, \theta)$, then the following is a weak PBE:

- $z^* = (\Phi_i + \Phi_j, 0)$, $r^*(z) = Y$ if and only if $z_j \geq 0$ and $g(z, N) = 1$ for every z. The general office belief following (z, N) equals its prior belief.

The trade in this equilibrium however is inefficient as the units never trade externally. The efficiency of the non-intervention equilibrium renders it more plausible then other equilibria which result in some inefficient trades.

[11]Though not a split-the-surplus rule.

positive theory of firms' boundaries. Under no-commitment the model also fails to explain the observed use of mandated internal trade. I now consider an intermediate form of commitment, assuming that the general office cannot commit ex-ante to a transfer pricing policy but can commit at the interim stage once quality-enhancing investment s_i has been sunk, but before m_i has been taken.

The proximity between the announcement of the transfer pricing policy and actual trade suggests that this form of limited commitment is more plausible then ex-ante commitment. For one reason, management is less likely to change in this time period. Also, some level of commitment is available, as we observe mandated internal trade between divisions.[12] In the remaining analysis, I therefore maintain the assumption of interim commitment.

6.1 Choice of Market Orientation

Proceeding backwards, I consider first the interim choice of market orientation m_i, given investments (s_1, s_2) and a realization of θ. Our assumptions on the bargaining imply that the units expect to trade efficiently given (s_1, s_2, θ) and their decisions on (m_1, m_2). Consider now the choice of (m_1, m_2) given each of the organizational forms:

Integration + Mandated Internal trade

Mandating internal trade effectively implies that $m_i = \underline{m}$ is optimal for $i \in \{1, 2\}$. The equilibrium choices of market orientation are therefore $m_1^* = m_2^* = \underline{m}$. The total gains from trade under mandated internal trade, from the point of view of the general office that does not observe the realization of θ, are

$$\Phi\left(s_1, \underline{m}\right) + \Phi\left(s_2, \underline{m}\right). \tag{3}$$

The cost of quality-enhancing effort and the cost of incentives (in terms of risk) which are already sunk are omitted.

Nonintegration and Integration + Exchange autonomy

[12] An alternative interpretation of interim commitment is the following. Suppose that the opportunity cost of market orientation, k, is observed only at the interim period, but that unlike θ it is observed by everyone in the firms' hierarchy. Provided that k can vary substantially, it may be too costly for the general office to commit ex-ante to a transfer pricing policy, even if it is within its power. To put it differently, the temptation to renege interim on an earlier decision may be too great. To accommodate such an interpretation, the analysis to follow has to be slightly modified, but it is conjectured that the essence of the results should nevertheless hold.

Given (b_i, b_j) and (s_i, s_j), unit i manager's equilibrium choice of market orientation m_i^* solves:

$$m_i^* = \arg\max_{m_i} a_i + b_i \left[\overline{\pi}_i \left(s_i, m_i, s_j, m_j^*, \theta\right)\right] - c(s_i) - r(b_i)^2 \sigma^2.$$

As s_i and b_i are already determined at the interim stage, this is equivalent to a maximization over $\overline{\pi}_i \left(s_i, m_i, s_j, m_j^*, \theta\right)$. When the parties trade internally in a continuation equilibrium, the part of $\overline{\pi}_i$ that depends on m_i can be seen from (1) to be

$$\frac{\Gamma(s_i, m_i, \theta) + \Phi(s_i, m_i)}{2} = s_i * h(m_i, \theta, k),$$

where

$$h(m, \theta, k) \equiv \frac{x(m, k) + \gamma(m, \theta)}{2}. \tag{4}$$

By assumption 1, $h(m, \theta, k)$ is concave in m and so h_m changes sign at most once on $[\underline{m}, \overline{m}]$. If the choice of m_i does not affect whether the trade is internal or not, then the optimal choice of market orientation given internal trade is:

$$\widetilde{m}(\theta, k) = \arg\max_{m_i \in [\underline{m}, \overline{m}]} s_i * h(m_i, \theta, k). \tag{5}$$

Given the multiplicative form of Φ and Γ, $\widetilde{m}(\theta, k)$ is independent of s_i and $\frac{\partial \widetilde{m}(\theta,k)}{\partial \theta} = -\frac{\gamma_{m\theta}}{x_{mm}+\gamma_{mm}} \geq 0$. When the discussion does not involve a comparison across k, I write $\widetilde{m}(\theta)$ for short, but the dependence in k should not be ignored.

I now characterize the equilibrium choice (m_1^*, m_2^*) given (s_1, s_2) and a realization of θ. In general, multiple equilibria are possible.

Lemma 4 *In any equilibrium of the subgame beginning in the interim stage,* $(m_1^*, m_2^*) = (\overline{m}, \overline{m})$ *if trade is external in the continuation equilibrium, and* $(m_1^*, m_2^*) = (\widetilde{m}(\theta), \widetilde{m}(\theta))$ *if trade is internal.*

Proof. If the units trade externally in the continuation equilibrium then it is clear that $m_i^* = \overline{m}, \forall i \in \{1, 2\}$. Otherwise each manager can deviate profitably to $\overline{m}$ while keeping trade external. Now suppose that in the continuation equilibrium the units trade with one another. Then $m_i^* \leq \widetilde{m}(\theta)$, $\forall i \in \{1, 2\}$ or else a deviation to $\widetilde{m}(\theta)$ increases the value of internal trade and is profitable as $\widetilde{m}(\theta)$ is optimal given internal trade. Suppose then that $m_i^* < \widetilde{m}(\theta)$ for some $i \in \{1, 2\}$. Clearly we cannot have $\Phi(s_1, m_1^*) + \Phi(s_2, m_2^*) > \Gamma(s_1, m_1^*, \theta) + \Gamma(s_2, m_2^*, \theta)$ or else, as $h(m_i, \theta, k)$ is increasing in m_i for $m_i \leq \widetilde{m}(\theta)$, some small increase in m_i^* would still result in internal trade and would be profitable. Therefore

$\Phi(s_1, m_1^*) + \Phi(s_2, m_2^*) = \Gamma(s_1, m_1^*, \theta) + \Gamma(s_2, m_2^*, \theta)$. Given the bargaining rule, each unit's profit equals its disagreement payoff $\Gamma(s_i, m_i^*, \theta)$. But then each manager can deviate to $m_i = \overline{m}$, guaranteeing himself a payoff of $\Gamma(s_i, \overline{m}, \theta)$ at least, and hence a contradiction, since $\Gamma(s_i, \overline{m}, \theta) > \Gamma(s_i, m_i^*, \theta)$ given $m_i^* < \widetilde{m}(\theta) \leq \overline{m}$ and $\Gamma_m > 0$. Therefore if trade is internal, $m_i^* = \widetilde{m}(\theta)$, $\forall i \in \{1, 2\}$. ■

Proposition 1 *For every* (s_1, s_2), *there exist* $\widetilde{\theta}(s_1, s_2) \in [\underline{\theta}, \overline{\theta}]$ *and* $\widehat{\theta}(s_1, s_2) \in [\underline{\theta}, \overline{\theta}]$, $\widehat{\theta}(s_1, s_2) < \widetilde{\theta}(s_1, s_2)$, *such that the equilibria of the subgame starting with the choice of* (m_1, m_2) *are characterized as follows:*

1. *For* $\theta \in \left[\underline{\theta},\ \widehat{\theta}(s_1, s_2)\right]$, $(m_1^*, m_2^*) = (\widetilde{m}(\theta), \widetilde{m}(\theta))$ *is the unique equilibrium and the units trade internally.*
2. *For* $\theta \in \left[\widehat{\theta}(s_1, s_2), \widetilde{\theta}(s_1, s_2)\right]$, *there are two pure strategy equilibria: one in which* $(m_1^*, m_2^*) = (\overline{m}, \overline{m})$ *and trade is external, and another where* $(m_1^*, m_2^*) = (\widetilde{m}(\theta), \widetilde{m}(\theta))$ *and trade is internal.*[13]
3. *For* $\theta \in \left[\widetilde{\theta}(s_1, s_2), \overline{\theta}\right]$, $(m_1^*, m_2^*) = (\overline{m}, \overline{m})$ *is the unique equilibrium and the units trade externally.*

Proof. Appendix. ■

As the next lemma shows, whenever multiple equilibria exist, the equilibrium in which trade is internal is Pareto-dominant.

Lemma 5 *If for* (s_1, s_2, θ) *an equilibrium of the interim subgame with* $(\widetilde{m}(\theta), \widetilde{m}(\theta))$ *and internal trade exists, then it is the Pareto-dominant equilibrium.*

Proof. By lemma 4, the only other possible equilibrium of the subgame has the units choose $m_i = \overline{m}$ each and trade externally. As each unit can always unilaterally guarantee itself at least $\Gamma(s_i, \overline{m}, \theta)$, profits under an $(\widetilde{m}(\theta), \widetilde{m}(\theta))$ equilibrium has to be at least $\Gamma(s_i, \overline{m}, \theta)$. But these are exactly the profits in an equilibrium with $(\overline{m}, \overline{m})$. ■

In what follows, I focus on equilibria of the complete game in which a Pareto-dominant equilibrium is played in the subgame starting with the choice of (m_1, m_2). Define the cutoff point between internal and external trade under nonintegration and integration+exchange autonomy by $\theta^E(s_1, s_2) = \widetilde{\theta}(s_1, s_2)$. The pareto-dominant equilibrium has

$$m_i^*(s_1, s_2, \theta) = \begin{cases} \widetilde{m}(\theta) & \text{if } \theta \leq \theta^E(s_1, s_2) \\ \overline{m} & \text{otherwise} \end{cases},$$

[13] There also exists a mixed strategy equilibrium.

for $i \in \{1,2\}$. If $s_1 = s_2 = s$, it is easily verified from (20) in the appendix, that $\theta^E(s,s)$ is independent of s which can then be omitted. In this case $x\left(\widetilde{m}\left(\theta^E\right),k\right) = \gamma\left(\overline{m},\theta^E\right)$ and $x\left(\widetilde{m}\left(\theta\right),k\right) \geq \gamma\left(\overline{m},\theta\right)$ if and only if $\theta \leq \theta^E$. An implication that is used extensively below is that

$$x\left(\widetilde{m}\left(\theta\right),k\right) \geq \frac{x\left(\widetilde{m}\left(\theta\right),k\right)+\gamma\left(\overline{m},\theta\right)}{2} \geq h\left(\widetilde{m}\left(\theta\right),\theta,k\right) \quad \text{for } \theta \leq \theta^E.$$

θ^E is also a function of k and we denote this explicitly whenever appropriate.

The expected total gains from trade amounts to:

$$\int_{\underline{\theta}}^{\theta^E(s_1,s_2)} \left\{\Phi\left(s_1,\widetilde{m}\left(\theta\right)\right)+\Phi\left(s_2,\widetilde{m}\left(\theta\right)\right)\right\} f\left(\theta\right)d\theta + \int_{\theta^E(s_1,s_2)}^{\overline{\theta}} \left\{\Gamma\left(s_1,\overline{m},\theta\right)+\Gamma\left(s_2,\overline{m},\theta\right)\right\} f\left(\theta\right)d\theta. \tag{6}$$

It is instructive to compare the choice of market orientation (m_1,m_2) made in an integrated firm under the two possible transfer pricing policies to the "first best" choices that would maximize the joint profits.

First Best (Integrated Firm)

Denote by $\left(m_1^{FB},m_2^{FB}\right)$ the first-best choices of market orientation – the choice of (m_1,m_2) that would have been chosen were m_i contractible and the general office informed about (s_1,s_2,θ). It is immediate to verify that $\forall i \in \{1,2\}$,

$$m_i^{FB}\left(\theta\right) = \begin{cases} \underline{m} & \text{if } \Phi\left(s_1,\underline{m}\right)+\Phi\left(s_2,\underline{m}\right) > \Gamma\left(s_1,\overline{m},\theta\right)+\Gamma\left(s_2,\overline{m},\theta\right), \\ \overline{m} & \text{Otherwise.} \end{cases}$$

Define $\theta^{FB}(s_1,s_2)$ as the solution with respect to θ to $\Phi\left(s_1,\underline{m}\right)+\Phi\left(s_2,\underline{m}\right) = \Gamma\left(s_1,\overline{m},\theta\right)+\Gamma\left(s_2,\overline{m},\theta\right)$. Given the multiplicative form of Φ and Γ, $\theta^{FB}(s_1,s_2)$ is independent of $(s_{1,}s_2)$. The expected total gains from trade amounts to:

$$\int_{\underline{\theta}}^{\theta^{FB}} \left\{\Phi\left(s_1,\underline{m}\right)+\Phi\left(s_2,\underline{m}\right)\right\} f\left(\theta\right)d\theta + \int_{\theta^{FB}}^{\overline{\theta}} \left\{\Gamma\left(s_1,\overline{m},\theta\right)+\Gamma\left(s_2,\overline{m},\theta\right)\right\} f\left(\theta\right)d\theta. \tag{7}$$

I now compare the cutoff level θ^E between internal and external trade under exchange autonomy to the first-best cutoff level θ^{FB} defined above.

Lemma 6 *For every (s_1,s_2), $\theta^E(s_1,s_2) \leq \theta^{FB}$.*

Proof. $\forall(s_1,s_2)$ and $\forall\theta \leq \theta^E(s_1,s_2)$, $\left(\widetilde{m}\left(\theta\right),\widetilde{m}\left(\theta\right)\right)$ is an equilibrium, and so we know that

$$\begin{aligned} s_1\gamma(\overline{m},\theta) &\leq \frac{s_1x\left(\widetilde{m}\left(\theta\right),k\right)+s_2x\left(\widetilde{m}\left(\theta\right),k\right)+s_1\gamma\left(\widetilde{m}\left(\theta\right),\theta\right)-s_2\gamma(\widetilde{m}\left(\theta\right),\theta)}{2}, \\ s_2\gamma(\overline{m},\theta) &\leq \frac{s_2x\left(\widetilde{m}\left(\theta\right),k\right)+s_1x\left(\widetilde{m}\left(\theta\right),k\right)+s_2\gamma\left(\widetilde{m}\left(\theta\right),\theta\right)-s_1\gamma(\widetilde{m}\left(\theta\right),\theta)}{2}. \end{aligned}$$

Adding up we get $(s_1 + s_2)\,\gamma(\overline{m}, \theta) \leq (s_1 + s_2)\,x\left(\widetilde{m}\left(\theta\right), k\right)$, and consequentially as $x\left(\widetilde{m}\left(\theta\right), k\right) \leq x\left(\underline{m}, k\right)$,

$$(s_1 + s_2)\,\gamma(\overline{m}, \theta) \leq (s_1 + s_2)\,x\left(\underline{m}, k\right),$$

and hence $\theta \leq \theta^{FB}$. Therefore $\theta^{E}\left(s_1, s_2\right) \leq \theta^{FB}$. ■

Many of the managers in multidivisional corporations interviewed by Eccles (1985) felt that whenever an exchange autonomy policy was employed, profitable trading opportunities within the firm were foregone, and "too little" internal trade was taken relative to what they perceived as optimal. The lemma above shows that indeed exchange autonomy can lead to inefficiently low levels of internal trade. Managers do not specialize enough towards their internal partners and as a result they find it more profitable to trade outside later on.

6.2 Choice of Incentives and Investments

I now turn to the choice of incentives and quality in the ex-ante stage. Consider first the choice of incentives $\{(a_i, b_i)\}_{i=1}^{2}$. In a **non-integrated** setting, unit i's owner chooses (a_i, b_i) to satisfy

$$\max_{a_i, b_i, s_i} \int_{\underline{\theta}}^{\theta^E(s_i, s_j)} \overline{\pi}\left(s_i, \widetilde{m}\left(\theta\right), s_j, \widetilde{m}\left(\theta\right), \theta\right) f\left(\theta\right) d\theta + \int_{\theta^E(s_i, s_j)}^{\overline{\theta}} \Gamma\left(s_i, \overline{m}, \theta\right) f\left(\theta\right) d\theta - Ew_i$$

subject to

$$(i) \quad s_i \in \arg\max_{s} \begin{array}{l} a_i + b_i \left(\int_{\underline{\theta}}^{\theta^E(s, s_j)} \overline{\pi}\left(s, \widetilde{m}\left(\theta\right), s_j, \widetilde{m}\left(\theta\right), \theta\right) f\left(\theta\right) d\theta + \int_{\theta^E(s, s_j)}^{\overline{\theta}} \Gamma\left(s, \overline{m}, \theta\right) f\left(\theta\right) d\theta \right) \\ -c\left(s\right) - r b_i^2 \sigma^2 \end{array}$$

$$\begin{array}{l} (ii) \quad a_i + b_i \left(\int_{\underline{\theta}}^{\theta^E(s_i, s_j)} \overline{\pi}\left(s_i, \widetilde{m}\left(\theta\right), s_j, \widetilde{m}\left(\theta\right), \theta\right) f\left(\theta\right) d\theta + \int_{\theta^E(s_i, s_j)}^{\overline{\theta}} \Gamma\left(s_i, \overline{m}, \theta\right) f\left(\theta\right) d\theta \right) \\ \quad - c\left(s_i\right) - r b_i^2 \sigma^2 \geq \underline{U} \end{array}$$

Given the linearity of the wage contract w_i, the base salary a_i is chosen to satisfy the manager's individual rationality constraint with equality. One can then substitute the wage term into the objective and drop this constraint. Given the functional forms in assumption

1, and after omitting constant terms, the program can be written as follows:

$$
\begin{aligned}
&\max_{b_i,s_i} \int\limits_{\underline{\theta}}^{\theta^E(s_i,s_j)} \left[s_i \tfrac{x(\widetilde{m}(\theta),k)+\gamma(\widetilde{m}(\theta),\theta)}{2} + s_j \tfrac{x(\widetilde{m}(\theta),k)-\gamma(\widetilde{m}(\theta),\theta)}{2}\right] f(\theta)\, d\theta \\
&\qquad + \int_{\theta^E(s_i,s_j)}^{\overline{\theta}} s_i * \gamma(\overline{m},\theta) f(\theta)\, d\theta - c(s_i) - rb_i^2\sigma^2 \\
&\text{subject to} \\
&s_i \in \arg\max_s b_i \int_{\underline{\theta}}^{\theta^E(s,s_j)} \left(s \tfrac{x(\widetilde{m}(\theta),k)+\gamma(\widetilde{m}(\theta),\theta)}{2} + s_j \tfrac{x(\widetilde{m}(\theta),k)-\gamma(\widetilde{m}(\theta),\theta)}{2}\right) f(\theta)\, d\theta\ + \\
&b_i \int_{\theta^E(s,s_j)}^{\overline{\theta}} s * \gamma(\overline{m},\theta) f(\theta)\, d\theta - c(s).
\end{aligned}
\tag{8}
$$

In a similar fashion, one derives the program for an integrated firm. The choice of incentives in an integrated firm is conditioned on the transfer pricing policy that would be chosen in equilibrium interim. For **exchange autonomy** (E), optimal incentives are chosen to satisfy

$$
\max_{\{b_i,s_i\}_{i=1}^2} \sum_{i=1}^2 s_i \left[\int\limits_{\underline{\theta}}^{\theta^E(s_i,s_j)} x(\widetilde{m}(\theta),k) f(\theta)\, d\theta + \int\limits_{\theta^E(s_i,s_j)}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\, d\theta\right] - \sum_{i=1}^2 c(s_i) - \sum_{i=1}^2 rb_i^2\sigma^2
\tag{9}
$$

$$
\begin{aligned}
&\text{subject to} \\
&s_i \in \arg\max_s\ b_i \int_{\underline{\theta}}^{\theta^E(s,s_j)} \left(s \tfrac{x(\widetilde{m}(\theta),k)+\gamma(\widetilde{m}(\theta),\theta)}{2} + s_j \tfrac{x(\widetilde{m}(\theta),k)-\gamma(\widetilde{m}(\theta),\theta)}{2}\right) f(\theta)\, d\theta + \\
&b_i \int_{\theta^E(s,s_j)}^{\overline{\theta}} s * \gamma(\overline{m},\theta) f(\theta)\, d\theta - c(s), \quad i \in \{1,2\}
\end{aligned}
$$

For **mandated internal trade** (M), the optimal incentives are chosen to satisfy

$$
\begin{aligned}
&\max_{\{b_i,s_i\}_{i=1}^2} \sum_{i=1}^2 s_i \int\limits_{\underline{\theta}}^{\overline{\theta}} x(\underline{m},k) f(\theta)\, d\theta - \sum_{i=1}^2 c(s_i) - \sum_{i=1}^2 rb_i^2\sigma^2 \\
&\text{subject to} \\
&s_i \in \arg\max_s b_i * s \int_{\underline{\theta}}^{\overline{\theta}} \tfrac{x(\underline{m},k)}{2} f(\theta)\, d\theta - c(s), \quad i \in \{1,2\}
\end{aligned}
\tag{10}
$$

Incentives can be seen therefore to maximize the joint gains from trade under the incentive compatibility constraints.

We restrict attention to a symmetric choice of incentives, $b_1 = b_2 = b^j(k)$ for $j \in \{NI, M, E\}$[14]. Given this, there is a continuation equilibrium in which $s_1 = s_2 = s^j(k)$ for $j \in \{NI, M, E\}$. Under exchange autonomy (E) and non-integration (NI), the investments in this symmetric continuation equilibrium are characterized by the following first-order condition:

$$
b^j \left[\int_{\underline{\theta}}^{\theta^E} h(\widetilde{m}(\theta),\theta,k) f(\theta)\, d\theta + \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\, d\theta\right] - c'\left(s^j\right) = 0 \quad j \in \{NI, E\},
$$

[14]Given the symmetry of the problem, symmetric incentives may be (but are not proved to be here) the equilibrium choice of incentives when asymmetric incentives are allowed.

where $h(m,\theta,k)$ is defined in (4), and $\theta^E(s_1,s_2)$ independent of (s_1,s_2) given that $s_1 = s_2 = s^j$. The derivative with respect to the limits of integration can be seen to be proportional to $s^j[x(\widetilde{m}(\theta^E),k) - \gamma(\overline{m},\theta^E)]$, which equals zero as $x(\widetilde{m}(\theta^E),k) = \gamma(\overline{m},\theta^E)$.

Define $\widehat{s}(\cdot) = c'^{-1}(\cdot)$. As $c' \geq 0$, $c'' \geq 0 \rightarrow \widehat{s}' \geq 0$. We can then write:

$$
\begin{aligned}
s^{NI}(k) &= \widehat{s}\left(b^{NI}(k)\left[\int_{\underline{\theta}}^{\theta^E} h(\widetilde{m}(\theta),\theta,k) f(\theta)\,d\theta + \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\,d\theta\right]\right), \qquad (11)\\
s^{E}(k) &= \widehat{s}\left(b^{E}(k)\left[\int_{\underline{\theta}}^{\theta^E} h(\widetilde{m}(\theta),\theta,k) f(\theta)\,d\theta + \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\,d\theta\right]\right),
\end{aligned}
$$

and similarly for mandated internal trade

$$s^M(k) = \widehat{s}\left(b^M(k)\left[\int_{\underline{\theta}}^{\overline{\theta}} \frac{x(\underline{m},k)}{2} f(\theta)\,d\theta\right]\right).$$

Comparing the programs for non-integration, (8), and exchange autonomy, (9), above, we note that the choice of s_i is equal given the same level of incentives, but as $x(\widetilde{m}(\theta),k) \geq h(\widetilde{m}(\theta),\theta,k)$ for $\theta \leq \theta^E$, the return to an increase in incentives is higher under exchange autonomy, and so managers receive higher-powered incentives. The common owner internalizes the full effect that an increase in s_i has on the gains from internal trade, whereas the owner of a nonintegrated unit internalizes only one half of that gain plus an additional smaller part through the increase in external trade value. Other comparisons between incentives under various regimes, are in general ambiguous.

The total ex-ante expected gains from trade for the various organizational forms are then defined as follows:

$$
\begin{aligned}
V_{(EX)}^{NI}(k) &= 2s^{NI}(k)\left[\int_{\underline{\theta}}^{\theta^E} x(\widetilde{m}(\theta),k) f(\theta)\,d\theta + \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\,d\theta\right] - 2c\left(s^{NI}(k)\right) - 2r\left(b^{NI}\right)^2\sigma^2\\
V_{(EX)}^{E}(k) &= 2s^{E}(k)\left[\int_{\underline{\theta}}^{\theta^E} x(\widetilde{m}(\theta),k) f(\theta)\,d\theta + \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\,d\theta\right] - 2c\left(s^{E}(k)\right) - 2r\left(b^{E}\right)^2\sigma^2\\
V_{(EX)}^{M}(k) &= 2s^{M}(k)\int_{\underline{\theta}}^{\overline{\theta}} x(\underline{m},k) f(\theta)\,d\theta - 2c\left(s^{M}(k)\right) - 2r\left(b^{M}\right)^2\sigma^2 \qquad (12)
\end{aligned}
$$

6.3 Optimal Interim Transfer Pricing Policy in an Integrated Firm

In this section, I consider the problem facing the general office of an integrated firm when it comes to choose a transfer pricing policy in the interim stage. Given that it set up divisions' managers incentives at $b_i = b$ for $i \in \{1,2\}$, the general office expects them to invest $s_i = s(b)$ each. As the actual investments are not observed, the choice between transfer pricing policies is based on these expected values. Define the expected gains from trade at the interim stage, given the different transfer pricing policies as follows:

$$V^M_{(IN)}(b,k) \equiv 2s(b)\int_{\underline{\theta}}^{\overline{\theta}} x(\underline{m},k)\,f(\theta)\,d\theta, \tag{13}$$

$$V^E_{(IN)}(b,k) \equiv 2s(b)\left\{\int_{\underline{\theta}}^{\theta^E} x(\widetilde{m}(\theta,k),k)\,f(\theta)\,d\theta + \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m},\theta)\,f(\theta)\,d\theta\right\}. \tag{14}$$

The terms reflecting investment costs and incentives' costs are omitted as they are already sunk. As argued before, θ^E does not depend on the investments as $s_1 = s_2$.

For the purpose of comparison, define also the interim first-best

$$V^{FB}_{(IN)}(b,k) = 2s(b)\left\{\int_{\underline{\theta}}^{\theta^{FB}} x(\underline{m},k)\,f(\theta)\,d\theta + \int_{\theta^{FB}}^{\overline{\theta}} \gamma(\overline{m},\theta)\,f(\theta)\,d\theta\right\}. \tag{15}$$

Each of the transfer pricing policies entail a cost relative to the interim first-best. Mandating internal trade leads to inefficient internal trade if $\theta \in \left[\theta^{FB},\overline{\theta}\right]$. An exchange autonomy results in excessive market orientation (or too little specialization) for $\theta \in \left[\underline{\theta},\theta^{E}\right]$, and inefficient external trade for $\theta \in \left[\theta^{E},\theta^{FB}\right]$. The interim decision on a transfer pricing policy requires a comparison of those costs. It is optimal for the general office to mandate internal trade if the expected value of the residual profits under such policy is larger than under exchange autonomy and vice-versa. Therefore the general office would mandate internal trade if and only if [15]

$$\Delta M_{(IN)}(k) \equiv (1-b)\left(V^M_{(IN)}(b,k) - V^E_{(IN)}(b,k)\right) \geq 0.$$

Given assumption 1, the expression for $\Delta M_{(IN)}(k)$ can be written as follows:

$$\begin{aligned}\Delta M_{(IN)}(k) &= (1-b)\left(V^M_{(IN)}(b,k) - V^E_{(IN)}(b,k)\right) = \\ &(1-b)*2s(b)\int_{\underline{\theta}}^{\overline{\theta}} x(\underline{m},k)\,f(\theta)\,d\theta - \\ &(1-b)*2s(b)\left[\int_{\underline{\theta}}^{\theta^E(k)} x(\widetilde{m}(\theta,k),k)\,f(\theta)\,d\theta + \int_{\theta^E(k)}^{\overline{\theta}} \gamma(\overline{m},\theta)\,f(\theta)\,d\theta\right]\end{aligned} \tag{16}$$

[15] If the general office does not commit to any transfer pricing policy at the interim stage, it has effectively chosen exchange autonomy. See the discussion on no commitment, Section 5.

The important thing to note is that the sign of $\Delta M_{(IN)}(k)$ and therefore the optimal choice of transfer pricing interim does not depend on b or s.

I now turn to study how $\Delta M_{(IN)}(k)$, and correspondingly the optimal interim transfer pricing policy, varies with k. I make the following additional assumption:

Assumption 3 *For every θ and every k, $-x_m \frac{x_{mk}}{x_{mm}+\gamma_{mm}} + x_k \leq 0$ at $m = \widetilde{m}(\theta, k)$.*

The total effect of a change in k on $x(\widetilde{m}(\theta, k), k)$, and hence on the value of internal trade under exchange autonomy is $\frac{d\ x(\widetilde{m}(\theta,k),k)}{dk} = x_m \frac{d\widetilde{m}}{dk} + x_k$. It is composed of a negative direct effect x_k and an indirect effect $x_m \frac{d\widetilde{m}}{dk}$. Differentiating the first-order condition of (5) with respect to (m, k) yields $\frac{d\widetilde{m}}{dk} = -\frac{x_{mk}}{x_{mm}+\gamma_{mm}} \leq 0$. The assumption states that the direct effect is stronger than the indirect effect, so that the total effect is negative.

Lemma 7 *If $\Delta M_{(IN)}(k) \geq 0$ for some $k \in [\underline{k}, \overline{k}]$, then $\Delta M_{(IN)}(k') \geq 0$, $\forall k' \geq k$.*

Proof. $V^M_{(IN)}$ is independent of k, as by assumption 1, $x(\underline{m}, k) = 1$, $\forall k$. Consider then the effect on $V^E_{(IN)}$. The derivative of $\int_{\underline{\theta}}^{\theta^E(k)} x(\widetilde{m}(\theta, k), k) f(\theta) d\theta + \int_{\theta^E(k)}^{\overline{\theta}} \gamma(\overline{m}, \theta) f(\theta) d\theta$ with respect to the limits of integration equals zero as $\gamma(\overline{m}, \theta^E) = x(\widetilde{m}(\theta^E, k), k)$. Therefore

$$\frac{\partial V^E_{(IN)}(k)}{\partial k} = 2s(b) \int_{\underline{\theta}}^{\theta^E} \frac{d[x(\widetilde{m}(\theta, k), k)]}{dk} f(\theta) d\theta \leq 0,$$

given assumption 3. ∎

The lemma implies that a cutoff value $k_1 \in [\underline{k}, \overline{k}]$ exists such that a policy of mandated internal trade is optimal interim if and only if $k \geq k_1$. If the opportunity cost of market orientation is not too high, the general office prefers to allow the divisions to choose their trading partners themselves so that profitable external trading opportunities would not be forgone.

6.4 A Comparison to the Optimal Ex-ante Transfer Pricing Policy

The optimal choice of transfer pricing policy at the interim stage does not depend on the level of quality-enhancing investments (s_1, s_2) which are already sunk. If on the other hand the general office was able to commit ex-ante to a transfer pricing policy, its decision would also take into account the effect on those investments. From an ex-ante perspective, the total gains from trade under mandated internal trade are higher in expectation than under exchange autonomy if $\Delta M_{(EX)} \geq 0$, where:

$$\Delta M_{(EX)}(k) = V^M_{(EX)}(k) - V^E_{(EX)}(k) = \tag{17}$$
$$2\left\{ s^M(k) \int_{\underline{\theta}}^{\overline{\theta}} x(\underline{m}, k) f(\theta) d\theta - c\left(s^M(k)\right) - r\left(b^M(k)\right)^2 \sigma^2 \right\} -$$
$$2\left\{ s^E(k) \int_{\underline{\theta}}^{\theta^E} x(\widetilde{m}(\theta), k) f(\theta) d\theta + s^E(k) \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m}, \theta) f(\theta) d\theta - c\left(s^E(k)\right) - r\left(b^E(k)\right)^2 \sigma^2 \right\}.$$

$s^M(k)$, $s^E(k)$ are the symmetric equilibrium investments in quality given mandated internal trade and exchange autonomy respectively (defined in (11)), and $b^M(k), b^E(k)$ are the optimal incentives. The next proposition shows that as commitment to a transfer pricing policy is only possible at the interim stage, internal trade is mandated in "too many" cases compared to the ex-ante optimum. The reason is that ex-ante the general office also takes into account the adverse effect that mandating internal trade has on the investment levels.

Proposition 2 *All trades that would have been mandated ex-ante are also mandated interim: For all $k \in [\underline{k}, \overline{k}]$, $\Delta M_{(EX)}(k) \geq 0 \Rightarrow \Delta M_{(IN)}(k) \geq 0$.*

Proof. Suppose that $\Delta M_{(EX)}(k) \geq 0$. Denote by $s^{\overline{M}}$ the symmetric equilibrium quality investment when exchange autonomy is chosen ex-ante and incentives are set to b^M. As b^E is optimal, a revealed preference argument then implies that

$$\begin{aligned} V^E_{(EX)} &= 2\left(s^E \int_{\underline{\theta}}^{\theta^E} x(\widetilde{m}(\theta), k) f(\theta) d\theta + s^E \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m}, \theta) f(\theta) d\theta - c\left(s^E\right) - r\left(b^E\right)^2 \sigma^2 \right) \\ &\geq 2\left(s^{\overline{M}} \int_{\underline{\theta}}^{\theta^E} x(\widetilde{m}(\theta), k) f(\theta) d\theta + s^{\overline{M}} \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m}, \theta) f(\theta) d\theta - c\left(s^{\overline{M}}\right) - r\left(b^M\right)^2 \sigma^2 \right). \end{aligned}$$

Therefore as $\Delta M_{(EX)} = V^M_{(EX)} - V^E_{(EX)} \geq 0$,

$$\begin{aligned} &2\left(s^M \int_{\underline{\theta}}^{\overline{\theta}} x(\underline{m}, k) f(\theta) d\theta - c\left(s^M\right) \right) \\ &\qquad - 2\left(s^{\overline{M}} \int_{\underline{\theta}}^{\theta^E} x(\widetilde{m}(\theta), k) f(\theta) d\theta + s^{\overline{M}} \int_{\theta^E}^{\overline{\theta}} \gamma(\overline{m}, \theta) f(\theta) d\theta - c\left(s^{\overline{M}}\right) \right) \\ &\qquad\qquad \geq \Delta M_{(EX)} \geq 0\ , \end{aligned}$$

where

$$
\begin{aligned}
s^{M} &= \widehat{s}\left(b^{M}\int\limits_{\underline{\theta}}^{\overline{\theta}}\frac{x\left(\underline{m},k\right)}{2}f\left(\theta\right)d\theta\right), \\
s^{\overline{M}} &= \widehat{s}\left(b^{M}\left(\int\limits_{\underline{\theta}}^{\theta^{E}}h\left(\widetilde{m}\left(\theta\right),\theta,k\right)f\left(\theta\right)d\theta+\int\limits_{\theta^{E}}^{\overline{\theta}}\gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta\right)\right).
\end{aligned}
$$

By definition $h\left(\widetilde{m}\left(\theta\right),\theta,k\right)\geq h\left(\underline{m},\theta,k\right)\geq\frac{x(\underline{m},k)}{2}$ $\forall\theta$, and $\forall\theta\geq\theta^{E}$, $\gamma\left(\overline{m},\theta\right)\geq x\left(\widetilde{m}\left(\theta\right),k\right)\geq h\left(\widetilde{m}\left(\theta\right),\theta,k\right)$ implying $\gamma\left(\overline{m},\theta\right)\geq\frac{x(\underline{m},k)}{2}$. Therefore

$$
\int_{\underline{\theta}}^{\theta^{E}}h\left(\widetilde{m}\left(\theta\right),\theta,k\right)+\int_{\theta^{E}}^{\overline{\theta}}\gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta\geq\int_{\underline{\theta}}^{\overline{\theta}}\frac{x\left(\underline{m},k\right)}{2},
$$

and $s^{\overline{M}}\geq s^{M}$ as $\widehat{s}'\geq0$. Now the function $s*\int_{\underline{\theta}}^{\theta^{E}}x\left(\widetilde{m}\left(\theta\right),k\right)f\left(\theta\right)d\theta+s*\int_{\theta^{E}}^{\overline{\theta}}\gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta-c\left(s\right)$ is concave in s, and is maximized at

$$
s^{*}=\widehat{s}\left(\int_{\underline{\theta}}^{\theta^{E}}x\left(\widetilde{m}\left(\theta\right),k\right)f\left(\theta\right)d\theta+\int_{\theta^{E}}^{\overline{\theta}}\gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta\right).
$$

As $x\left(\widetilde{m}\left(\theta\right),k\right)\geq h\left(\widetilde{m}\left(\theta\right),\theta,k\right)$ then $s^{*}\geq s^{\overline{M}}\geq s^{M}$. Its value at $s^{\overline{M}}$ is therefore higher than at s^{M}. We can then substitute s^{M} for $s^{\overline{M}}$ in the inequality above, simplify and conclude that

$$
2\left(\int\limits_{\underline{\theta}}^{\overline{\theta}}x\left(\underline{m},k\right)f\left(\theta\right)d\theta\right)-2\left(\int\limits_{\underline{\theta}}^{\theta^{E}}x\left(\widetilde{m}\left(\theta\right),k\right)f\left(\theta\right)d\theta+\int\limits_{\theta^{E}}^{\overline{\theta}}\gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta\right)\geq0,
$$

which implies that $\Delta M_{(IN)}\left(k\right)\geq0$. ■

Interim, the general office mandates trade in all cases (i.e. for all values of k) it would mandate trade ex-ante if able to commit. Note that if $\theta^{E}<\overline{\theta}$, then $s^{\overline{M}}>s^{M}$, and we can strengthen the result above and claim that $\Delta M_{(EX)}\left(k\right)=0\Rightarrow\Delta M_{(IN)}\left(k\right)>0$. This implies that for values of k at which the general office is indifferent between transfer pricing policies ex-ante it strictly prefers to mandate internal trade in the interim stage.

The results of proposition 2 are summarized in the following diagram:

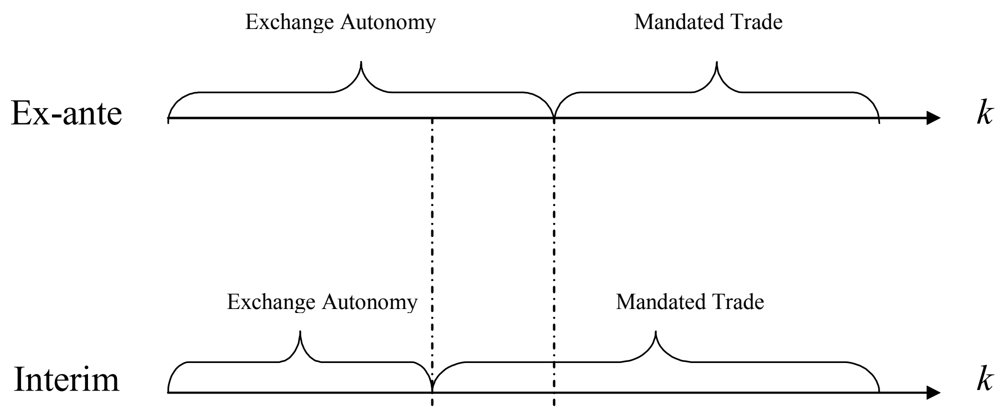

Figure 2: Ex-ante vs. interim optimal policies

6.5 Choice of Organizational Form

A long-standing tradition in the literature on the vertical organization of firms is to assume that trade is organized in a form that maximizes the joint (ex-ante) gains from trade (see for example Hart (1995)). In the environment discussed here, the value of the integrated firm is constrained because the transfer pricing policy is chosen at an interim stage. Whenever the ex-ante optimal transfer pricing policy is the same as the interim optimum, integration must be (at least) weakly optimal. But as demonstrated in proposition 2, the general office tends to mandate internal trade in too many instances. In those cases, where exchange autonomy (E) is optimal ex-ante but mandated internal trade (M) is interim optimal, the optimal organizational form is found by a comparison of a the total gains from trade under a non-integrated relationship, $V^{NI}_{(EX)}$, to that under integration with mandated internal trade, $V^{M}_{(EX)}$. The comparison trades off the advantages of "coordinating" the incentives of the two managers (shared by all integrated forms) and of optimal specialization for internal trades under mandated internal trade, with the advantages of non-integration in terms of the flexibility to trade externally when profitable, and the improved incentives for quality-enhancing investment (for a fixed level of incentives). I summarize this as follows:

Lemma 8 *Suppose that the organizational form maximizes the ex-ante gains from trade. Then for any* $k \in \left[\underline{k}, \overline{k}\right]$,

1. *If* $V^{E}_{(IN)}(k) \geq V^{M}_{(IN)}(k)$ *then the units would be integrated and be given an exchange autonomy.*

2. If $V^E_{(IN)}(k) < V^M_{(IN)}(k)$ then if $V^{NI}_{(EX)}(k) < V^M_{(EX)}(k)$ the units would be integrated and internal trade would be mandated, otherwise the units would be nonintegrated firms.

Proof.

1. From proposition 2, $V^E_{(IN)}(k) \geq V^M_{(IN)}(k) \Rightarrow V^E_{(EX)}(k) \geq V^M_{(EX)}(k)$. As $V^E_{(EX)}(k) \geq V^{NI}_{(EX)}(k)$ for every k, the ex-ante optimum is also the interim optimum and therefore E is the optimal organizational form.

2. If $V^E_{(IN)}(k) < V^M_{(IN)}(k)$, a transfer policy M would be chosen in an integrated firm interim and the ex-ante comparison is then between M and NI.

■

7 Comparative Statics

In this section I study how changes in the trading environment affect the optimal choice of organizational form. I have already noted in the previous section that the choice of organizational form depends on the value of k, the opportunity cost of market orientation. Variation in k may be interpreted in one of two ways: first, as cross-sectional differences in the specialization of inputs within an industry; second, as intrafirm variation between different inputs that are manufactured in-house, with similar implications. Lemma 7 and proposition 2 together imply that exchange autonomy is optimal both interim and ex-ante for $k < k_1$, where k_1 is the interim cutoff value between transfer pricing policies, and therefore integration is necessarily weakly optimal for those values of k. Trades in which internal value only moderately decrease if the input is not fully specialized, are more likely to be taken internally within an integrated firm. In these cases, the divisions would be given exchange autonomy and allowed to trade externally. If the cost of insufficient specialization is higher, exchange autonomy would not be sustainable and the production of the input would either be outsourced to an independent supplier (if the provision of incentives to the units is more important) or done in-house (if specialization is more important), in which case internal trade would be mandated. Integrated firms with very specialized needs (so that $V^M_{(EX)} > V^E_{(EX)}$) would mandate internal trade.

7.1 Distribution of external trade opportunities

What is the effect of a favorable change in the distribution of external trade opportunities? I consider this question for the special case of $\gamma(m, \theta) = (m + \theta)$. For this form, as can be seen from (5), the choice of orientation, $\widetilde{m}(\theta, k)$, is independent of θ. A first-order stochastic dominance increase in $f(\theta)$ then raises the total gains from trade under non-integration and integration+exchange autonomy both from an interim, (14), and ex-ante, (12), perspectives, but has no effect on the gains from trade if internal trade is mandated.[16] By lemma 8, there are several possible effects:

1. If $V^M_{(IN)}(k) < V^E_{(IN)}(k)$ before the shift in $f(\theta)$, the optimal organizational form both before and after is integration + exchange autonomy (E).

2. If $V^E_{(IN)}(k) < V^M_{(IN)}(k)$ before the shift in $f(\theta)$ and $V^E_{(IN)}(k) \geq V^M_{(IN)}(k)$ after, then if prior to the change $V^M_{(EX)}(k) < V^{NI}_{(EX)}(k)$ as well, the organization of trade would change from non-integration (NI) to integration+exchange autonomy (E). If prior to the change $V^M_{(EX)}(k) \geq V^{NI}_{(EX)}(k)$, we would observe a change from M into E.

3. If $V^E_{(IN)}(k) < V^M_{(IN)}(k)$ before and after the shift in $f(\theta)$, then by the second part of lemma 8, the effect is determined by comparing the ex-ante gains from trade between non-integration and integration + mandated internal trade. For $k \in [\underline{k}, \overline{k}]$ such that $V^M_{(EX)}(k) > V^{NI}_{(EX)}(k)$ before the shift and $V^M_{(EX)}(k) < V^{NI}_{(EX)}(k)$ after, the vertical structure would change from integration to non-integration.

One conclusion that we can draw is that a first order stochastic dominance increase in the distribution of θ results in fewer instances (fewer values of k) for which integration+mandated trade is optimal. But an analysis that ignores the internal organization of integrated firms (i.e. the transfer pricing policy) but focuses narrowly on the choice of vertical structure (integration vs. nonintegration) may find no clear effect.

7.2 Cooperativeness of investment

Up to this point, we have limited attention to "self-investments" in quality: investments by unit i that increase the value of i's internal and external trade options, but have no effect on the value of j's trade with other partners. The literature on incomplete contracting and vertical integration (Che and Hausch (1999), Whinston (2003)) has also considered a second type of investment, called "cooperative investment". An investment by unit i is said to be cooperative when it increases the value of unit j's external trading options. An example of a cooperative investment is a buyer who deploys a team of quality specialists to help a

[16]For the general case, as $\frac{\partial \widetilde{m}}{\partial \theta} \geq 0$ and $x_m \leq 0$, such shift in the distribution of θ results in a decrease in the expected value of trade conditional on it being internal. The overall effect on the total gains from trade under exchange autonomy and nonintegration is therefore ambiguous.

supplier streamline its operation. Such an investment also increases the competitiveness of the supplier in working with other buyers. Here we briefly consider the implications of introducing a cooperative element to the investments in quality by the units. We reformulate assumption 1 as follows:

Assumption 4

1. $\Phi(s_i, s_j, m_i) = [(1-\lambda)s_i + \lambda s_j] * x(m_i, k)$.
2. $\Gamma(s_i, s_j, m_i, \theta) = [(1-\lambda)s_i + \lambda s_j] * \gamma(m_i, \theta)$.

$(1-\lambda)s_j$ is the "selfish" part of unit j's investment and λs_j is the "cooperative" part. $\lambda = 0$ is the self-investment case. In appendix A.2, it is shown that the analysis of the previous sections carries over to cooperative investments of this form, provided that λ is small enough. It is also shown that a small increase in λ has no effect on $V^{M}_{(EX)}$ but lowers $V^{NI}_{(EX)}$ and $V^{E}_{(EX)}$.

Consider the effect of a small increase in the "cooperativeness" of investments (an increase in λ) on the optimal organizational form. The interim choice of transfer pricing policy in an integrated firm does not change, as it is not affected by the level of investments. By the second part of lemma 8 then, the change in organizational form is a result of a relative change in values of nonintegration and mandated internal trade. Following an increase in λ, $V^{NI}_{(EX)}(k)$ decreases but $V^{M}_{(EX)}(k)$ remains constant for every k. Thus mandated trade is optimal for all values of k for which it was optimal before the change. As investments become more cooperative we would therefore observe less nonintegration.

The intuition for this result is simple. Whenever external trade is an option, an increase in the cooperativeness of investments has an adverse effect on the units' incentives to invest, as it raises the extent to which investment improves the trading partner's outside options and bargaining position. It is interesting however to note the difference with a similar exercise in the Property Rights model. There, an increase in cooperativeness of both parties' investments has an ambiguous effect on the probability of integration (for details, see Whinston (2003)).

Finally in the case of pure cooperative investments ($\lambda = 1$), the general office faces a "time-consistency" problem opposite to the one under pure self investments. It mandates internal trade in fewer instances (fewer values of k) than it would find optimal ex-ante. Whereas ex-ante, mandating trade may be beneficial to improve investment incentives, at the interim stage investment are sunk, and the benefits of exchange autonomy loom larger. Integration therefore dominates nonintegration as whenever exchange autonomy is optimal ex-ante it is also optimal interim.

8 More Elaborate Mechanisms

In the previous discussion, the general office was restricted to choose between one of two alternatives: either mandate internal trade or grant the divisions exchange autonomy. While such policies correspond to common transfer pricing practices observed in firms, one may wish to explore the implications of allowing the general office an additional freedom on this regard, particularly as both of these policies are suboptimal and do not achieve the interim first-best.

Assume therefore that the general office can use mechanisms of the type discussed in Moore and Repullo (Moore and Repullo (1988)) for implementation in complete-information environments to elicit the realization of the commonly observed state variable θ from the divisions. Knowing θ, the general office can then implement the interim first-best by mandating internal trade if and only if $\theta \leq \theta^{FB}$. Given exchange autonomy, when $\theta > \theta^{FB}$ the divisions would both choose $m_i = \overline{m}$ and trade externally, as by lemma 6, $\theta^E \leq \theta^{FB}$. Even without resorting to such complex mechanisms, the general office can implement the interim first-best rather easily, employing the following simple mechanism: With equal probabilities, one of the divisions is delegated the right to decide whether trade is internal or external. The division then has to publicly announce its decision which is thereafter implemented. If trade is internal, the transfer price is determined by bargaining as before. The mechanism is executed at the beginning of the interim stage, prior to the orientation decisions, (m_1,m_2). If internal trade is to follow both divisions choose $m_i = \underline{m}$ and otherwise $m_i = \overline{m}$ for $i = 1, 2$.

Given ex-ante investments (s_1, s_2) , division i chooses internal trade if and only if

$$\frac{(s_i + s_{-i})\, x\left(\underline{m}, k\right)}{2} \geq s_i \gamma\left(\overline{m}, \theta\right).$$

Note that for a symmetric ex-ante choice $s_1 = s_2 = s$ the inequality above is equivalent to that for the jointly optimal rule

$$x\left(\underline{m}, k\right) \geq \gamma\left(\overline{m}, \theta\right).$$

Furthermore it can be shown that the choice of investments is symmetric in equilibrium, similarly to what was shown in section 6.2.

Note however, and this is the key point, that from an ex-ante perspective, the use of such a mechanism interim can have a detrimental effect on the incentives to invest. To see this, suppose that the general office implements the interim first-best using a mechanism that elicits θ. Then the ex-ante problem under integration is

$$\max_{b,s_1,s_2} \sum_{i=1}^{2} s_i \left[\int_{\underline{\theta}}^{\theta^{FB}} x(\underline{m},k) f(\theta)\, d\theta + \int_{\theta^{FB}}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\, d\theta \right] - \sum_{i=1}^{2} c(s_i) - 2rb^2\sigma^2$$

subject to

$$s_i \in \arg\max_s b * s \left[\int_{\underline{\theta}}^{\theta^{FB}} \frac{x(\underline{m},k)}{2} f(\theta)\, d\theta + \int_{\theta^{FB}}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\, d\theta \right] - c(s),$$

for $i \in \{1,2\}$

where terms that do not depend on s_i are dropped from the agents' incentive constraints. Compare this to the non-integration program in (8). For a given level of incentives $b_1 = b_2 = b$, the symmetric investments under integration can be seen to be $s_1 = s_2 = s^I(b)$, where

$$s^I(b) = \widehat{s}\left(b\left(\int_{\underline{\theta}}^{\theta^{FB}} \frac{x(\underline{m},k)}{2} f(\theta)\, d\theta + \int_{\theta^{FB}}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\, d\theta \right)\right)$$

and $\widehat{s}(\cdot)$ was defined above. Under non-integration $s_1 = s_2 = s^{NI}(b)$, where

$$s^{NI}(b) = \widehat{s}\left(b\left(\int_{\underline{\theta}}^{\theta^{E}} h(\widetilde{m}(\theta),\theta,k) f(\theta)\, d\theta + \int_{\theta^{E}}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\, d\theta \right)\right).$$

Now for all θ, $\frac{x(\underline{m},k)}{2} \leq h(\underline{m},\theta,k) \leq h(\widetilde{m}(\theta),\theta,k)$. And for $\theta \geq \theta^E$, $\gamma(\overline{m},\theta) \geq h(\widetilde{m}(\theta),\theta,k)$ by definition. Therefore for every level of incentives, b, investments are higher under non-integration than under integration. Integration does have the advantage of incentives coordination and here also of efficiency over the orientation decisions.

While we would not attempt to directly compare the solutions to the two programs here, we note a couple of points. First, an increase in the opportunity cost of market orientation, k, clearly favors integration here, as it lowers the value of a non-integrated organization but has no effect on that of an integrated one. Second, non-integration can indeed be optimal. Consider the case of costless market orientation: $x(m,\underline{k}) = 1$, for all m. In that case $\widetilde{m}(\theta) = \overline{m}$ for all θ and $\theta^E = \theta^{FB}$. The two programs can therefore be written as follows: Under integration

$$\max_b \; s^I(b) \left[\int_{\underline{\theta}}^{\theta^{FB}} x(\overline{m},k) f(\theta)\, d\theta + \int_{\theta^{FB}}^{\overline{\theta}} \gamma(\overline{m},\theta) f(\theta)\, d\theta \right] - c\left(s^I(b)\right) - rb^2\sigma^2$$

(the constant of two multiplying the objective is dropped) and under non-integration

$$\max_{b}\ s^{NI}(b)\left[\int_{\underline{\theta}}^{\theta^{FB}}\frac{x(\overline{m},k)+\gamma(\overline{m},\theta)}{2}f(\theta)\,d\theta+\int_{\theta^{FB}}^{\overline{\theta}}\gamma(\overline{m},\theta)\,f(\theta)\,d\theta\right]-c\left(s^{NI}(b)\right)-rb^{2}\sigma^{2}$$

We have argued that $s^{NI}(b) > s^{I}(b)$ above. The term in square brackets is bigger under integration, but the difference can be made arbitrarily small for particular choice of distribution $f(\cdot)$. For example consider a distribution with the following properties:

$$\Pr\left(\theta\in\left[\theta^{FB}-\mu,\ \theta^{FB}\right]\right)=1-\varepsilon$$

for some $\mu > 0$ and $\varepsilon > 0$ small. Provided that γ_{θ} is bounded, for every $\delta > 0$, there are some μ, ε such that the difference between the two terms is no larger than δ. Under this conditions, as incentives are more "productive" under non-integration, the non-integrated owners would set more powered incentives than the general office of an integrated firm. Consequentially, the joint surplus from integration under non-integration would be larger as well.[17]

While costless market orientation is an extreme case (in fact the general office would do equally well interim to allow exchange autonomy), a continuity argument implies that this holds true also for "close by" cases where market orientation bears a small price (in which case the general office would strictly favor using the mechanism interim).

To summarize, the introduction of more complex mechanisms that may be played in the interim stage, does not change the essence of the results. Non-integration is viable as a "commitment device", guaranteeing the incentives of managers to make ex-ante investments, and is more likely to be observed the less costly the distortion of the orientation decision is. As the "conventional" transfer pricing policies are no longer used here, the more elaborate results obtained in section 6.5 regarding the vertical structure of firms are no longer relevant in this setup. The results of this section highlights once again the insights established in section 5 on the value of "hierarchial ignorance" in organizations. The general office may be worse off being able to learn the private information held by the division managers, as by doing so it limits their real authority which, due to the contractual incompleteness, is essential for their incentives to invest.

[17] In fact the joint surplus under integration when the use of such mechanisms is feasible can be lower than in the case where only two transfer policies are considered. Surplus is clearly higher in those cases where trade was mandated ($k > k_1$), but may be lower for cases where exchange autonomy is granted ($k \leq k_1$). The argument follows similar lines to those used above.

9 Conclusions

In this paper, I have discussed the determinants of vertical organization in a framework that accounts for some important characteristics of real-world firms: First, operative decisions are in the hands of middle managers who are rewarded with incentives contracts based on the results of their units. Second, managers' decisions are subject to approval and intervention by the top management of the firm, and third managers are better informed regarding the affairs of their divisions than their superiors in the firm's hierarchy. I demonstrated that a key factor explaining whether a trade occurs within the firm or on the market is whether the general office of an integrated firm can credibly delegate the choice of trading partners to the managers of the trading divisions. Whenever such delegation is sustainable, integration of the trade is beneficial. I show that this is satisfied not only under the assumption of full commitment power, but also, interestingly, if the general office has no commitment power at all. In the latter case, asymmetric information within the firm's hierarchy on the values of different trading opportunities implies that at the time of trade, the general office finds it optimal to let the divisions choose trading partners, so that profitable opportunities for trading externally are not foregone. An explanation of the boundaries of the firm emerges only when we assume that the general office retains some limited commitment power. Specifically, in the context examined, I have shown that the general office of the firm faces a "time consistency" problem. It tends to mandate internal trades in more instances than would have been optimal if full commitment was possible. This has an adverse effect on the investments taken by the divisions' managers. Whenever such inconsistency arises, it may be optimal to have the trade conducted between non-integrated, independent parties.

A Appendix

A.1 Proofs

Proof of Proposition 1: Assume without loss of generality that $s_1 > s_2$. The proof builds on the following three lemmas.

Lemma 9 *$\exists\widehat{\theta}(s_1, s_2)$ such that $(\overline{m}, \overline{m})$ are chosen in equilibrium if and only if $\theta \geq \widehat{\theta}(s_1, s_2)$.*

Proof. The proof proceeds in several steps:

1. To check if $(\overline{m}, \overline{m})$ is an equilibrium, only deviations to $m_i = \widetilde{m}(\theta)$ need to be considered.

   For each unit $i \in \{1, 2\}$, a deviation to $m_i < \overline{m}$ can only be profitable if trade in the continuation equilibrium is internal. A deviation to $\widetilde{m}(\theta)$ is then optimal if it leads to internal trade. Suppose then that the optimal deviation that induces internal trade is to $\widehat{m}_i < \widetilde{m}(\theta)$. Optimality of $\widehat{m}_i$ implies that the total gains from internal and external trade are equal:

$$s_i x(\widehat{m}_i, k) + s_j x(\overline{m}, k) = s_i \gamma(\widehat{m}_i, \theta) + s_j \gamma(\overline{m}, \theta),$$

   or else, as $h(m, \theta)$ is increasing in m for $m \leq \widetilde{m}(\theta)$, trade would also be internal for a slightly higher m_i, and i's profits would be larger. But in this case, unit i's profits equals its disagreement payoff $s_i \gamma(\widehat{m}_i, \theta)$, which is smaller than $s_i \gamma(\overline{m}, \theta)$.

2. For each $i \in \{1, 2\}$, there exists a $\widehat{\theta}(s_i, s_j) \in (\underline{\theta}, \overline{\theta})$, such that a profitable deviation by unit i from $(\overline{m}, \overline{m})$ exists if and only if $\theta < \widehat{\theta}(s_i, s_j)$.

   A deviation by unit i is profitable given θ if

$$s_i \gamma(\overline{m}, \theta) < \frac{s_i x(\widetilde{m}(\theta), k) + s_i \gamma(\widetilde{m}(\theta), \theta) + s_j x(\overline{m}, k) - s_j \gamma(\overline{m}, \theta)}{2}. \qquad (18)$$

   Rearranging,

$$(s_i + s_j)\gamma(\overline{m}, \theta) < s_i [x(\widetilde{m}(\theta), k) + \gamma(\widetilde{m}(\theta), \theta) - \gamma(\overline{m}, \theta)] + s_j x(\overline{m}, k). \qquad (19)$$

   The term on the left is increasing in θ, as $\gamma_\theta \geq 0$. The derivative with respect to θ of right-hand side term is, by the envelope theorem,

$$\frac{d[x(\widetilde{m}(\theta), k) + \gamma(\widetilde{m}(\theta), \theta) - \gamma(\overline{m}, \theta)]}{d\theta} = \gamma_\theta(\widetilde{m}(\theta), \theta) - \gamma_\theta(\overline{m}, \theta) \leq 0,$$

since $\gamma_{m\theta} > 0$ and $\widetilde{m}(\theta) < \overline{m}$. It can also be verified that given assumption 2, $\forall (s_1, s_2)$ and $\forall i \in \{1, 2\}$, there exists a profitable deviation for $\theta = \underline{\theta}$ and there is no profitable deviation for $\theta = \overline{\theta}$. A unique cutoff $\widehat{\theta}(s_i, s_j) \in (\underline{\theta}, \overline{\theta})$ thus exists, such that a deviation by $i \in \{1, 2\}$ is profitable if and only if $\theta < \widehat{\theta}(s_i, s_j)$.

3. If $s_1 > s_2$ then $\widehat{\theta}(s_1, s_2) > \widehat{\theta}(s_2, s_1)$.

   First note that by definition of $\widetilde{m}(\theta)$ in(5),

$$x(\widetilde{m}(\theta), k) + \gamma(\widetilde{m}(\theta), \theta) - \gamma(\overline{m}, \theta) > x(\overline{m}, k).$$

   Therefore if $s_1 > s_2$ then

$$\begin{aligned} & s_1 [x(\widetilde{m}(\theta), k) + \gamma(\widetilde{m}(\theta), \theta) - \gamma(\overline{m}, \theta)] + s_2 x(\overline{m}, k) \\ > \ & s_2 [x(\widetilde{m}(\theta), k) + \gamma(\widetilde{m}(\theta), \theta) - \gamma(\overline{m}, \theta)] + s_1 x(\overline{m}, k). \end{aligned}$$

   Comparing condition (19) above for $i = 1$ and $i = 2$ then, the left-hand side term is equal while the right-hand side is larger for $i = 1$. As the right-hand side term was shown to be decreasing in θ, this implies that $\widehat{\theta}(s_1, s_2) > \widehat{\theta}(s_2, s_1)$.

$(\overline{m}, \overline{m})$ is therefore an equilibrium if and only if $\theta \geq \widehat{\theta}(s_1, s_2)$. ■

Lemma 10 $\exists \widetilde{\theta}(s_1, s_2)$ *such that* $(\widetilde{m}(\theta), \widetilde{m}(\theta))$ *is an equilibrium if and only if* $\theta < \widetilde{\theta}(s_1, s_2)$.

Proof. A deviation from $(\widetilde{m}(\theta), \widetilde{m}(\theta))$ by unit i is profitable given (s_1, s_2, θ) if

$$s_i \gamma(\overline{m}, \theta) > \frac{s_i x(\widetilde{m}(\theta), k) + s_i \gamma(\widetilde{m}(\theta), \theta) + s_j x(\widetilde{m}(\theta), k) - s_j \gamma(\widetilde{m}(\theta), \theta)}{2}. \tag{20}$$

The net gains from deviation are increasing in θ. To see this, rearrange the condition above as follows:

$$s_i [2\gamma(\overline{m}, \theta) - x(\widetilde{m}(\theta), k) - \gamma(\widetilde{m}(\theta), \theta)] - s_j [x(\widetilde{m}(\theta), k) - \gamma(\widetilde{m}(\theta), \theta)] > 0,$$

and take a derivative of the left-hand side term with respect to θ. This equals

$$s_i [2\gamma_\theta(\overline{m}, \theta) - \gamma_\theta(\widetilde{m}(\theta), \theta)] - s_i (x_m + \gamma_m) \frac{\partial \widetilde{m}(\theta)}{\partial \theta} + s_j \gamma_\theta(\widetilde{m}(\theta), \theta) - s_j (x_m - \gamma_m) \frac{\partial \widetilde{m}(\theta)}{\partial \theta}.$$

The second term above equals zero by the definition of $\widetilde{m}(\theta)$. All other terms are positive. Thus for $i \in \{1, 2\}$, a deviation is profitable if and only if $\theta \geq \widetilde{\theta}(s_i, s_j)$ for some $\widetilde{\theta}(s_i, s_j) \in (\underline{\theta}, \overline{\theta})$.

To compare $\widetilde{\theta}(s_1, s_2)$ and $\widetilde{\theta}(s_2, s_1)$, rewrite the condition for profitable deviation as follows:

$$s_i\left[2\gamma(\overline{m}, \theta) - \gamma\left(\widetilde{m}(\theta), \theta\right)\right] + s_j\gamma(\widetilde{m}(\theta), \theta) > (s_i + s_j)\, x\left(\widetilde{m}(\theta), k\right). \tag{21}$$

As $2\gamma(\overline{m}, \theta) - \gamma\left(\widetilde{m}(\theta), \theta\right) \geq \gamma(\overline{m}, \theta) \geq \gamma\left(\widetilde{m}(\theta), \theta\right)$,

$$\begin{aligned}
& s_1\left[2\gamma(\overline{m}, \theta) - \gamma\left(\widetilde{m}(\theta), \theta\right)\right] + s_2\gamma(\widetilde{m}(\theta), \theta) \\
> \;& s_2\left[2\gamma(\overline{m}, \theta) - \gamma\left(\widetilde{m}(\theta), \theta\right)\right] + s_1\gamma(\widetilde{m}(\theta), \theta),
\end{aligned}$$

The term on the right in (21) is equal for $i = 1, 2$, whereas the term on the left is larger for $i = 1$. Therefore whenever unit 2 has a profitable deviation so does unit 1. Hence $\widetilde{\theta}(s_1, s_2) < \widetilde{\theta}(s_2, s_1)$, and $(\widetilde{m}(\theta), \widetilde{m}(\theta))$ is an equilibrium if and only if $\theta < \widetilde{\theta}(s_1, s_2)$. ■

Below, the abbreviations $\widehat{\theta}_1 \equiv \widehat{\theta}(s_1, s_2)$ and $\widetilde{\theta}_1 \equiv \widetilde{\theta}(s_1, s_2)$ would be used at times .

Lemma 11 $\widehat{\theta}(s_1, s_2) \leq \widetilde{\theta}(s_1, s_2)$.

Proof. First, for every θ,

$$\begin{aligned}
& \frac{s_i x\left(\widetilde{m}(\theta), k\right) + s_i\gamma\left(\widetilde{m}(\theta), \theta\right) + s_j x\left(\widetilde{m}(\theta), k\right) - s_j\gamma(\widetilde{m}(\theta), \theta)}{2} \\
\geq \;& \frac{s_i x\left(\widetilde{m}(\theta), k\right) + s_i\gamma\left(\widetilde{m}(\theta), \theta\right) + s_j x\left(\overline{m}, k\right) - s_j\gamma(\overline{m}, \theta)}{2}.
\end{aligned}$$

Suppose $\widehat{\theta}_1 > \widetilde{\theta}_1$ so that $s_i\gamma(\overline{m}, \widehat{\theta}_1) > s_i\gamma(\overline{m}, \widetilde{\theta}_1)$. As (18) holds with equality for $\widehat{\theta}_1$ and (20) holds in equality for $\widetilde{\theta}_1$, we can write

$$\begin{aligned}
& \frac{s_i x\left(\widetilde{m}\left(\widehat{\theta}_1\right), k\right) + s_i\gamma\left(\widetilde{m}\left(\widehat{\theta}_1\right), \widehat{\theta}_1\right) + s_j x\left(\overline{m}, k\right) - s_j\gamma(\overline{m}, \widehat{\theta}_1)}{2} \\
> \;& \frac{s_i x\left(\widetilde{m}\left(\widetilde{\theta}_1\right), k\right) + s_i\gamma\left(\widetilde{m}\left(\widetilde{\theta}_1\right), \widetilde{\theta}_1\right) + s_j x\left(\widetilde{m}\left(\widetilde{\theta}_1\right), k\right) - s_j\gamma(\widetilde{m}\left(\widetilde{\theta}_1\right), \widetilde{\theta}_1)}{2} \\
\geq \;& \frac{s_i x\left(\widetilde{m}\left(\widetilde{\theta}_1\right), k\right) + s_i\gamma\left(\widetilde{m}\left(\widetilde{\theta}_1\right), \widetilde{\theta}_1\right) + s_j x\left(\overline{m}, k\right) - s_j\gamma(\overline{m}, \widetilde{\theta}_1)}{2},
\end{aligned}$$

where the second inequality follows from the first claim above. But as the right-hand side of (18) is decreasing in θ, a contradiction. ■

The proof of Proposition 1 then follows immediately from the three lemmas above.

A.2 Cooperative investments

In here, I demonstrate that the analysis of this paper can be extended to cover cooperative investments. I therefore posit (assumption 4) that the respective values of trade are as follows:

$$\begin{aligned}\Phi\left(s_i, s_j, m_i\right) &= \left[\left(1-\lambda\right)s_i + \lambda s_j\right] * x\left(m_i, k\right),\\ \Gamma\left(s_i, s_j, m_i, \theta\right) &= \left[\left(1-\lambda\right)s_i + \lambda s_j\right] * \gamma\left(m_i, \theta\right).\end{aligned}$$

Consider first the choice of market orientation (m_1, m_2). Unit $i \in \{1,2\}$ is "in charge" of an investment of value

$$\widetilde{s_i} = \left(1-\lambda\right)s_i + \lambda s_j.$$

The analysis of the subgame starting with the choice of market orientation then follows exactly that of subsection 6.1 for the self-investments case, where s_i is replaced by $\widetilde{s_i}$. The interim decision on the transfer pricing policy is determined by the sign of

$$\Delta M_{(IN)}\left(k\right) = \left(1-b\right) * \widetilde{s}\left(b\right)\left[\int_{\underline{\theta}}^{\theta^E(k)} x\left(\widetilde{m}\left(\theta,k\right),k\right)f\left(\theta\right)d\theta + \int_{\theta^E(k)}^{\overline{\theta}} \gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta\right],$$

which does not depend on the equilibrium investments $s_1 = s_2 = \widetilde{s}\left(b\right)$.

Consider next the choice of investments in the ex-ante stage. Given incentives b_i, unit i's investment s_i under non-integration and exchange autonomy solves:

$$\max_s b_i * \left[\begin{array}{l}\displaystyle\int_{\underline{\theta}}^{\theta^E}\left\{\begin{array}{l} s\frac{(1-\lambda)(x(\widetilde{m}(\theta),k)+\gamma(\widetilde{m}(\theta),\theta))+\lambda(x(\widetilde{m}(\theta),k)-\gamma(\widetilde{m}(\theta),\theta))}{2}+\\ s_j\frac{(1-\lambda)(x(\widetilde{m}(\theta),k)-\gamma(\widetilde{m}(\theta),\theta))+\lambda(x(\widetilde{m}(\theta),k)+\gamma(\widetilde{m}(\theta),\theta))}{2}\end{array}\right\}f\left(\theta\right)d\theta+\\ \displaystyle\int_{\theta^E}^{\overline{\theta}} s*\left(1-\lambda\right)\gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta\end{array}\right] - c\left(s\right).$$

And therefore

$$s_i = \widehat{s}\left(b_i\left[\int_{\underline{\theta}}^{\theta^E}\frac{x\left(\widetilde{m}\left(\theta\right),k\right)+\left(1-2\lambda\right)\gamma\left(\widetilde{m}\left(\theta\right),\theta\right)}{2}f\left(\theta\right)d\theta + \int_{\theta^E}^{\overline{\theta}}\left(1-\lambda\right)\gamma\left(\overline{m},\theta\right)f\left(\theta\right)d\theta\right]\right),$$

where $\widehat{s}\left(\cdot\right)$ is defined in subsection 6.2. Under mandated trade, s_i is a solution to

$$\max_s b_i * s\int_{\underline{\theta}}^{\overline{\theta}}\frac{\left(1-\lambda\right)x\left(\underline{m},k\right)+\lambda x\left(\underline{m},k\right)}{2}f\left(\theta\right)d\theta - c\left(s\right).$$

Therefore

$$s_i = \widehat{s}\left(b_i\frac{x\left(\underline{m},k\right)}{2}\right).$$

As $\widehat{s}' > 0$, an increase in λ, other things being equal, has no effect on investment under mandated trade, but lowers investment under exchange autonomy and non-integration.

Finally, we argue that the proof of proposition 2 is still valid for the case of cooperative investments, provided that λ is small enough. The only part of the proof that has to be verified is the comparison between s^M and $s^{\overline{M}}$, the symmetric equilibrium investments that would have been chosen under exchange autonomy, given incentives of b^M:

$$\begin{aligned} s^M &= \widehat{s}\left(b^M \int_{\underline{\theta}}^{\overline{\theta}} \frac{x\left(\underline{m}, k\right)}{2} f\left(\theta\right) d\theta\right), \\ s^{\overline{M}} &= \widehat{s}\left(b^M \left[\int_{\underline{\theta}}^{\theta^E} \frac{x\left(\widetilde{m}\left(\theta\right), k\right) + \left(1 - 2\lambda\right)\gamma\left(\widetilde{m}\left(\theta\right), \theta\right)}{2} f\left(\theta\right) d\theta + \int_{\theta^E}^{\overline{\theta}} \left(1 - \lambda\right)\gamma\left(\overline{m}, \theta\right) f\left(\theta\right) d\theta\right]\right). \end{aligned}$$

For $\lambda = 0$, $s^{\overline{M}} > s^M$, following the argument made in proposition 2. Therefore as $s^{\overline{M}}$ is continuous in λ, there exists a $\overline{\lambda}$ such that $s^{\overline{M}} > s^M$ for $\lambda < \overline{\lambda}$. The rest of the proposition then holds under this restriction.

References

Aghion, P. and Tirole, J.: 1997, Formal and real authority in organizations, *Journal of Political Economy* **105**(1), 1–29.

Baker, G. P.: 1992, Incentive contracts and performance measurement, *Journal of Political Economy* **100**(3), 598–614.

Bolton, P. and Scharfstein, D. S.: 1998, Corporate finance, the theory of the firm, and organizations, *Journal Fo Economic Perspectives* **12**(4), 95–114.

Chandler, A.: 1962, *Strategy and Structure: Chapters in the History of the Industrial Enterprise*, M.I.T. Press, Cambridge, Mass.

Che, Y.-K. and Hausch, D. B.: 1999, Cooperative investments and the value of contracting, *American Economic Review* **89**(1), 125–147.

Coase, R.: 1937, The nature of the firm, *Economica* **4**(16), 386–405.

Eccles, R. G.: 1985, *The Transfer Pricing Policy. A Theory for Practice*, Lexington Books.

Grossman, S. J. and Hart, O.: 1986, The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy* **94**, 691–719.

Hart, O.: 1995, *Firms, Contracts and Finacial Structure*, Calderon Press, Oxford.

Hart, O. and Moore, J.: 1990, Property rights and the nature of the firm, *Journal of Political Economy* **98**, 1119–1158.

Holmstrom, B. and Milgrom, P.: 1987, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica* **55**(2), 303–328.

Holmstrom, B. and Milgrom, P.: 1991, Multitask principal-agent analyses: Incentive contracts, asset ownership and job design, *Journal of Law, Economics, and Organization* **7**, 24–52.

Holmstrom, B. and Tirole, J.: 1991, Transfer pricing and organizational design, *Journal of Law and Economic Organization* **7**(2), 201–228.

Klein, B., Crawford, R. and Alchian, A.: 1978, Vertical integration, appropriable rents, and the competitive contracting process, *Journal of Law and Economics* **21**, 297–326.

Moore, J. and Repullo, R.: 1988, Subgame perfect implementation, *Econometrica* **56**, 1191–1220.

Osborne, M. J. and Rubinstein, A.: 1990, *Bargaining and Markets*, Academic Press, San Diego, CA.

Whinston, M. D.: 2003, On the transaction costs determinants of vertical integration, *Journal of Law, Economics, and Organization* **19**(1), 1–23.

Williamson, O. E.: 1975, *Markets and Hierarchies: Analysis and Antitrust Implications*, The Free Press, New York.

Williamson, O. E.: 1985, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, Free Press, New York.

Williamson, O. E.: 2000, The new institutional economics: Taking stock, looking ahead, *Journal of Economic Literature* **38**(3), 595–613.