Foundations and Trends® in
Accounting
Vol. 3, No. 4 (2008) 223–313

DOI: 10.1561/1400000013

Internal Pricing

By Tim Baldenius

Contents

Foundations and Trends® in
Accounting
Vol. 3, No. 4 (2008) 223–313

DOI: 10.1561/1400000013

Internal Pricing*

Tim Baldenius

Graduate School of Business, Columbia University, New York, NY 10027, USA, tb171@columbia.edu

Abstract

This monograph focuses on the use of incomplete contracting models to study transfer pricing. Intrafirm pricing mechanisms affect division managers' incentives to trade intermediate products and to undertake relationship-specific investments so as to increase the gains from trade. Letting managers negotiate over the transaction is known to cause holdup (underinvestment) problems. Yet, in the absence of external markets, negotiations frequently outperform cost-based mechanisms, because negotiations aggregate cost and revenue information more efficiently into prices. This result is established in a symmetric information setting and confirmed, with some qualification, for bargaining under incomplete information. In the latter case, trading and investment efficiency can be improved by adding non-financial performance measures to a divisional performance measurement system. When the intermediate product can also be sold in an imperfectly competitive external market, internal discounts on external market prices are shown often to improve the efficiency of intrafirm trade and of upfront investments.

* The author would like to thank Brian Mittendorf (the reviewer), Neale O'Connor and Stefan Reichelstein for their helpful comments.

1

Introduction

Evaluating the performance of divisional managers is a key issue in decentralized firms. Compensation for these managers is based for the most part on divisional profit metrics. Disaggregating firmwide into divisional performance, however, is often complicated by intrafirm trade. Ultimately one division will book the revenues when a final product is sold, yet the value thus created may be determined to a large extent by intermediate goods and services provided by many divisions. Without proper internal pricing schemes to account for these resources, the contributing divisions will have little or no incentives to render any (high quality) inputs. This monograph proposes an incomplete contracting model to evaluate the strengths and weaknesses of commonly used pricing schemes.

The Arrow–Debreu mechanism efficiently aggregates valuations of market participants into market-clearing prices. Yet, two of the underlying assumptions — atomistic buyers and sellers, and homogeneous goods — are generally violated for intrafirm transactions. The reasons for firms to vertically integrate often revolve around specialized (non-commoditized) inputs and less than competitive markets to source these inputs from.[1] Purchasing such inputs from external sources requires paying monopoly rents or risking information spillage in case of proprietary technologies. Within vertically integrated firms, there are typically a small number of sellers and buyers for any given input; in fact, the extreme case of bilateral monopoly is common. It is well-known that a laissez-faire mechanism performs less well in such "thin" markets (Myerson and Satterthwaite, 1983).

In his classic study, Hirshleifer (1956) has shown that pricing intermediate products at marginal cost — broadly defined to include opportunity costs if the intermediate good is traded externally — achieves full efficiency. While conceptually important, this result is of limited practical value as the firm's central office (hereafter, HQ) typically lacks the information necessary to determine marginal cost at the product level. Commonly used transfer pricing mechanisms therefore are more decentralized in nature and fall into three broad categories: negotiated, cost-based, and market-based.[2] HQ's role is to choose a mechanism from among these candidates and, possibly, to make further design choices, e.g., under cost-based pricing whether to use standard or actual and full or variable costs, and whether to add a profit markup; whether to make internal adjustments under market-based pricing; and whether to force internal sourcing ("exclusivity clauses") under negotiations.

At an early stage, division managers can often increase the gains from trade by investing in fixed assets. If these assets are relationship-specific (they cannot be redeployed easily), and contracts are incomplete (fully contingent contracts are infeasible), a "holdup" problem arises if prices are negotiated *ex post* (Williamson, 1985). The investing party anticipates that in the process of negotiation it will have to split

[1] These reasons include: to increase control over inputs, to capture upstream/downstream profits, to increase barriers to entry, and to facilitate investments in specific assets, e.g., Joskow (2005).

[2] For example, Horngren et al. (2007). For *multinational* enterprises (where taxation plays a crucial role), Ernst and Young (2008) report that market-based transfer pricing is most commonly used for tangible and intangible goods and for financing, whereas the cost-plus method is most frequently used to price services rendered internally. Note that negotiated transfer pricing is not admissible for tax reporting. In practice this categorization can be fuzzy at times. For instance, cost-based (market-based) transfer prices are often subject to markups (discounts), and determining those adjustments sometimes involves an element of negotiations.

the surplus with the other party, and therefore underinvests. Thus, HQ has to provide managers with appropriate *ex post* trading and *ex ante* investment incentives. To that end, it is useful to think of internal pricing schemes as mechanisms that determine: (i) the expected gains from trade for given investments and (ii) an allocation of bargaining power that determines the split of this surplus. It is these two factors combined that determine the divisions' upfront investment incentives.

In the absence of reliable and competitive market prices for intermediate goods, the most common internal pricing methods are negotiated and standard cost-based, as actual costs are often hard to verify for HQ. The main problem associated with standard cost pricing is that cost standards are often set (or influenced) by the upstream division, and thus inflated.[3] This in turn gives rise to a double-marginalization problem (Tirole, 1988). The main disadvantages of negotiations are haggling costs and the holdup problem. Using the above taxonomy, standard-cost pricing (with seller reporting discretion) confers more bargaining power to the selling division, yet the expected overall surplus is suboptimal. Ignoring haggling costs at first by assuming that the managers bargain under symmetric information, the double marginalization problem under standard-cost pricing is shown, in many cases, to be so severe as to make negotiations the preferred regime, despite the attendant holdup problem. Standard-cost pricing is particularly ill-suited if buyer investments are important, due to a holdup problem of its own: the seller will opportunistically submit an even higher "cost" quote upon observing that the buyer has invested in the transaction and thereby increased his willingness-to-pay. Thus, the buyer will be reluctant to invest under this pricing regime because the overall expected surplus will be smaller *and* he has basically no bargaining power.

Market-based transfer pricing is a viable mechanism if goods or services comparable to those transferred internally are traded in external markets. Firms often combine market-based pricing with *intracompany discounts*. This practice is usually rationalized with cost differences such as reduced transportation of bad debt costs. With imperfectly

[3] In Vaysman (1996) markups over cost arise endogenously as a result of informational rents earned by the upstream division.

competitive external markets, however, a more fundamental reason for internal discounts exists: to mitigate double-marginalization problems. Discounts are shown to be particularly effective (and under certain condition achieve first-best) if upstream capacity is constrained, whereas they can actually reduce firmwide profit with if capacity is abundant. Yet, even then, the performance of market-based pricing (with suitable internal discounts) converges to first-best as the importance of the external market relative to internal transfers becomes large.

Allowing for specific investments, it is shown that buyer investments under market-based pricing suffer from a holdup problem similar to that under standard cost-based pricing: the seller will raise the external market price (and thereby the transfer price) in response to the buyer investing. The holdup problem is alleviated, however, by the fact that the seller loses external business, as he is forced to raise the external price so as to capture some of the downstream rents. Even if imperfectly competitive, external markets thus provide some protection for investments for otherwise powerless internal buyers. Specific investments are shown to add further impetus for internal discounts; in particular they foster investments downstream.

The symmetric-but-nonverifiable information model has become the standard "work horse" in the incomplete contracting literature, mostly for reasons of analytic tractability, as generalized Nash bargaining results in full *ex post* efficiency (by the Coase Theorem). However, the assumption of symmetric information across divisions is unrealistic and, in the context of intrafirm transactions, conceptually problematic: if the divisions know each others' costs and revenues, why doesn't HQ? After all, it is HQ that designs the accounting system. Therefore, I will revisit the performance comparison of negotiated and standard cost-based transfer pricing in a model variant where each division has some private information at the bargaining stage. Negotiations (via sealed-bid bargaining) then also fail to realize all gains from trade *ex post*, i.e., this model variant picks up haggling costs.[4]

[4] This is in line with Williamson (2000) and Hart and Moore (2008) who stress the importance of *ex post* frictions for the optimal choice of governance mechanisms. Myerson and Satterthwaite (1983) employ a general mechanism-design approach to demonstrate that

Even with asymmetric information at the transaction stage, however, negotiations often generate greater firmwide *ex post* gains from trade than standard-cost pricing. Overall, the model provides the following prescriptions: HQ should choose the pricing mechanism that allocates more bargaining power to the party that: (i) has more private information (to minimize trade distortions) and (ii) more substantial investment opportunities (to minimize aggregate holdup problems), all else equal. A confounding factor to (i) and (ii), however, is that holdup problems tend to diminish with private information on the part of the investing party. To illustrate, suppose Manager X has some (or all) bargaining power but only Manager Y has an investment opportunity. If X is uncertain about Y's reservation price, then X will have to bid (or price) more "carefully" or otherwise risk foregoing the transaction altogether. Private information thus provides partial protection from holdup.

The sealed-bid mechanism for modeling bargaining under asymmetric information can be fruitfully employed also to study how compensation contracts affect divisional investment and trading incentives. Since it is rooted in non-cooperative game theory, one can trace explicitly how the managers' bargaining strategies are affected by the incentive contracts. Specifically, it is shown that HQ can alleviate trade and investment inefficiencies by adding non-financial performance measures (NPMs) to contracts that otherwise depend only on divisional profits. This is in line Bouwens and van Lent (2007) who have shown empirically that firms' use of NPMs is increasing in the extent of cross-divisional cooperation.

A key theme running through this paper is that HQ plays a rather limited role in mediating individual transactions. This captures the stylized empirical fact that in most divisionalized firms HQ designs the broad "rules of the game" by choosing a pricing mechanism and compensation contracts, but usually does not get involved in pricing on a product-by-product basis.

bargaining under asymmetric information (given budget balancing) will always result in inefficiency.

In the following, I make no attempt at an exhaustive survey of the transfer pricing literature but focus on select articles that have studied commonly used pricing schemes using incomplete contracting models: Baldenius et al. (1999), Baldenius (2000), Sahay (2003), Baldenius and Reichelstein (2006), Baiman and Baldenius (2009), and Pfeiffer et al. (2009). Most importantly, this paper ignores studies concerned with optimal mechanism design, "strategic" transfer pricing in oligopoly, or taxation.[5] Moreover, a number of restrictive assumptions will be made. First, the transaction is one-shot.[6] Second, I ignore firmwide profit sharing and instead (for the most part) assume compensation is tied solely to divisional income, which appears descriptive; managers are risk neutral.[7] Lastly, upfront contracts cannot be written due to lack of *ex ante* describability of the product to be traded later. Again, this appears descriptive as formal upfront contracts between divisions are rarely observed.[8]

The remainder of this monograph is organized as follows. Section 2 develops the basic symmetric information model to compare the performance of cost-based and negotiated pricing in the absence of external input markets. Section 3 considers market-based pricing and the role of internal price adjustments; it ignores investments and focuses solely on

[5] For selective references for the omitted topics see Harris et al. (1982), Wagenhofer (1994), Vaysman (1996, 1998) on the mechanism-design approach to internal pricing; Alles and Datar (1998), Gox (2000), Narayanan and Smith (2000), Shor and Chen (2009) on strategic transfer pricing; and Halperin and Srinidhi (1991), Baldenius et al. (2004), Hyde and Choe (2005), Johnson (2006) on transfer pricing and taxes. See Gox and Schiller (2007) for a comprehensive survey.

[6] With infinitely repeated transactions and patient players, equilibria can be supported in which the holdup problem essentially disappears, by the folk theorem. Che and Sakovics (2004) show that even in a one-shot game, but with multiple rounds of "staged" investments, high investments can be sustained in equilibrium despite the fact that holdup problems remain. Dutta and Reichelstein (2009) study holdup problems in connection with investments in long-lived assets.

[7] Studying the tradeoff between risk premia and holdup, Holmstrom and Tirole (1991) and Anctil and Dutta (1999) derive optimal weights on divisional and firmwide performance measures.

[8] Rogerson (1992), Edlin and Reichelstein (1995), Wielenberg (2000), Bockem and Schiller (2004), among others, have shown that non-contingent upfront contracts, to be renegotiated once uncertainty is resolved, can resolve the holdup problem. See also Maskin and Tirole (1999).

trading incentives. Section 4 adds investments to the model of Section 3 and shows that investment opportunities further strengthen the case for internal adjustments. Section 5 reconsiders the initial analysis of Section 2 for the case of asymmetrically informed divisional managers. Section 6 concludes.

2

Intrafirm Trade Absent External Markets

As argued in the Introduction, one of the key reasons for firms to start making their own inputs is lack of (competitive) markets to source these inputs from. In this section — as well as in Section 5 — I take this notion to the extreme by ignoring external input markets altogether. Absent market prices as a benchmark, internal prices will be either negotiated or based on upstream production cost. This section — based on Baldenius et al. (1999, henceforth BRS) — compares these two classes of mechanisms. It lays the foundation for subsequent sections, which will extend this model in several directions.

2.1 Model with Symmetrically Informed Managers

Consider a firm consisting of two divisions and headquarters (HQ). The firm is decentralized in that HQ's role is confined to choosing a transfer pricing mechanism. The divisions operate in separate markets except for an intermediate good that is made by Division 1 (the supplier or seller) and transferred to Division 2 (the buyer). Division 2 further processes this input and sells a final product to the external market. No close substitutes for the intermediate good are traded in external markets, i.e., Division 1 cannot sell the good externally and Division 2 cannot buy it externally. Thus, the divisions face each other in a bilateral monopoly situation.

Each division can make specific investments to increase the gains from trade. The supplying division can reduce its variable production cost, e.g., by installing more efficient equipment. The buying division can increase its net revenue, e.g., by investing in reducing the cost of further processing the intermediate good or in sales promotions for the final product. The investments have to be sunk at a point in time when final costs and revenues are still uncertain. This uncertainty is captured by the (possibly multidimensional) random state variable θ. Divisional managers are assumed risk-neutral, and they maximize their own division's income, e.g., because their compensation is a linear function of divisional profit and they do not incur any personal disutility (or empire benefits) from the transaction at hand. Throughout this monograph I will ignore firmwide profit sharing; Section 5.6 will look more closely at compensation contracts and their effect on trade and investments.[1] The sequence of events is summarized in Figure 2.1.

At Date 1, the seller chooses his specific investment I_1 from the interval $[0, \bar{I}_1]$ and, similarly, the buyer chooses $I_2 \in [0, \bar{I}_2]$. These investments generate divisional fixed costs of $w_1(I_1)$ for the selling division and $w_2(I_2)$ for the buying division. Contingent on θ and $I \equiv (I_1, I_2)$,

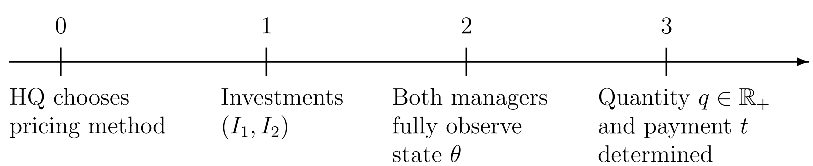

Fig. 2.1 Timeline with symmetrically informed division managers.

[1] As in most incomplete contracting models without additional agency problems, underinvestment problems could be mitigated by basing the managers' salaries on firm-wide profit. However, in most cases internal transactions constitute only "additional business" for the divisions. While unmodeled here, moral hazard problems associated with other projects keep profit sharing from attaining first-best; see Holmstrom and Tirole (1991) or Anctil and Dutta (1999). Eccles (1985), Merchant (1989), and Bushman et al. (1995) observe that divisional performance measures generally are the main drivers of division managers' compensation.

Division 1's variable costs and Division 2's revenues (net of finishing costs) from trading a quantity $q \in \mathbb{R}_+$ are given by[2]

$$C(q,\theta,I_1) \equiv [c(\theta) - I_1]q, \quad R(q,\theta,I_2) = r(q,\theta) + I_2 q. \tag{2.1}$$

That is, Division 1's variable cost function is linear, whereas Division 2's "base revenue" function, $r(q,\theta)$, can take any arbitrary shape. Note that the firm-wide returns to the divisional investments are proportional to the quantity traded. It is assumed for now that the upstream division has sufficient capacity to produce the requested q units. When adding external selling opportunities for the intermediate good, Section 3 will explicitly account for capacity constraints.[3]

State uncertainty is fully resolved (for the division managers) at Date 2, so that at Date 3 each manager knows his own valuation of the intermediate good as well as that of the respective other division. This is a strong assumption, which will be relaxed in Section 5. However, neither θ nor the investments I are verifiable to the firm's HQ, which rules out contingent transfer rules.[4] Instead, at Date 3 the transfer pricing mechanism determines the actual transfer quantity, q, and the transfer price, $t \in \mathbb{R}_+$, that Division 2 pays to Division 1.

The First-Best Solution. A natural benchmark is the *first-best solution* that would obtain if HQ were to observe θ at Date 2 and could either directly choose investments or verifiably instruct the divisions to invest a certain amount at Date 1. By backward induction, for given q, I, and θ, the firm's contribution margin is

$$M(q,\theta,I) \equiv R(q,\theta,I_2) - C(q,\theta,I_1).$$

Assume that $M(\cdot,\theta,I)$ has a unique interior maximizer, $q^*(\theta,I) > 0$, for any θ, I. Let $M^*(\theta,I) \equiv M(q^*(\theta,I),\theta,I)$ and $\pi^*(I) \equiv \mathbb{E}_\theta[M^*(\theta,I)]$

[2] Assume that $\bar{I}_2 \le c(\theta) - \bar{I}_1$ for any θ.

[3] If it were common knowledge that the upstream division has scarce capacity while only supplying an internal buyer, then the production decision would be trivial and could be efficiently centralized at very low informational requirements.

[4] One possible interpretation is that both division managers observe θ and (I_1, I_2). Alternatively, a manager may simply observe the valuation of the other division without being able to disentangle the effects of θ and I. The model could be amended to allow for Division 1 to remain uncertain about its actual cost until some date later than Date 3. At the trading stage, both divisions then would only know that the unit cost of the intermediate good is $c(\theta, I_1) + \epsilon$, where ϵ is some unbiased random variable. With risk-neutral division managers, such an extension would have no impact on the results below.

$-w_1(I_1) - w_2(I_2)$ denote firm-wide contribution margin and expected profit, respectively, conditional on efficient quantity choices (throughout this monograph, $\mathbb{E}$ is the expectation operator). First-best investments are then given by

$$I^* \equiv (I_1^*, I_2^*) \in \arg\max_I \ \pi^*(I). \tag{2.2}$$

Throughout this section, assume that for any I_j the function $\pi^*(I_i, I_j)$ is single-peaked in I_i with a unique interior maximizer $I_i^*(I_j),\ i \neq j$. By the envelope theorem, the following first-order conditions then will be necessary and sufficient for an optimum:

$$\begin{aligned} -\mathbb{E}_\theta\left[\frac{\partial}{\partial I_1} C(q^*(\theta, I^*), \theta, I_1^*)\right] - w_1'(I_1^*) \\ = \mathbb{E}_\theta[q^*(\theta, I^*)] - w_1'(I_1^*) = 0, \end{aligned} \tag{2.3}$$

$$\begin{aligned} \mathbb{E}_\theta\left[\frac{\partial}{\partial I_2} R(q^*(\theta, I^*), \theta, I_2^*)\right] - w_2'(I_2^*) \\ = \mathbb{E}_\theta[q^*(\theta, I^*)] - w'(I_2^*) = 0. \end{aligned} \tag{2.4}$$

To maximize expected firmwide profit, upstream investments should be made up to the point where the marginal investment cost equals the expected savings in variable cost. Similarly, downstream investments should be made up to the point where the marginal investment cost equals the expected marginal increase in net revenues. Given the separability properties embedded in Equation (2.1), both these marginal investment benefits equal the expected trading quantity.

The Decentralized Solution. Now consider the delegation game where operating decisions are made by division managers who take the transfer pricing regime as given. For given transfer price t and trading quantity q, divisional contribution margins are

$$M_1 = tq - C(q, \theta, I_1)$$

for the upstream division and

$$M_2 = R(q, \theta, I_2) - tq$$

for the downstream division. As stated above, Manager i aims to maximize his own division's realized profit of $M_i - w_i(I_i)$. Clearly, for given quantity q, the transfer price t is irrelevant for the firm-wide profit, as it "cancels out". Yet, t affects the division managers' payoffs and thereby their incentives. The following analysis will focus on two particularly popular transfer pricing mechanisms.

Standard-Cost Transfer Pricing. The most common criticism of standard cost transfer pricing is that standards are determined largely by a party that has a vested interest in biasing the standard, namely the supplying division. Profit centers are generally engaged in many transactions and often find ways to allocate overhead costs across products and services. To capture this feature in the model, suppose for now that the selling division quotes a "standard unit cost" of t^s, and the buying division can only choose the quantity to be traded at this transfer price. That is, the selling division effectively has monopoly power. Since in practice there may be limit for it to overstate costs, Section 2.5 will address such constraints.

The selling division anticipates that for any transfer price, t, the buyer will order a quantity $Q(t,\theta,I_2) \in \arg\max_q \{R(q,\theta,I_2) - tq\}$, where $Q(\cdot)$ is the inverse of the buyer's marginal revenue (or willingness-to-pay) curve, i.e., $Q(\cdot,\theta,I_2) \equiv R'^{-1}(\cdot,\theta,I_2)$. Thus, the seller quotes (inflated) unit costs of $t^s(\theta,I)$, such that:[5]

$$t^s(\theta,I) \in \arg\max_t \{[t - c(\theta) - I_1]Q(t,\theta_2,I_2)\}.$$

Since $t^s(\theta,I)$ will exceed the true unit cost $c(\theta) - I_1$, this scheme suffers from a *double marginalization* bias (Tirole, 1988), as $q^s(\theta,I) \equiv Q(t^s(\theta,I),\theta,I_2) \leq q^*(\theta,I)$.

Negotiated Transfer Pricing. Now suppose that at Date 3 the divisional managers bargain over the transfer to take place. The standard approach for modeling bargaining under symmetric information is the (generalized) Nash bargaining solution whereby the selling division receives a share $\eta \in [0,1]$ of the attainable contribution margin. Thus, the parties agree on the quantity $q^*(\theta,I)$ and a unit transfer price, t^n,

[5] As mentioned in Footnote 4 of this section, all results to follow would hold if the seller's cost were subject to residual uncertainty at Date 3. In that case, t^s would be a "true" standard (i.e., expected) cost, albeit still biased.

which is implicitly defined by the equations:

$$t^n q^*(\theta,I) - C(q^*(\theta,I),\theta,I_1) = \eta M^*(\theta,I) \tag{2.5}$$

$$R(q^*(\theta,I),\theta,I_2) - t^n q^*(\theta,I) = (1-\eta)M^*(\theta,I). \tag{2.6}$$

The sharing parameter η measures the seller's bargaining power, with the special case of $\eta = \frac{1}{2}$ representing the Nash bargaining solution.[6]

The firm now faces the problem of choosing a transfer pricing mechanism that simultaneously deals with two interrelated incentive problems: managers should have incentives: (i) to trade the efficient quantities of the intermediate good, *ex post*, and (ii) to undertake investments, *ex ante*.

2.2 Investment by the Supplying Division

For now I disregard investments by the buying division, i.e., $I_1 \geq \bar{I}_2 = 0$ and suppress I_2 to save on notation. The buyer's net revenue function then collapses to $r(q,\theta)$. Now turn to the performance comparison of the two transfer pricing mechanisms. While this may appear to require tedious calculations of closed-form solutions for expected profits, the preceding discussion suggests a "short-cut:" since Nash bargaining leads to *ex post* efficient transfers, while standard-cost pricing in general does not, for the latter scheme to be preferred, it has to generate better investment incentives. Put differently, a sufficient condition for negotiations to dominate standard-cost transfer pricing is that $I_1^s < I_1^n$. This follows from the assumed single-peakedness of the first-best expected profit function $\pi^*(I_i)$ along each argument, together with the fact that negotiations result in underinvestment, as shall be demonstrated now.

As argued in the Introduction, divisional investment incentives are determined by the expected firmwide gains from trade, together with the allocation of bargaining power. Under negotiations, the seller's

[6] The disagreement point here is zero for each division. This follows from the assumptions that the investments are relationship-specific (hence they cannot be redeployed to other uses) and upfront fixed-price contracts a la Edlin and Reichelstein (1995) are not feasible.

investment problem reads:

$$I_1^n \in \arg\max_{I_1} \{\eta \mathbb{E}_\theta[M^*(\theta, I_1)] - w_1(I_1)\},$$

which, by the envelope theorem, yields the necessary first-order condition

$$\begin{aligned} -\eta\left(\mathbb{E}_\theta\left[\frac{\partial}{\partial I_1} C(q^*(\theta, I_1^n), \theta, I_1^n)\right]\right) &- w_1'(I_1^n) \\ &= \eta \mathbb{E}_\theta[q^*(\theta, I_1^n)] - w_1'(I_1^n) = 0. \end{aligned} \tag{2.7}$$

By revealed preference, a comparison of Equation (2.7) with Equation (2.3) shows that the seller will underinvest relative to first-best for any $\eta < 1$; the seller internalizes the full fixed investment cost but only a share less than one of the attendant returns.

Under standard-cost transfer pricing, the selling division invests according to:

$$\begin{aligned} I_1^s \in \arg\max_{I_1} \{&\mathbb{E}_\theta[t^s(\theta, I_1^s) \cdot Q(t^s(\theta, I_1^s), \theta) \\ &- C(Q(t^s(\theta, I_1^s), \theta), \theta, I_1^s)] - w_1(I_1)\}. \end{aligned}$$

Since the seller will subsequently choose $t^s(\cdot)$ sequentially rationally so as to maximize his payoff, the envelope theorem yields the following necessary first-order condition:

$$-\mathbb{E}_\theta\left[\frac{\partial}{\partial I_1} C(q^s(\theta, I_1^s), \theta, I_1^s)\right] - w_1'(I_1^s) = \mathbb{E}_\theta[q^s(\theta, I_1^s)] - w_1'(I_1^s) = 0. \tag{2.8}$$

Comparing Equation (2.7) with Equation (2.8) reveals a tradeoff: the *firm-wide* increase in contribution margin associated with seller investments is larger under negotiations, because Coasian bargaining ensures *ex post* efficiency. On the other hand, the seller has the power to make a take-it-or-leave-it offer under the standard-cost-based regime and hence is able to extract a greater share of the (smaller) firmwide surplus, as compared with negotiations. (The buyer still earns some share of the surplus as the seller is constrained to linear pricing.)

To evaluate this tradeoff, it is critical to understand the determinants of the trade distortions under standard-cost pricing, in particular how these distortions relate to the shape of the buying division's demand function for the input (which in turn is derived of the demand this division faces for its final product). With linear production costs, it is well-known that the quantity sold by a monopolist who charges a uniform per unit price is exactly half the efficient quantity in the special case of a linear demand function. If, in addition, $\eta = \frac{1}{2}$ (Nash bargaining), then the seller would invest the exact same amount under either pricing regime, because then $\eta q^*(\theta, I_1) \equiv q^s(\theta, I_1)$. The first result now generalizes this observation by showing that the double-marginalization problem is particularly severe for convex demand functions.

Proposition 2.1. Suppose only the supplier invests and divisional costs and revenues are as in Equation (2.1). If $\eta \in [\frac{1}{2}, 1]$ and $r'(q, \theta)$ is convex in q, then negotiated transfer pricing strictly dominates standard cost-based transfer pricing.

If the buyer's (inverse) demand function for the intermediate good is convex, then $q^s(\theta, I_1) \leq \frac{1}{2} q^*(\theta, I_1)$, for any θ and I_1. The trade distortions under cost-based pricing will then be particularly severe, thereby dampening the seller investment incentives. If, at the same time, the seller has at least equal bargaining power ($\eta \geq \frac{1}{2}$), the severity of the holdup problem will be bounded from above and the seller will invest more under negotiations. On the other hand, if $r'(q, \theta)$ is strictly concave and $\eta < \frac{1}{2}$, then $I_1^n < I_1^s$ holds, and one needs to trade off the improved investments against the associated quantity distortion under standard cost transfer pricing. In order for this latter regime to generate higher expected firmwide profit than negotiations, the function $r'(q, \theta)$ will have to be sufficiently concave.

An extreme example of a concave internal demand function for the intermediate good obtains if the buyer needs a fixed quantity, say $\bar{q}(\theta)$, and is willing to pay up to $\bar{p}(\theta)$ for this quantity. That is, $Q(t, \theta) = \bar{q}(\theta)$ for any transfer price $t \leq \bar{p}(\theta)$; and $Q(t, \theta) = 0$ otherwise. In this extreme example standard-cost pricing achieves first-best performance, as the selling division can extract the entire surplus by quoting a transfer price $t^s(\theta) = \bar{p}(\theta)$, i.e., the double-marginalization problem disappears. Even for demand functions that are everywhere downward sloping, first-best can be attained (absent buyer investments) by dropping the linear pricing constraint under standard-cost pricing, i.e., by letting the seller charge a two-part tariff where the lump sum extracts the buying division's rent. Note that within the model framework this could also be interpreted as negotiations with $\eta = 1$.

2.3 Investment by the Buying Division

Consider now the complementary case where only investments by the buyer are of importance, i.e., $I_2 \geq \bar{I}_1 = 0$ (thus, I_1 will be suppressed in this subsection). Given the lopsided distribution of bargaining power under standard-cost pricing, one would suspect this regime to fare poorly for buyer investments, as the expected firmwide surplus will be reduced by double marginalization *and* the investing party has no bargaining power. This intuition will be confirmed below in the sense that negotiated transfer pricing can be shown to dominate in a wider class of environments than in the preceding section.

Negotiated transfer pricing experiences a qualitatively similar holdup problem as in the seller investment setting. This is evidenced by the buyer's first-order investment condition:

$$(1-\eta)\mathbb{E}_\theta[q^*(\theta, I_2^n)] - w_2'(I_2^n) = 0. \tag{2.9}$$

Under standard-cost pricing, the buying division anticipates that the seller will raise the cost quote (in general) in response to observing higher buyer investment so as to take advantage of the increased willingness-to-pay. The buying division's investment problem reads as

$$\begin{aligned} I_2^s \in \arg\max_{I_2} \{ \mathbb{E}_\theta [& R(Q(t^s(\theta, I_2), \theta, I_2), \theta, I_2) \\ & - t^s(\theta, I_2) \cdot Q(t^s(\theta, I_2), \theta, I_2)] - w_2(I_2) \} . \end{aligned}$$

Differentiating with respect to I_2 yields

$$\mathbb{E}_\theta \left[\left(1 - \frac{\partial t^s(\theta, I_2^s)}{\partial I_2} \right) q^s(\theta, I_2^s) \right] - w_2'(I_2^s) = 0. \tag{2.10}$$

Condition (2.10) exhibits a "new" holdup effect afflicting buyer investments under standard-cost transfer pricing. While the price charged by the selling division, $t^s(\theta, I_2)$, need not always be increasing in I_2, in many cases it will be. Specifically, it can be shown that (i) $\partial t^s / \partial I_2 \geq 0$, if $r'(q,\theta)$ is log-concave in q and (ii) $\partial t^s/\partial I_2 \geq \frac{1}{2}$, if $r'(q,\theta)$ is concave in q.[7] This leads to the next result:

Proposition 2.2. Suppose only the buyer invests and divisional costs and revenues as in (2.1). If $\eta \in \left[0, \frac{1}{2}\right]$ and $r'(q,\theta)$ is log-concave in q, then negotiated transfer pricing strictly dominates standard-cost transfer pricing.

Under the conditions stated in Proposition 2.2, $I_2^n \geq I_2^s$. If the internal demand function is concave in q, then, as stated above, $\partial t^s/\partial I_2 \geq \frac{1}{2}$, i.e., the holdup problem under cost-based pricing is particularly severe. At the same time $q^s(\cdot) < q^*(\cdot)$, which implies that the buyer's investment incentives will be stronger under negotiations. The same will hold, but for a different reason, if $r'(\cdot,\theta)$ is convex but log-concave. Convexity of $r'(\cdot,\theta)$ ensures that $q^s(\cdot) \leq \frac{1}{2}q^*(\cdot)$, i.e., the quantity distortion under cost-based pricing is especially severe. Log-concavity of $r'(\cdot,\theta)$ on the other hand ensures that $\partial t^s/\partial I_2 \geq 0$. These two conditions together again imply that the buyer will invest more under negotiations.

Consider the special case of a linear internal demand function for the intermediate good (i.e., $R(\cdot,\theta,I_2)$ is quadratic in q.) Propositions 2.1 and 2.2 then imply that negotiations are the preferred mechanism (i) for any $\eta \geq \frac{1}{2}$ if only the seller invests and (ii) for any $\eta \leq \frac{3}{4}$ (i.e., a wider range of parameters) if only the buyer invests.

2.4 Bilateral Investments

I now combine the preceding results and consider bilateral (simultaneous) investments. To assess the relative performance of the transfer pricing schemes, I focus on pure-strategy Nash equilibria of the induced investment games. The interdependence of the investment decisions can be most easily exemplified for negotiated transfer pricing. As can be seen from Equations (2.7) and (2.9), the divisions' investment incentives

[7] See Baldenius (2000). A differentiable function $f(x)$ is log-concave if the ratio $(f'(x)/f(x))$ is decreasing in x. At the same time, if the internal demand curve exhibits constant elasticity, then the optimal monopoly price is decreasing in positive demand shifts, i.e., $\partial t^s/\partial I_2 < 0$.

depend on the expected trading quantity. This quantity is increasing in the investments of each division. As a consequence, the reaction curves are positively sloped, i.e., divisional investments are strategic complements (e.g., Milgrom and Roberts, 1990). For the sake of brevity, I will sidestep the issue of equilibrium existence here and simply *assume* that there exist unique interior pure-strategy equilibria under the respective pricing regimes.[8]

A pure-strategy Nash equilibrium under negotiations is characterized by the following simultaneous equations:

$$\eta \mathbb{E}_\theta[q^*(\theta, I^n)] = w_1'(I_1^n) \quad \text{and} \quad (1-\eta)\mathbb{E}_\theta[q^*(\theta, I^n)] = w_2'(I_2^n), \qquad (2.11)$$

where $q^*(\cdot)$ is given by $r'(q^*(\cdot), \theta) + I_2 = c(\theta) - I_1$. Under rather mild conditions (see Footnote 8), one can show that this Nash equilibrium is characterized by underinvestment, i.e., $(I_1^n, I_2^n) < (I_1^*, I_2^*)$.[9] A pure-strategy Nash equilibrium under standard-cost pricing is given by

$$\mathbb{E}_\theta\left[q^s(\theta, I^s)\right] = w_1'(I_1^s) \quad \text{and} \quad \mathbb{E}_\theta\left[\left(1 - \frac{\partial t(\theta, I^s)}{\partial I_2}\right) q^s(\theta, I^s)\right] = w_2'(I_2^s). \qquad (2.12)$$

The following result essentially invokes the "intersection" of the conditions used in the unilateral investment scenarios:

Proposition 2.3. Suppose divisional costs and revenues as in Equation (2.1), both divisions invest, and unique interior pure-strategy Nash equilibria in investments exist under either regime. If $\eta = \frac{1}{2}$ (Nash bargaining), then negotiated transfer pricing strictly dominates standard-cost transfer pricing if $r'(q, \theta)$ is convex and log-concave in q.[10]

---

[8] BRS show that the following three conditions together are sufficient for existence of a unique interior pure-strategy equilibrium under negotiations: (i) $\pi^*(I_1, I_2)$ has a unique interior maximizer I^*; (ii) for any fixed value I_j, $\pi^*(I_i, I_j)$ is concave in I_i on the interval $[0, I_i^*]$, for $i, j = 1, 2,\ j \neq i$; and (iii) $w_i'(0) = 0$, for $i = 1, 2$. The equilibrium analysis is more complicated under standard-cost transfer pricing. Since the holdup term $\partial t^s / \partial I_2$ is itself a function of I_2, the reactions curves may not be well behaved; for that reason, BRS consider mixed-strategy equilibria under this regime.

[9] See BRS for a formal proof. Let strict vector inequalities $x \equiv (x_1, \ldots, x_n) > y \equiv (y_1, \ldots, y_n)$ postulate that $x_i > y_i, i = 1, \ldots, n$, whereas weak inequalities $x \equiv (x_1, \ldots, x_n) \geq y \equiv (y_1, \ldots, y_n)$ be equivalent to $x_i \geq y_i, i = 1, \ldots, n$, with at least one inequality strict.

[10] Note that a convex function is log-concave provided its second derivative is "sufficiently small."

The proof of Proposition 2.3 revolves around showing that, under the stated conditions, both divisions invest less under standard-cost transfer pricing than under negotiations. The reason is that straightforward generalizations of Propositions 2.1 and 2.2 imply that, holding constant the amount invested by one of the divisions — say, Division k — the respective other division — Division $j, j \neq k$ — will invest more under negotiations. This, combined with the fact that investments are strategic complements, establishes Proposition 2.3 for bilateral investments.

A special case of convexity combined with log-concavity is, of course, linearity. To illustrate, consider the special case where $\eta = \frac{1}{2}$ (Nash bargaining) while, at the same time, the buyer's demand function is linear in q. In that case, $q^s(\theta, I) \equiv \frac{1}{2}q^*(\theta, I)$, so that $I_1^n(I_2) \equiv I_1^s(I_2)$ while $I_2^n(I_1) > I_2^s(I_1)$, for all I_1. This case is depicted in Figure 2.2. By strategic complementarity of I_1 and I_2 (i.e., upward sloping reaction curves), even the selling division will invest more in equilibrium under negotiations, because it anticipates greater buyer investment than under the cost-based scheme. Evaluating the strength of this strategic complementarity, it can be shown for linear internal demand functions, $r'(q, \theta)$, and identical quadratic fixed cost functions, $w_i(\cdot)$, that negotiated transfer pricing strictly dominates standard-cost transfer

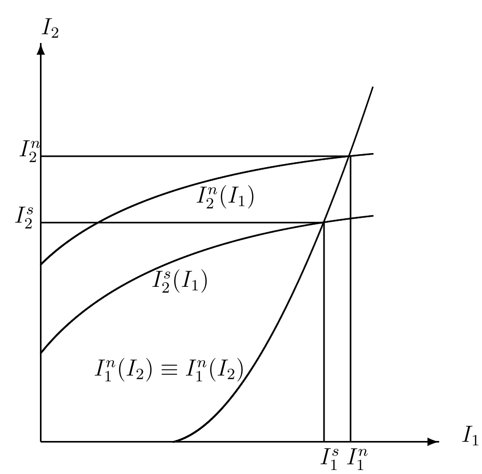

Fig. 2.2 Linear marginal revenues and $\eta = 1/2$.

pricing for any $\eta \in \left[\frac{1}{4}, \frac{3}{4}\right]$ (BRS, Corollary to Proposition 3). Negotiations thus dominate for a wide range of bargaining scenarios in this special case.

2.5 Alternative Approaches to Cost-Based Pricing

When implementing cost-based transfer pricing, firms have to make a number of design choices, most notably whether to use actual or standard costing (and how to derive those standards), and whether to include profit markups.[11] Empirically, firms prefer standard over actual cost, mostly because standard cost-based prices provide better incentives for efficiency improvements upstream. Another empirical finding is that firms frequently use cost-plus markups as a means to allocate part of the firmwide surplus to the upstream division. The preceding model of standard cost-pricing was somewhat extreme in that the supplying division was assumed to face no constraints whatsoever when self-reporting its cost. Clearly in this case, there is no role for markups, as they would simply be "undone" by the upstream division. While reporting discretion on the part of the seller and the associated distortions are the most commonly cited downside of standard-cost pricing, in practice this discretion may be limited by benchmarking of costs or internal audits.

If HQ had precise information about the seller's variable cost and only buyer investments are important, then setting the transfer price equal to the *actual* upstream variable costs (i.e., the Hirshleifer solution) would achieve full efficiency. If seller investments are critical, however, such a method would not work well as the seller would be unwilling to invest. In that case, adding a profit markup for the upstream division increases firmwide profit as it (partially) restores upstream investment incentives. In this context, Sahay (2003) points out that percentage markups over actual variable cost on the other hand would create perverse efficiency incentives as the seller would ultimately earn a contribution margin equal to some percentage of his

[11] Another dimension of choice is between full and variable costs as the cost basis. In the context of this model, however, this distinction is vacuous as fixed costs are by definition nonverifiable and hence cannot be included in the price.

actual cost for each unit traded. Thus cost-plus markups should be *additive* in nature, i.e., stated in absolute dollars per unit transferred. Percentage markups over actual variable cost on the other hand would create perverse cost reduction incentives as the seller would ultimately earn a contribution margin equal to some percentage of his actual cost for each unit traded.

To illustrate the role of additive cost-plus markups in providing investment incentives to the seller, suppose the actual variable unit cost of $c(\theta) - I_1$ (but not its individual components) was indeed observable to HQ at Date 3.[12] Denoting the markup by $m \geq 0$, the transfer price would be given by $t^a(\theta, I_1, m) \equiv c(\theta) - I_1 + m$ (with superscript "a" for *actual-cost* pricing). The seller's objective function when choosing I_1 at Date 1 then reads:

$$I_1^a(m) \in \arg\max_{I_1} \{\mathbb{E}_\theta[m \cdot Q(t^a(\theta, I_1, m), \theta)] - w_1(I_1)\}$$

with the corresponding first-order condition

$$m \cdot \mathbb{E}_\theta[-Q'(t^a(\theta, I_1^a(m), m), \theta)] - w_1'(I_1^a(m)) = 0.$$

If $w_1'(0) = 0$, then $I_1^a(m)$ is readily shown to be strictly positive for $m > 0$, and it is an increasing function. For positive markups, the seller will indeed invest as this will increase the expected quantity, for which he earns a unit contribution margin of \$$m$. HQ's objective is to choose m so as to maximize firmwide expected profit:

$$\begin{aligned}\pi(m) \equiv \mathbb{E}_\theta\{&R(Q(\bar{t}(\theta, m), \theta), \theta) - [c(\theta) - I_1^a(m)]Q(\bar{t}(\theta, m), \theta)\} \\ &- w_1(I_1^a(m)),\end{aligned} \tag{2.13}$$

where $\bar{t}(\theta, m) \equiv t^a(\theta, I_1^a(m), m)$. To see that imposing some positive additive markup over cost is always optimal, it suffices to show that $\pi'(0) > 0$. Recall that for any transfer price t the buyer will choose $Q(t, \theta)$ so that $R'(Q(t, \theta), \theta) = t$. Moreover, at $m = 0$, the seller will choose zero investment, so that $t^a(\theta, 0, 0) = c(\theta)$ and thus $Q(t^a(\theta, 0, 0), \theta) = q^*(\theta, I_1 = 0)$. By the envelope theorem, therefore:

$$\pi'(0) = \left(\frac{dI_1^a(m = 0)}{dm}\right) q^*(\theta, 0) > 0.$$

[12] Note that observability of $c(\theta) - I_1$ does not imply that the seller's investment I_1 can be inferred, because θ is unobservable.

The quantity distortion introduced by such a small markup is of second order, while the beneficial investment effect is of first order. The optimal markup, m^*, as derived in Sahay (2003), maximizes Equation (2.13) over all m by trading off the attendant double-marginalization distortions with the improved seller investments.[13]

A comprehensive comparison of various cost-based transfer pricing methods is found in Pfeiffer et al. (2009); see also Gox and Schiller (2007). In addition to (*ex ante*) type uncertainty as modeled above, their model also allows for *ex post* uncertainty in that, when ordering its input units, the buying division has incomplete information about the seller's cost. More precisely, at the trading stage the seller knows his variable production cost with certainty but only observes a noisy signal about the buyer's revenues, while the buyer observes perfectly his revenues but only a noisy signal of the upstream cost. This allows for a continuous parametrization of information asymmetry between HQ and the divisions. Cost standards can be determined by HQ or self-reported by the upstream division. The more precise the cost information held by HQ, the more effective it is to price at centrally determined standard cost. Note that such a system fully protects divisional investments from holdup, because the transfer price is fixed upfront (see also Baldenius, 2000). The main downside of centrally determined cost standards is lack of flexibility when HQ's information about cost realizations becomes coarser. Pfeiffer et al. (2009) derive thresholds for such a centralized mechanism to be dominated by decentralized ones.

Pricing at either centrally determined standard cost or at actual-cost-plus-markup suffer from *ex- post* trading inefficiencies because the buyer's imperfect cost information. Letting the upstream division self-report its cost (as was assumed in this section) has the advantage that the production cost is fully impounded (if in a biased way) in the price. As a result, Pfeiffer et al. (2009) show that this mechanism dominates the other cost-based mechanisms in settings where (i) there is

[13] Section 5 below will revisit profit markups in an asymmetric information setting. Markups (over properly defined standard cost) then may be advantageous even in the absence of investment opportunities.

sufficient *ex ante* uncertainty (i.e., centralization the standard setting process is problematic) and (ii) the buyer faces significant *ex post* cost uncertainty.[14]

[14] Pfeiffer et al. (2009) also "expand the contract set" for cost-based pricing mechanisms by allowing for profit markups over actual cost that effectively split the contribution margin between seller and buyer in a predetermined fashion. This method is closely related to negotiated transfer pricing in terms of the resulting divisional incentives, except that the profit split parameter (the η in the above analysis) can now be chosen endogenously by HQ. By revealed preference such "contribution-margin" pricing outperforms negotiations (where η is exogenous) and hence also tends to dominate decentralized standard-cost pricing unless the buyer faces high cost uncertainty when placing the order.

3

Tying Internal Prices to Market Prices

The analysis up to this point has assumed away any external markets for the intermediate good. The justification for that approach was that firms often start making their own (non-commoditized) inputs for lack of external markets to source them from. In many cases, however, external input markets do exist, yet they are far from the competitive, price-taking, and ideal. This section analyzes market-based transfer pricing when external input markets are imperfectly competitive. Borrowing from Baldenius and Reichelstein (2006, henceforth BR), I shall ignore relationship-specific investments for now and focus solely on *ex post* trading efficiency. Section 4, below, will sketch a framework that brings together external input markets and specific investments.

Hirshleifer (1956) has shown that market-based internal pricing ensures efficient decentralization with perfectly competitive markets. While this result breaks down for imperfect markets, survey evidence shows that, nonetheless, many firms tie transfer prices to market prices. It is therefore important to understand the distortions arising under this method, and to study additional instruments at firms' disposal to remedy them. In particular, this section highlights the role played by *internal discounts* in improving the efficiency of market-based internal pricing.

3.1 Model with External Input Markets

To study internal pricing in the presence of external input markets, consider a variation of the preceding model where Division 1 sells its output to Division 2 *and* to external customers. To keep matters simple, Division 2 in this model can only buy from Division 1.[1] Since relationship-specific investments are ignored for now, I will use *contribution margin* instead of profit as the measure of performance, both on the divisional level and the corporate level.

Revenues and costs are again affected by the state variable $\theta \in \Theta$. Division 1's production costs are assumed to be linear in the total quantity with unit variable cost of $c(\theta)$. The function $c(\cdot)$ is differentiable and varies nontrivially in θ (otherwise simple cost-based transfer pricing would ensure efficient decentralization). For now, internal and external sales are assumed to have the same costs. When Division 1 sets an external price of p, it faces a demand of $Q_e(p,\theta)$ for its output from external customers. For q_i internally transferred units, Division 2 will earn a net revenue of $R_i(q_i,\theta)$ from selling final goods. The added subscript "i" is used to distinguish Division 2's final goods revenues from the revenues $R_e(q_e,\theta) = q_e P_e(q_e,\theta)$ collected by Division 1 from its external sales of intermediate goods, where $P_e(\cdot,\theta)$ is the inverse of $Q_e(\cdot,\theta)$. With Division 1 having external market access, it is important to consider capacity constraints at this division.[2]

The First-Best Solution. The *first-best* quantities in case HQ could observe θ are denoted by $(q_i^*(\theta), q_e^*(\theta))$. They are found by solving the following program:

$$\mathcal{P}_{4.1}: \quad \max_{q_i,q_e} M(q_i,q_e,\theta) \equiv q_e P_e(q_e,\theta) + R_i(q_i,\theta) - c(\theta)[q_e + q_i],$$
$$\text{subject to:} \quad q_i + q_e \leq K(\theta). \tag{3.1}$$

[1] This specification appears descriptive of many cases where upstream divisions are launched as captive suppliers of downstream profit centers, and gradually expand their activities to include some external sales. This may reflect the proprietary nature of the product. In connection with the Paine Chemical Company, Eccles and White (1988, S35) cite managers who allege high transaction cost for switching to a potential outside supplier.

[2] Opportunity costs arising from external selling opportunities are a key rationale for firms to choose market-based transfer pricing, e.g., Horngren et al. (2007).

Here, $K(\theta)$ denotes a capacity constraint at the upstream division, which may vary with the underlying state θ. It is easy to see that, regardless of the capacity utilization, $R_i'(q_i^*(\theta),\theta) = R_e'(q_e^*(\theta),\theta)$ has to hold. If Equation (3.1) is binding, then $q_e^*(\theta) = K(\theta) - q_i^*(\theta)$; whereas if it is slack, then $R_j'(q_j^*(\theta),\theta) = c(\theta),\ j = i, e$. Let $p^*(\theta) \equiv P_e(q_e^*(\theta),\theta)$ denote the external market price at the first-best quantity.

The Decentralized Solution. In practice, operating decisions are generally delegated to division managers. The managers aim to maximize their own division's contribution margin, which for given transfer price t equals

$$M_1 = [p - c(\theta)]q_e + [t - c(\theta)]q_i$$

for the upstream division; for the downstream division we have

$$M_2 = R_i(q_i,\theta) - tq_i.$$

While fully observable to the division managers, the realization of θ is again assumed to be unobservable to HQ (and the divisions cannot verify θ to HQ).

To reiterate, this section is only concerned with analyzing the strengths and weaknesses of market-based pricing. With linear production costs and symmetrically informed managers, the first-best solution could be attained by decoupling internal from external prices and instead giving the buying manager unfettered pricing power for the internally traded quantities. Such a scheme will rarely be effective in practice though, as the assumption of the buying division knowing with certainty the relevant upstream costs is somewhat unrealistic. To avoid notational clutter, however, I shall stick to the symmetric information framework throughout this section.[3]

Decisions in the decentralized firm are made in the following order. Initially, HQ specifies a transfer pricing rule $t(p)$, which determines the internal price as a function of the external market price. For a given transfer pricing rule, $t(\cdot)$, and state, θ, the selling division will choose

[3] BR explicitly employ informational assumptions under which giving all bargaining power to the buyer would not achieve the first-best solution. Here, I only implicitly appeal to such arguments and instead exclude such regimes by fiat.

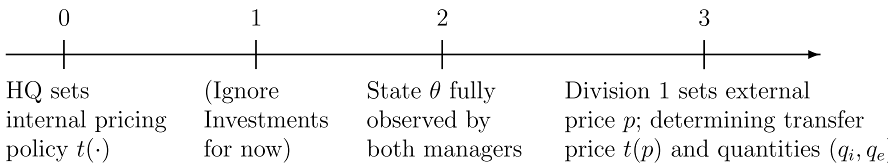

Fig. 3.1 Timeline for market-based transfer pricing.

the external sales price, $p(\theta)$, as the solution to the following program:

$$\mathcal{P}_{4.2}: \quad \max_p \; M_1(p,\theta \mid t(\cdot)) \equiv [p - c(\theta)]Q_e(p,\theta) + [t(p) - c(\theta)]Q_i(t(p),\theta) \tag{3.2}$$

subject to:

$$Q_i(t(p),\theta) \in \arg\max_{q_i}\{M_2(q_i,\theta \mid p,t(\cdot)) = R_i(q_i,\theta) - t(p)q_i\},$$

$$Q_i(t(p),\theta) + Q_e(p,\theta) \leq K(\theta). \tag{3.3}$$

In managing its capacity, the upstream division must give priority to the internal buyer before selling to outside parties.

Aside from the standard differentiability assumptions, assume further that both revenue functions, $R_k(\cdot,\theta)$, are concave in their respective quantity arguments, q_k, $k = i,e$. Finally, it will be convenient to decompose Division 1's objective in Equation (3.2) into contribution margins associated with internal and external sales, $M_{1,k}(p,\theta) \equiv [p - c(\theta)]Q_k(p,\theta)$, $k = i,e$, and assume that both these functions are concave in p for any θ. Denote by $p_k^m(\theta) \in \arg\max_p M_{1,k}(p,\theta)$ the respective monopoly prices that the upstream division would charge if it could freely price discriminate between the two markets absent any capacity constraints. Lastly, let $q_k^m(\theta) \equiv Q_k(p_k^m(\theta),\theta)$ be the corresponding unrestricted monopoly quantities.

Accounting textbooks suggest that market-based transfer pricing performs well with scarce upstream capacity, e.g., Horngren et al. (2007). The reason is that the marginal cost of a unit transferred internally then equals the foregone external contribution margin. The following analysis confirms this intuition. At the same time, one might expect a stronger case for internal discounts in case of abundant capacity, in the sense that, if discounts increase efficiency with scarce capacity, then *a fortiori* this should be true also for abundant capacity. After all, with abundant capacity the marginal cost of a unit transferred internally is simply its variable production cost, while externally the firm wants the upstream division to price monopolistically; thus internal transfers should be discounted "sharply". As Section 3.3 shows, however, this intuition is incomplete.

3.2 Binding Upstream Capacity Constraint

To streamline the exposition, I focus for the most part on pricing rules which determine the internal transfer price by subtracting an additive discount Δ from the externally charged price, that is, $t(p) = p - \Delta$.[4] Given the discount Δ and the state θ, the upstream division selects an external market price $p(\Delta, \theta)$ that solves Program $\mathcal{P}_{4.2}$ for $t(p) = p - \Delta$. Suppose HQ has sufficient information about the state of the industry to predict that capacity will be sufficiently scarce in that Equation (3.3) will be binding for all states.

I will ask, first, if starting from unadjusted market-based transfer pricing ($\Delta = 0$), imposing a discount can improve firmwide performance. Technically, this amounts to signing $M'(\Delta = 0)$, where $M(\Delta) \equiv \mathbb{E}_\theta[M(Q_i(p(\Delta,\theta) - \Delta, \theta), Q_e(p(\Delta,\theta),\theta),\theta)]$ is the firmwide expected contribution margin for any Δ. Note that with constrained capacity, the corporate objective is to maximize total (net) revenue with the available upstream capacity. In setting the market price, however, the upstream division does not internalize the entire internal revenue of $R_i(\cdot, \theta)$ obtained by the downstream division, but only that division's willingness-to-pay for the intermediate product. If transfers are valued at market price, this will again cause a double marginalization bias.[5] Internal discounts may alleviate this bias.

Second, I ask the more ambitious question if a simple internal discount rule (not necessarily additive though) can ever achieve the first-best solution. A necessary condition for eliminating the double

[4] BR instead study proportional discounts, which appear to be more common in practice but make the model more complex for their lack of multiplicative separability. The results are largely unaffected by this modification.

[5] This feature of the model is consistent with the observation in Eccles and White (1988, S31).

marginalization bias is that the discount be equal to the difference between external market price, $p^*(\theta)$, and the associated marginal revenue, $R'_e(q_e^*(\theta),\theta)$, *at the first-best quantity*. This in turn requires that the transfer pricing rule sets the internal discount inversely to the price elasticity of demand at $p^*(\theta)$. The downstream division will then face the appropriate decentralizing price equal to the external marginal revenue at the optimal quantity $q_e^*(\theta)$. Note, however, that for this to be sufficient for efficient delegation, the upstream division must not have incentives to to deviate from $p^*(\theta)$ externally. To formalize this notion, denote by

$$\varepsilon(p,\theta) = -\left(\frac{dQ_e(p,\theta)}{dp}\right)\left(\frac{p}{Q_e(p,\theta)}\right),$$

the price elasticity of external demand. For brevity, define $\varepsilon^*(\theta) \equiv \varepsilon(p^*(\theta),\theta)$.

Proposition 3.1.

(i) If the capacity constraint (3.3) is binding at $p(0,\theta)$ for any θ, then $M'(\Delta = 0) > 0$, i.e., imposing an internal discount improves firmwide performance.

(ii) Suppose $q_i^*(\theta) < q_i^m(\theta)$ and $q_e^*(\theta) < q_e^m(\theta)$ for any θ. If there exists a function $f(\cdot)$ such that $f(p^*(\theta)) = \varepsilon^*(\theta)$ for all θ, then the transfer pricing rule

$$t(p) = p\left(1 - \frac{1}{f(p)}\right) \tag{3.4}$$

implements the first-best solution, provided $t(p)$ is increasing in p.

Proposition 3.1 shows that with scarce upstream capacity, imposing some internal discount is always valuable (Part (i)), and, in some cases, properly designed internal discount rules can achieve full efficiency (Part (ii)). Consider again the bias in Division 1's pricing problem at $\Delta = 0$. The marginal (opportunity) cost of shipping one unit internally is not the external market *price*, but the external *marginal revenue,* which of course will always be less than the market price. Imposing a discount, Δ, induces the upstream division to increase the market price, while at the same time reducing the transfer price, $p(\Delta,\theta) - \Delta < p(0,\theta)$, in such a manner that some units of the intermediate good will be *redirected* from external to internal sales where they yield higher marginal (net) revenues to the firm. (For small Δ, by continuity, the capacity constraint remains binding, in that $Q_i(p(\Delta,\theta) - \Delta,\theta) + Q_e(p(\Delta,\theta),\theta) = K(\theta)$; thus Division 1's total output remains unaffected by Δ.) As a result, $M'(\Delta = 0) > 0$.[6]

One way to interpret internal discounts is that HQ forces the seller to price discriminate in favor of the internal buyer, whereas for $\Delta = 0$ the seller has to price uniformly. The imposition of a discount always makes the internal buyer better off in equilibrium (as the transfer price will go down). If the unrestricted monopoly price is greater externally than internally, i.e., $p_e^m(\theta) \geq p_i^m(\theta)$, then the discount will make the seller better off also, as he can move both internal and external prices closer to their respective monopoly values. On the other hand, if $p_e^m(\theta) < p_i^m(\theta)$, then a discount makes the seller worse off, but this is more than offset by the resulting increase in the downstream division's payoff.

Part (ii) of Proposition 3.1 demonstrates that, under certain conditions, a simple *multiplicative* discount rule (independent of θ, which HQ does not observe) achieves full efficiency. It postulates an invertibility property requiring that for every external price $p^*(\theta)$ there is at most one corresponding price elasticity $\varepsilon^*(\theta)$. Note that by definition of the price elasticity of demand, external marginal revenue for any external price p_e is equal to $R_e'(Q_e(p_e,\theta),\theta) = p_e(1 - \frac{1}{\varepsilon(p_e,\theta)})$. Suppose for now that the upstream division does indeed choose the optimal external price $p^*(\theta)$ (to be verified shortly). By construction of the internal pricing rule, the transfer price will then be equal to $R_e'(q_e^*(\theta),\theta)$, which will induce the downstream division to demand the first-best quantity $q_i^*(\theta)$. It remains to confirm that the upstream division has no incentives to deviate from the first-best external price.

[6] As the proof demonstrates, Proposition 3.1, Part (i) holds pointwise, i.e., for any state θ. The same will hold for Proposition 3.4, below.

Under the stated conditions, it cannot do better by raising the price above $p^*(\theta)$. Externally, it would move further away from the monopoly price, $p_e^m(\theta)$, because $q_e^*(\theta) \leq q_e^m(\theta)$; the same argument would apply to internal sales, provided $t(p)$ is increasing. On the other hand, the capacity constraint prevents the upstream division from lowering the external price below $p^*(\theta)$. Thus, the upstream division will be effectively "wedged in" between its desire to extract downstream rents and the binding capacity constraint.

When the state of the world is complex (i.e., θ is multidimensional), the invertibility condition at the heart of Proposition 3.1, Part (ii), will generally break down. On the other hand, it is straightforward to construct examples in which θ is of low dimensionality such that the result holds. An important application is the case where the price elasticity of demand does not depend on θ directly but only through the price $p^*(\theta)$, as in the following constant elasticity scenario:

Corollary 3.2. If the external market demand exhibits constant price elasticity, i.e., $Q_e(p,\theta) = a_e(\theta)p^{-\varepsilon}$ with $\varepsilon > 1$, a proportional discount rule of the form $t(p) = (1 - \delta^o)p$, with $\delta^o = \frac{1}{\varepsilon}$, achieves the efficient outcome.

The key to this result is that the price elasticity does not depend on θ. Constant elasticity demand functions of course are a staple in the industrial organization literature and practice, where the price elasticity is generally viewed as a measure of the competitiveness of a market. This finding is therefore consistent with the conventional wisdom that transferring at (unadjusted) market price becomes approximately efficient as the external market becomes more competitive, i.e., as ε grows "large".

3.3 Abundant Upstream Capacity

The case for intracompany discounts seems intuitively even more compelling when it is commonly known that production capacity at the upstream division will be effectively unconstrained, as the marginal cost of internal transfers then is simply $c(\theta)$. If internal transactions are valued at the prevailing market price, the upstream division will set a price that is an average of the external and the internal monopoly prices. The first-best solution, on the other hand, would call for external sales to be priced at $p_e^m(\theta)$ and internal units to be transferred at marginal cost, $c(\theta)$.[7] Hence, intuition suggests that internal transfers should be discounted "even more" than in the case of a binding upstream capacity constraint.

Figure 3.2 illustrates this intuition (suppressing the state variable θ) for the special case of linear inverse demand functions for the input, P_e and MR_i (recall that the downstream division's demand for the input is derived from its marginal revenue for the final product). It depicts the first-best quantities and the internal transfer price, $t^* \equiv MR_i(q_i^*(\theta), \theta)$, that would implement $q_i^*(\theta)$ in a decentralized firm, *provided the upstream division indeed were to charge a price* $p^*(\theta)$ *externally.* We saw above that the first-best solution with scarce upstream capacity (Figure 3.2a) calls for $MR_e(q_e^*(\theta), \theta) = MR_i(q_i^*(\theta), \theta)$ and $q_i^*(\theta) + q_e^*(\theta) = K(\theta)$. That is, a necessary condition for efficiency to

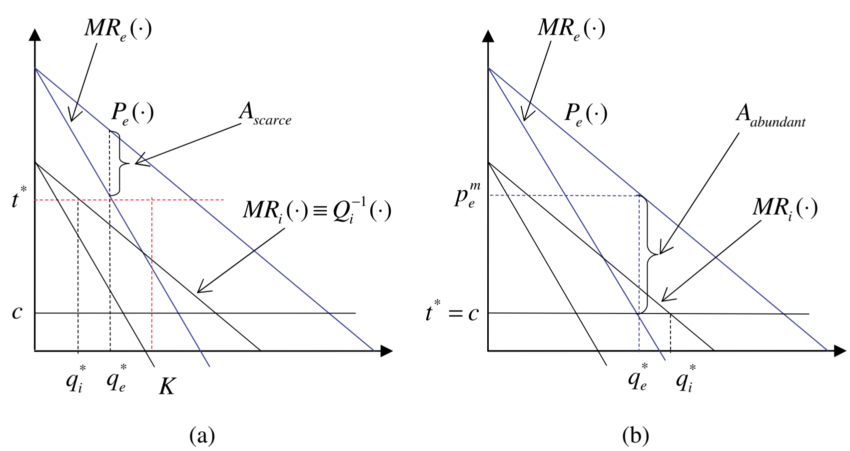

Fig. 3.2 (a) Scarce capacity. (b) Abundant capacity.

[7] With abundant capacity, the upstream capacity can be viewed as a "quasi-cost center" with regard to internal transfers. Since the two sales channels are effectively decoupled, there is no opportunity cost of internal transfers; the textbook solution would thus call for marginal cost transfers (Horngren et al., 2007), leaving aside upstream investments.

obtain in either capacity scenario is that the internal discount be equal to $\Delta = A(q_e^*(\theta), \theta)$, where $A(q_e, \theta) \equiv P_e(q_e, \theta) - MR_e(q_e, \theta)$ denotes the distance between external price and marginal revenue at the efficient quantity. Key to the above intuition now is that as long as the demand function $Q_e(\cdot, \theta)$ is not too convex in p, $A(q_e, \theta)$ will be increasing in q_e (this obviously holds for linear demand functions).[8] Together with the fact that $q_e^*(\theta)$ is greater with abundant than with scarce capacity, this yields $A_{abundant} \geq A_{scarce}$, hence suggesting a stronger case for discounts when capacity is abundant (Figure 3.2b).

However, this intuition is incomplete as it ignores the upstream division's pricing behavior. As argued above, when capacity is constrained, the adjustment in the external price triggered by the imposition of a discount always moves the sales quantities in the right direction so as to improve total profit. This will not always be the case with abundant capacity:

Proposition 3.3. Suppose upstream capacity is effectively unconstrained. Then:

(i) The sign of $M'(\Delta = 0)$ is indeterminate.
(ii) There does not exist a transfer pricing rule $t(\cdot)$ that induces the efficient outcome.

Part (i) of Proposition 3.3 states that (small) internal discounts may make the firm worse off with abundant upstream capacity. BR provide an example where the downstream division faces an external order of fixed size, and hence its demand for the input is inelastic up to a reservation price $\bar{p}_i$, and zero above that price. If the input quantity required by the internal buyer is large, then absent a discount the seller will set

[8] More precisely, $A(q_e, \theta)$ is increasing q_e for any θ, if and only if the external demand function $Q_e(p_e, \theta)$ is log-concave in p_e for any θ. In demonstrating this I shall suppress θ to save on notation. Note that $A(q_e) \equiv P_e(q_e) - R'_e(q_e)$ so that $A'(q_e) = -[P''_e(q_e)q_e + P'_e(q_e)]$. Clearly, $A'(q_e) > 0$ if $P_e(\cdot)$ is a concave function. (Hence, so is its inverse $Q_e(\cdot)$.) Moreover, by definition, $q_e \equiv Q_e(P_e(q_e))$ and therefore $P'(q_e) \cdot Q'_e(P_e(q_e)) = 1$ and $P''_e(q_e) = -Q''_e(P_e(q_e))/[Q'_e(P_e(q_e))]^3$. Using both these conditions shows that $A'(q_e) < 0$ if and only if $Q''_e(P_e(q_e)) \cdot Q_e(P_e(q_e)) > [Q'_e(P_e(q_e))]^2$, which is equivalent to $Q_e(\cdot)$ being log-convex in p_e.

the price equal to $\bar{p}_i$ to fully extract the firmwide surplus. Suppose now HQ imposes an internal discount. The seller will simply raise the external price in such a way that the resulting discounted transfer price again equals $\bar{p}_i$. Firm-wide gains from internal transfers will be unaffected, while those from external sales by Division 1 will be reduced if the external monopoly price is less than $\bar{p}_i$, because the external price will be pushed away even farther from the external monopoly price.[9]

On the other hand, the impact of intracompany discounts is unambiguously positive if the external monopoly price exceeds the internal one. Higher discounts always translate into higher external but lower internal prices. That is, for any θ and any $\Delta_{oo} > \Delta_o \geq 0$, $p(\Delta_{oo}, \theta) \geq p(\Delta_o, \theta)$ and $p(\Delta_{oo}, \theta) - \Delta_{oo} \leq p(\Delta_o, \theta) - \Delta_o$, by simple revealed preference arguments.[10] Therefore, if $p_e^m(\theta) > p_i^m(\theta)$, an internal discount effectively allows the supplying division to engage in "benign" price discrimination: not just do the two prices move closer to their respective monopoly values (from the upstream division's perspective), but the internal transfer price moves in the direction of marginal cost, which is the desired transfer price from a corporate perspective. As a result, both divisions will be better off. Conversely, a discount may be harmful if it raises internal trade only marginally and, at the same time, $p_e^m(\theta) < p_i^m(\theta)$, as is the case for instance in the fixed-quantity scenario in the preceding paragraph.

Now consider the impossibility result in Part (ii) of Proposition 3.1. Suppose the invertibility condition in Proposition 3.1, Part (ii), is satisfied but capacity is abundant so that $q_e^*(\theta) + q_i^*(\theta) < K(\theta)$ for any θ. Recall that the transfer pricing function stated in that earlier result constitutes a necessary condition for efficient decentralization. For it also to be sufficient, the supplying division must have incentives indeed

[9] Note that the conclusion of this example could also be obtained if internal demand was "somewhat" elastic (for prices below the reservation price) and therefore the firm was to face a double marginalization problem.

[10] This is implied by the additive nature of the discount and the fact that the upstream division's profits from internal and external sales are each concave in price. The upstream division's pricing behavior is less well behaved in BR. With proportional discounts, the upstream division may have incentives to "overreact" in response to the imposition of an internal discount by increasing the external price in such an extreme manner that the resulting transfer price also increases.

to charge an external price of $p^*(\theta)$ in response to the transfer pricing rule. The driving force behind Part (ii) of Proposition 3.3 now is that the upstream division would always respond to this discount function by raising the external price above $p^*(\theta)$, because the loss of external profits at the monopoly price is of second order compared with a first-order gain on internal transactions.

The sharp contrast between Propositions 3.1 and 3.3 highlights the importance of capacity constraints for the performance of market-based transfer pricing. With constrained capacity the upstream division can be "wedged in" between the discount and the capacity constraint, in that any price increase beyond $p^*(\theta)$ would result in a first-order loss to the upstream division since $p^*(\theta)$ already exceeds the external monopoly price. With unconstrained upstream division, in contrast, the seller always has incentives to deviate upward with the external price.[11]

Linear Demand Scenario. Given the inconclusive result of Proposition 3.3, it is useful to examine the tradeoffs associated with intracompany discounts under abundant capacity in more detail. To that end, consider a setting in which both internal and external demand functions are linear:

$$Q_e(p,\theta) = \alpha_e(\theta) - \beta_e(\theta)p \quad \text{and} \quad Q_i(t,\theta) = \alpha_i(\theta) - \beta_i(\theta)t. \tag{3.5}$$

Denote the corresponding external price function by $P_e(q_e,\theta) = a_e(\theta) - b_e(\theta)q_e$, so that $a_e(\theta) \equiv \alpha_e(\theta)/\beta_e(\theta)$, and $b_e(\theta) \equiv 1/\beta_e(\theta)$. The buying division's net revenue from the intermediate product then is given by $R_i(q_i,\theta) = [a_i(\theta) - \frac{1}{2}b_i(\theta)q_i]q_i$, so that $R_i'(q_i,\theta) = a_i(\theta) - b_i(\theta)q_i$, where $a_i(\theta) \equiv \alpha_i(\theta)/\beta_i(\theta)$ and $b_i(\theta) \equiv 1/\beta_i(\theta_i)$.[12]

[11] In light of Corollary 3.2 and Proposition 3.3, Part (ii), it is natural to ask how a proportional discount rule of the form $t(p) = (1 - 1/\varepsilon)p$ fares if the external price elasticity, ε, is constant and capacity is abundant. As shown in BR, the residual loss becomes negligible as the external market becomes "large". It then becomes too costly (in terms of foregone external sales) for the upstream division to extract rents from the downstream division by raising the external price above $p^*(\theta)$.

[12] To avoid corner solutions, restrict attention to parameter values such that, absent any discount, the seller finds it advantageous to charge a price which attracts both external and internal sales. Specifically, this requires that for all θ: $p(0,\theta) = \frac{1}{2}\left[\frac{\alpha_e(\theta)+\alpha_i(\theta)}{\beta_e(\theta)+\beta_i(\theta)} + c(\theta)\right] <$ $\min\{a_e(\theta), a_i(\theta)\}$, where $(0,\theta)$ is the average monopoly price given linear demand curves.

For this linear demand scenario, it turns out that even if the upstream division may be worse off due to the internal discount, this will always be offset by the increase in profit recorded by the buying division.

Proposition 3.4. Suppose capacity is unconstrained and both internal and external demands are given by linear functions as in Equation (3.5). Introducing an internal discount then increases firmwide performance, i.e., $M'(\Delta = 0) > 0$.

Contrary to the message of Proposition 3.3, but in line with the constrained capacity setting, discounts are unambiguously beneficial with linear demand curves. The key to Proposition 3.4 is that with linear demand functions, the upstream division will raise the external price in response to an internal discount by just so much that the decrease in external quantity is offset by the increase in internally transferred units. More formally, let $q_i(\Delta,\theta) \equiv Q_i(p(\Delta,\theta) - \Delta,\theta)$ and $q_e(\Delta,\theta) \equiv Q_e(p(\Delta,\theta),\theta)$, then $d[q_i(\Delta,\theta) + q_e(\Delta,\theta)]/d\Delta \equiv 0$, for any θ.[13] Thus, total upstream output quantity will be unaffected by Δ, and it can be written therefore simply as $\bar{q}(\theta) \equiv q_i(\Delta,\theta) + q_e(\Delta,\theta)$. As a consequence, Δ will again only redirect units of the upstream division's output from external sales to internal transfers revenues (as in Proposition 3.1), which leaves the firm better off as internal transfers are subject to double marginalization at $\Delta = 0$.

While Proposition 3.3, Part (ii), has shown that no market-based transfer pricing function can ever achieve the first-best solution with abundant capacity, it is straightforward to compute the *optimal* additive discount (maximizing expected firmwide contribution margin) in the linear demand setting; see BR for details. Since total upstream quantity $\bar{q}(\theta)$ is unaffected by Δ, total production cost will be constant in Δ, and the sole effect will be on revenues. As argued above, for $\Delta = 0$ and the attendant (external and internal) price $p(0,\theta)$, the

[13] This property generalizes a well-known finding by Robinson (see Tirole, 1988): when a monopoly is required to charge the same price in two markets, it will sell a total quantity that is equal to the sum of the separate monopoly quantities. Clearly, the linearity of the two demand curves is crucial for this property to hold.

firmwide marginal revenue for externally sold intermediate goods is less than that for internally transferred units. As Δ increases, more units are redirected from external sales to internal transfers, which will eventually equate the respective marginal revenues. The optimal discount has exactly this feature, that the marginal revenues are the same for internally and externally shipped units, in expectation over θ.

To conclude this section, note that the preceding results also speak to a setting in which HQ does not know *a priori* whether capacity will be effectively constrained. For the case of linear demand curves, Propositions 3.1 and 3.4 taken together show that discounts unambiguously improve the corporate performance, regardless of the upstream capacity utilization.

3.4 Cost Differences

The model has so far ignored the possibility of cost differences between internal and external sales and instead aimed to develop a rationale for internal discounts solely based on double-marginalization problems. Yet, survey evidence on market-based transfer pricing often cites cost differences (e.g., SG&A, transportation costs) as a justification for imposing intracompany discounts. Propositions 3.1 and 3.4 can be extended easily to settings in which internal transfers are less costly than external transfers. If $c_i(\theta) = c(\theta) - k$ with $k \geq 0$, the marginal revenue of internal sales effectively increases by k dollars per unit and therefore the change in profit associated with a (small) intracompany discount is even larger.

At the same time, note that the reasoning in fixed-quantity scenario discussed after Proposition 3.3 also extends to settings with upstream cost differences. Even with internal sales being less costly than external sales, imposing an internal discount in that example would still reduce firmwide profit. Cost differences between internal and external sales thus are neither necessary nor sufficient for the desirability of intracompany discounts.

If discounts are desirable with identical costs across internal and external sales, however, then this will hold *a fortiori* also with cost differences. The question however is, at what rate should cost differences be reflected in discounts. While surveys and textbooks are typically vague about this, there is some anecdotal evidence that firms often use a "dollar-for-dollar" rule in that, a cost differential of $\$x$ should raise Δ by $\$x$ also. For the linear demand scenario, BR however show that the optimal discount increases in k at a rate less than one. Recall that the internal discount should be chosen so that, in equilibrium, the expected external marginal revenue exceeds the expected internal marginal revenue by the cost difference k. As k increases, HQ anticipates that, all else equal, the upstream division will lower the external price $p(\cdot)$, and thereby also the transfer price. Therefore, internal cost savings should not be reflected dollar-by-dollar in Δ^*.

4

Bringing Investments and External Markets Together

The preceding section has demonstrated that, even absent cost differences between internal and eternal sales, internal discounts can increase the expected firmwide contribution margin. The results were derived in a setting where the production technology was exogenously given, i.e., in the absence of relationship-specific investments. This section sketches a model that combines the specific investment model of Section 2 with the imperfect external market model of Section 3.

It is easy to see that Hirshleifer's (1956) efficiency result for perfectly competitive markets generalizes to settings with specific investments. If both divisions can buy and sell virtually unlimited quantities of the intermediate product at a given price in the external market, then not only will they trade efficiently, but they will also be protected from holdup. (A perfectly competitive market price is unaffected by the investment undertaken by an individual firm, and thus each division is the residual claimant for its respective marginal investment returns.) As argued above, however, the competitive market assumption is not very descriptive. In the following, I therefore retain the market structure assumed in Section 3, i.e., the buyer has no market access while the seller has pricing power in the external market. Invoking the taxonomy developed in the Introduction, the bargaining power then rests with the selling division, but the expected total surplus will be reduced under unadjusted market-based transfer pricing because of double marginalization. As the preceding section has shown, internal discounts can improve the efficiency of *ex post* trade, but it left open the question of how such discounts affect divisional investment incentives.

In the remainder of this section, I shall only show that Proposition 3.4 generalizes to a setting with divisional investments. That is, given linear demand functions for the intermediate good, the firm will be better off introducing some internal discount, factoring in that division managers will choose their investments endogenously (and non cooperatively) in response to this discount. In fact, the potential for such relationship-specific investments can be shown to add *further impetus* for HQ to adjust internal prices away from external market prices. The following analysis is merely a first pass at this topic, more work remains to be done, e.g., regarding the interplay of capacity constraints and divisional investment incentives.

4.1 Model with External Markets and Investments

The following model sketch retains a key feature from Section 3: the upstream division can sell internally and externally (where it enjoys pricing power) while the downstream divisions can only source internally. But now, as in Section 2, both the upstream and downstream divisions can invest in relationship-specific assets at Date 1. The buyer invests $I_2 \in [0, \bar{I}_2]$ so as to increase his final goods (net) revenues, while the seller invests $I_1 \in [0, \bar{I}_1]$ in reducing his variable manufacturing cost. Specifically, Division 1's variable costs are $C(\bar{q}, \theta, I_1) = [c(\theta) - I_1]\bar{q}$, where $\bar{q} \equiv q_i + q_e$ is the seller's total output. Division 1 sells q_e units of output externally at revenues of $R_e(q_e, \theta) = P_e(q_e, \theta)q_e$, and ships q_i

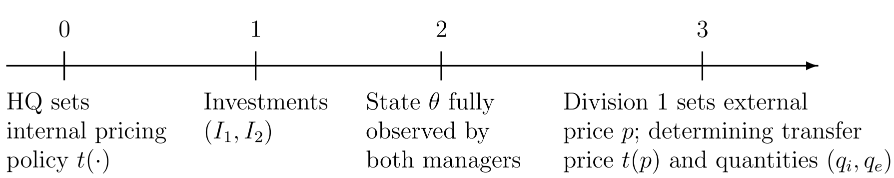

Fig. 4.1 Timeline for market-based transfer pricing with investments.

units internally. Division 2 further processes those q_i units and realizes revenues of $R_i(q_i, \theta, I_2) = r(q, \theta) + I_2 q_i$. That is, as in Section 2 above, investments affect marginal costs and revenues in an additively separable fashion.[1] As in Section 3, the externally and internally sold quantities are determined by the respective demand functions $Q_e(\cdot)$, which is the inverse of $P(\cdot)$, and $Q_i(\cdot)$, as derived from the downstream division's net revenue function R_i.

In the following analysis, any potential capacity constraint on the part of the upstream division is taken to be sufficiently lax so as never to be binding. The state variable θ is again assumed to be fully observed by both division managers but not HQ.

The First-Best Solution. Let $q_k^*(\theta, I),\ k = i, e$, denote the efficient quantities conditional on investments undertaken and the realized state. With abundant upstream capacity, they are determined as follows:

$$R_i'(q_i^*(\theta, I), \theta, I_2) = R_e'(q_e^*(\theta, I), \theta) = c(\theta) - I_1.$$

Furthermore, $M^*(\theta, I) \equiv R_i(q_i^*(\theta, I), \theta, I_2) + R_e(q_e^*(\theta, I), \theta) - [c(\theta) - I_1]\bar{q}^*(\theta, I)$ denotes the firmwide expected contribution margin given *ex post* efficient quantities. The first-best investments then solve the optimization problem:

$$I^* \in \arg\max_I \ \mathbb{E}_\theta[M^*(\theta, I)] - \sum_{j=1}^{2} w_j(I_j).$$

Given the separability assumptions made above, the first-best investment choices are given by:

$$\mathbb{E}_\theta\left[\bar{q}^*(\theta, I)\right] - w_1'(I_1^*) = 0,$$
$$\mathbb{E}_\theta\left[q_i^*(\theta, I^*)\right] - w_2'(I_2^*) = 0.$$

The marginal production cost savings associated with investing in I_1 apply to the entire upstream output, $\bar{q}$. In contrast, the marginal increase in downstream net revenues due to investing in I_2 applies only to the internally transferred units, q_i.

[1] Other investment scenarios are conceivable, e.g., the seller could invest in marketing so as to increase his external selling opportunities or invest in capacity expansion if this constraint becomes binding.

The Decentralized Solution. Now return to the delegation game where prices and quantities are chosen by divisional managers whose divisional performance measures are affected by the transfer price t. As in Section 3, I focus on additive internal adjustments to the external price, i.e., $t(p) = p - \Delta$. The respective divisional contribution margins then are:

$$\begin{aligned} M_1(p,\Delta,\theta,I) &= [p - c(\theta) + I_1]Q_e(p,\theta) \\ &\quad + [p - \Delta - c(\theta) + I_1]Q_i(p - \Delta,\theta,I_2), \\ M_2(p,\Delta,\theta,I) &= R_i(Q_i(p - \Delta,\theta,I_2),\theta,I_2) - [p - \Delta]Q_i(p - \Delta,\theta,I_2). \end{aligned}$$

Expected divisional profits equal $\pi_j = \mathbb{E}_\theta[M_j(\cdot)] - w_j(I_j)$, where $w_j(\cdot)$ denotes divisional fixed costs, with $w_j(I_j) \geq 0$, $w_j'(0) = 0$, and $w_j''(\cdot) > 0$.

As before, assume only the division managers observe θ and HQ's role is confined to choosing the transfer pricing rule (from among market-based mechanisms, i.e., the choice of Δ) at Date 0. The solution to this game closely mirrors that in Section 3 except for the added investment stage at Date 1. At Date 4, for any Δ, θ, and I, the selling division will choose $p(\Delta,\theta,I)$ as the solution to the pricing problem:

$$\max_p \; M_1(p,\Delta,\theta,I). \tag{4.1}$$

To streamline the notation, denote $q_i(\Delta,\theta,I) \equiv Q_i(p(\Delta,\theta,I) - \Delta,\theta,I_1)$; $q_e(\Delta,\theta,I) \equiv Q_e(p(\Delta,\theta,I),\theta)$; $M_j(\Delta,\theta,I) \equiv M_j(p(\Delta,\theta,I), \Delta,\theta,I)$; and $\pi_j(\Delta,I) \equiv \mathbb{E}_\theta[M_j(\Delta,\theta,I)] - w_j(I_j)$. At Date 1, for any internal pricing rule, $t(p)$, the division managers choose their respective investments in a pure-strategy Nash equilibrium, $I(\Delta) = (I_1(\Delta), I_2(\Delta))$, as follows:

$$I_j(\Delta) \in \arg\max_{I_j} \; \pi_j(\Delta, I_j, I_k(\Delta)), \quad j,k = 1,2, \quad k \neq j. \tag{4.2}$$

For ease of presentation, I focus on the case where, for any Δ, a unique interior pure-strategy Nash equilibrium exists.[2] Lastly, HQ aims at

[2] Uniqueness of Nash equilibrium for any Δ is ensured if the fixed cost functions $w_j(I_j)$ are sufficiently convex. The same holds for existence of a unique maximizer to the firmwide objective function in the first-best benchmark solution presented earlier.

maximizing over all Δ the expected firmwide profit, taking into account the equilibrium investment, pricing, and quantity choices:

$$\pi(\Delta) \equiv \sum_{j=1}^{2} \pi_j(\Delta, I(\Delta)).$$

I now show that Proposition 3.4, on the benefits of internal discounts, can be generalized to settings with relationship-specific investments. To that end, assume again that both the internal and the external markets are described by a linear-quadratic scenario in that $R_i(q_i,\theta,I_2) = \left[a_i(\theta) + I_2 - \frac{b_i(\theta)}{2} q_i\right] q_i$ and $P_e(q_e,\theta) = a_e(\theta) - b_e(\theta) q_e$. The upstream division then faces the following external and internal demand functions for given external price p, state θ, and buyer investment I_2:

$$\begin{aligned} Q_e(p,\theta) &= \alpha_e(\theta) - \beta_e(\theta) p, \\ Q_i(p-\Delta,\theta,I_2) &= \alpha_i(\theta) - \beta_i(\theta)[p - \Delta - I_2], \end{aligned} \tag{4.3}$$

where, as in Section 3.3, $a_j(\theta) \equiv \alpha_j(\theta)/\beta_j(\theta)$ and $b_j(\theta) \equiv 1/\beta_j(\theta)$, for $j = e,i$. It is straightforward to show that the total upstream quantity will again be invariant in Δ, due to the linearity of demand. Thus, let $\bar{q}(\theta,I) \equiv q_i(\Delta,\theta,I) + q_e(\Delta,\theta,I)$, for any Δ. Note that the specification in Equation (3.5) above constitutes a special case where $I_1 = I_2 = 0$.

The next result demonstrates that internal discounts remain desirable in the more general setting where division managers make specific investments contingent on Δ.

Proposition 4.1. Suppose capacity is effectively unconstrained, both division managers can make specific investments, internal and external demands are described by linear functions as in Equation (4.3), and upstream variable costs are $C(\bar{q},\theta,I_1) = [c(\theta) - I_1]\bar{q}$. Introducing an internal discount then increases the expected firmwide profit.

With unadjusted market-based transfer pricing ($\Delta = 0$), Section 3 (Proposition 3.4) has shown that the decentralized firm suffers from quantity distortions *ex post.* As I shall show momentarily, this inefficiency will be compounded by *ex ante* underinvestment. Internal discounts serve to alleviate *both* these problems.

To provide intuition for the underinvestment problem at $\Delta = 0$, and how it is alleviated by setting $\Delta > 0$, note that the first-order conditions for the divisions' investment choices (holding fixed the respective other division's investment) are:

$$\frac{\partial \pi_1}{\partial I_1} = \mathbb{E}_\theta[\bar{q}(\theta, I_1(\Delta, I_2), I_2)] - w_1'(I_1(\Delta, I_2)) = 0,$$
$$\frac{\partial \pi_2}{\partial I_2} = \mathbb{E}_\theta\left[\left(1 - \frac{\partial p(\cdot)}{\partial I_2}\right) q_i(\Delta, \theta, I_1, I_2(\Delta, I_1))\right] - w_2'(I_2(\Delta, I_1)) = 0,$$

following similar logic as in Equation (2.12). A pure-strategy Nash equilibrium in investments, $I(\Delta)$, constitutes a solution to these two simultaneous equations. Suppose $\Delta = 0$, that is, the firm uses unadjusted market-based transfer pricing. Then, as shown formally in the proof of Proposition 4.1, the divisions will underinvest for the following reasons. The selling division is residual claimant for all variable cost savings resulting from investing I_1. Yet these savings apply to a total output quantity that is less than first-best, i.e., $\bar{q}(\theta, I) < \bar{q}^*(\theta, I)$, due to double marginalization on internally transferred units.[3] The buying division's investment incentives are further muted by the holdup term $\partial p/\partial I_2$, as the selling division will opportunistically exploit any increase in I_2 by raising the external price p and thereby indirectly the transfer price. Lastly, the fact that investments are strategic substitutes yields that $I_j(\Delta = 0) < I_j^*,\ j = 1, 2$.

The holdup problem afflicting buyer investments is of qualitatively similar nature as, but less severe than, that afflicting standard cost-based transfer pricing in Section 2. Given linear demand functions, the external price can be readily shown to be:[4]

$$p(\Delta, \theta, I_1, I_2) = p(0, \theta, 0, 0) + \nu(\theta)\Delta - \frac{1}{2} I_1 + \frac{\nu(\theta)}{2} I_2, \tag{4.4}$$

where $\nu(\theta) \equiv \frac{\beta_i(\theta)}{\beta_i(\theta) + \beta_e(\theta)}$ and $p(0, \theta, 0, 0) = \frac{1}{2}\left[\frac{\alpha_e(\theta) + \alpha_i(\theta)}{\beta_e(\theta) + \beta_i(\theta)} + c(\theta)\right]$ equals the external price the selling division would set in the absence of any

[3] This result again uses the fact that the total quantity sold by a monopolist who can freely price discriminate over linear demand functions equals the total output he would sell if he were constrained to charge a uniform price.

[4] To avoid corner solutions, assume that, for all θ and $I_j \in [0, \bar{I}_j],\ j = e, i$: $p(0, \theta, I) < \min\{a_e(\theta), a_i(\theta) + I_2\}$.

discount or investments. Thus, the seller's optimal price is additively separable. While $p(\cdot)$ is indeed increasing in I_2, note, however, that this holdup problem will be alleviated by the presence of external markets. If, say, the external demand for the intermediate good is highly price-sensitive (i.e., $\beta_e(\cdot)$ is high), then $\nu(\theta)$ becomes small. In that case, buyer investments are largely protected from holdup, because the seller would lose too much in terms of foregone external sales if he were to raise $p(\cdot)$ to expropriate the increased surplus generated by I_2. For comparison, the price reaction under standard-cost transfer pricing for linear demand functions in Section 2 was $\partial t^s/\partial I_2 = \frac{1}{2} > \partial p/\partial I_2 = \nu(\theta)/2$. Thus, by tying internal prices to the external price, holdup problems can be alleviated.

The last step to establishing Proposition 3.4 is to show that the Nash equilibrium $I(\Delta)$ is monotonically increasing in the internal discount. That is, by raising Δ, HQ can alleviate the underinvestment problem. Recall that with linear demand functions the total upstream production quantity $\bar{q}(\Delta, \theta, I)$ is invariant to Δ, *holding constant investments*. Since the selling division's investment incentives hinge on $\bar{q}(\cdot)$, there is no direct effect of Δ on I_1. The buying division's investment incentives, however, depend on the *internally* traded quantity, which is increasing in Δ. Hence, the buying division will invest more for higher discounts.[5] Lastly, recall that divisional investments are strategic complements. In particular, as I_2 increases, total quantity $\bar{q}(\cdot)$ will expand. As a consequence, the selling division will also invest more, indirectly, as it anticipates greater total output. This in turn will raise the buyer's investment incentives even further, as greater seller investment translates into lower (external and internal) prices at the margin.

To summarize, in the new equilibrium, both divisions will invest more as a result of the imposition of Δ. As the Appendix shows, imposing a small discount starting from $\Delta = 0$ has three positive first-order effects: (i) the double-marginalization bias is partially remedied (as in Proposition 3.4); (ii) the buying division benefits from a reduction in the internal price as a result of the selling division investing more; and

[5] Since, by Equation (4.4), $p(\cdot)$ is additively separable in Δ and I_j, $\partial p/\partial I_2$ is independent of Δ.

(iii) the selling division benefits from an increase in internal transfers as a result of the buying division investing more. That is, divisional investment opportunities make internal discounts more desirable from a corporate perspective.

The preceding model sketch constitutes a mere first step in the direction of a comprehensive model addressing relationship-specific investments in firms that are simultaneously involved in intrafirm and external sales of similar intermediate products. A number of open questions remain, such as the role of capacity constraints and divisions' incentives to invest in relaxing those constraints.

5

Internal Pricing with Asymmetric Information

The canonical incomplete contracting model is built on the assumption that at the transaction stage all information relevant for the trading decision is commonly known to both trading parties — here, the divisional managers — but not verifiable to third parties. As a result, when free to negotiate over the transaction, the managers realize all gains from trade, and the only remaining friction afflicting the organization is the holdup problem. This emphasis on *ex ante* frictions has been critiqued by Williamson (2000) and, more recently, by Hart and Moore (2008, p. 3) who argue that "to develop more general and compelling theories of contracts and organizational form it is essential to depart from a world in which Coasian renegotiation always leads to *ex post* efficiency."

The main motivation for this framework is that the complexity surrounding many transactions often precludes fully contingent contracts. In other words, the parties to the transaction may know the relevant data, yet the "burden of proof" is prohibitively high when facing an outside party such as a court. For intrafirm transactions, however, the contract-enforcing party is the firm's HQ. Arguably, the burden of proof then should be lower, as HQ can instruct the divisions by fiat to take certain actions even on the basis of somewhat imperfect information. Moreover, it is difficult to argue that the firm's accounting system provides complete information to the division managers about their own and the respective other division's data while leaving HQ in the dark. After all, it is HQ that has designed and installed the control system in the first place. Hence, while the symmetric-but-unverifiable information assumption is commonly accepted by now for the study of *inter-firm* transactions, it appears conceptually more problematic for transactions *within* firms.

The second key argument in favor of the symmetric information model is its tractability, thanks to the generalized Nash solution (as in Section 2). Bargaining under incomplete information on the other hand poses bigger modeling challenges. This section proposes a tractable sealed-bid bargaining approach to model negotiations under incomplete information, which can be integrated into a multistage investment/trading game of the kind studied above. The main goal is to put the two pricing mechanisms analyzed in Section 2 on more equal footing by accounting for the fact that negotiations, too, will in general fail to realize all gains from trade *ex post*. This section builds on Baldenius (2000).

5.1 Model with Asymmetrically Informed Managers

The model closely resembles that of Section 2, with two exceptions (see Figure 5.1):

- The random state variable θ is decomposed into two components, θ_1 and θ_2, with θ_1 parameterizing the supplying division's relevant cost and θ_2 parameterizing the buyer's

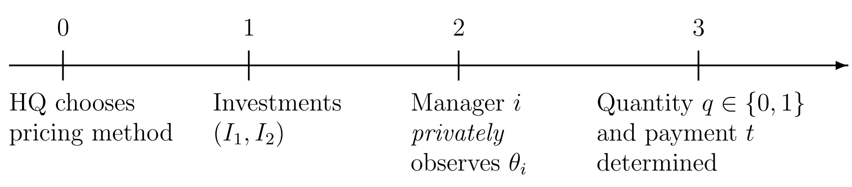

Fig. 5.1 Timeline with asymmetrically informed division managers.

relevant revenues. At Date 2, the manager of Division i *privately* observes his respective parameter θ_i, but not that of the other division, θ_j.
- The transfer decision at Date 3 is binary, $q \in \{0,1\}$; i.e., it involves a discrete project of fixed size.

The binary nature of trade allows for a simpler representation of the divisions' relevant cost and revenues. Specifically, the supplier's costs of manufacturing the goods and the buyer's relevant net revenues are given by:

$$C(\theta_1, I_1) \equiv \theta_1 - I_1, \quad R(\theta_2, I_2) = \theta_2 + I_2, \tag{5.1}$$

if $q = 1$, and zero otherwise. It is common knowledge that the "raw" cost and revenue parameters θ_1 and θ_2 are independently and uniformly distributed, i.e., $\theta_i \sim U[\underline{\theta}_i, \bar{\theta}_i]$, with densities $f_i(\theta_i) \equiv 1/\delta_i$ for $\delta_i \equiv \bar{\theta}_i - \underline{\theta}_i$, and $F_i(\theta_i)$ as the corresponding cumulative distributions. Investments thus move the relevant cost and revenue supports, in that the seller's *relevant cost* is uniformly distributed over $[\underline{\theta}_1 - I_1, \bar{\theta}_1 - I_1]$ whereas the buyer's *relevant revenue* is uniformly distributed over $[\underline{\theta}_2 + I_2, \bar{\theta}_2 + I_2]$. Let $\theta \equiv (\theta_1, \theta_2)$ and $I \equiv (I_1, I_2)$.

In line with Section 2, divisional investments I_i are unverifiable to HQ. Moreover, to motivate the delegation setting, for now I also assume that HQ does not know the exact divisional cost and revenue supports Θ_i. This specification seems descriptive for divisions that undertake many different transactions and investments at the same time. Hence HQ cannot therefore meaningfully intervene in the transaction other than by setting the transfer pricing policy. Section 5.5, below, will address the question whether HQ can do better by centralizing the process if it were to observe the type supports.

The First-Best Solution. From the viewpoint of the firm, the first-best solution calls for the transfer to be made *ex post* whenever the buyer's valuation exceeds the seller's cost:

$$q^*(\theta, I) = 1, \quad \text{if and only if} \quad \theta_2 + I_2 \geq \theta_1 - I_1. \tag{5.2}$$

Condition (5.2) characterizes the efficient trading rule HQ would implement, if it had complete information.[1] I will focus on the case where, *ex ante*, only one of the two divisions, say Division i, can undertake specific investments, while $\bar{I}_j = 0$. As a benchmark, let first-best expected firm profit as a function of I_i be denoted by

$$\pi^*(I_i) = M^*(I_i) - w_i(I_i), \quad i = 1, 2, \tag{5.3}$$

where $M^*(I_i) \equiv \mathbb{E}_\theta\left[(\theta_2 - \theta_1 + I_i)q^*(\theta, I_i)\right]$ is the firmwide expected contribution margin, provided the efficient quantity $q^*(\theta, I_i) = 1$ will later be traded at Date 3. Differentiating Equation (5.3) with respect to I_i, using the envelope theorem, yields the necessary first-order condition for efficient investments:

$$w_i'(I_i^*) = \text{Prob}\{q^*(\theta, I_i^*) = 1\}, \quad i = 1, 2. \tag{5.4}$$

Investments should be made up to the point where the marginal fixed cost equals the expected marginal benefit from investing. By Equation (5.1), the latter equals the equilibrium transfer probability.

The Decentralized Solution. Both managers are again assumed risk-neutral and motivated to maximize the expected profits of their own division.[2] Expected divisional profits consist of expected contribution margins, M_i, less the fixed cost from investing:

$$\begin{aligned}
\pi_1 &= M_1(\cdot) - w_1(I_1) \\
&= \int_{\Theta_1}\int_{\Theta_2} [t(\theta, I) - (\theta_1 - I_1)]q(\theta, I)\, dF_2(\theta_2) dF_1(\theta_1) \; - w_1(I_1) \\
\pi_2 &= M_2(\cdot) - w_2(I_2) \\
&= \int_{\Theta_1}\int_{\Theta_2} [(\theta_2 + I_2) - t(\theta, I)]q(\theta, I)\, dF_2(\theta_2) dF_1(\theta_1) \; - w_2(I_2)
\end{aligned}$$

for the selling and the buying division, respectively. The functions $\langle q(\cdot), t(\cdot)\rangle$ will be determined by the respective transfer pricing schemes,

[1] To avoid trivial decisions, I assume throughout that $0 < \mathbb{E}_\theta[q^*(\theta, I)] < 1$, for all I. That is, the relevant cost and revenue supports $[\underline{\theta}_1 - I_1, \overline{\theta}_1 - I_1]$ and $[\underline{\theta}_2 + I_2, \overline{\theta}_2 + I_2]$ intersect for any I.

[2] Kanodia (1991) studies the role of rent extraction in transfer pricing models with asymmetrically informed managers.

with attention confined to the two regimes studied in Section 2: (standard) cost-based and negotiated pricing.

Standard-cost Transfer Pricing. As in Section 2, Manager 1 is assumed to enjoy cost reporting discretion under this method. After observing his cost parameter θ_1, he reports a standard cost number, t^s, which then becomes the transfer price. This cost quote is assumed not to be audited by a corporate controller (or such audits are ineffective).[3] Since we are now dealing with a discrete trading quantity, $q \in \{0,1\}$, the buying division will accept the offer whenever $\theta_2 + I_2 \geq t^s$, which happens with probability $1 - F_2(t^s - I_2)$. Assuming that Manager 1 is not constrained at all in his cost reporting, he will act as a profit-maximizing monopolist who faces a customer with an uncertain valuation for the good:

$$t^s(\theta_1, I) = \arg\max_t \left\{(t - \theta_1 + I_1)[1 - F_2(t - I_2)]\right\}.$$

Given that θ_2 is uniformly distributed over $[\underline{\theta}_2, \overline{\theta}_2]$, differentiating this expression with respect to t yields the first-order condition:

$$t^s(\theta_1, I) = \begin{cases} \underline{\theta}_2 + I_2, & \text{if } \theta_1 - I_1 < 2\underline{\theta}_2 - \bar{\theta}_2 + I_2, \\ \frac{1}{2}(\bar{\theta}_2 + \theta_1 - I_1 + I_2), & \text{if } \theta_1 - I_1 \in [2\underline{\theta}_2 - \bar{\theta}_2 + I_2, \bar{\theta}_2 + I_2], \\ \theta_1 - I_1, & \text{if } \theta_1 - I_1 > \bar{\theta}_2 + I_2. \end{cases} \tag{5.5}$$

Like a monopolist selling to a buyer with unknown reservation price (e.g., Maskin and Riley, 1984), Manager 1 marks up his true production cost of $\theta_1 - I_1$ by an amount that trades off his contribution margin with the risk of refusal by Manager 2. The resulting trading rule under standard-cost transfer pricing reads:

$$q^s(\theta, I) = 1, \quad \text{if and only if} \quad \theta_2 + I_2 \geq t^s(\theta_1, I). \tag{5.6}$$

Again, this formulation of standard-cost transfer pricing is extreme in ignoring any reporting constraints for the selling division; Section 5.5 will relax this assumption.

[3] Again one could amend the model by adding *ex post* cost shocks so that actual costs are given by $\theta_1 - I_1 + \varepsilon$, with ε being a random variable with zero mean. Then the seller would quote t^s before observing ε, i.e., without precise knowledge of his actual cost. That is, the seller would be *better* informed than HQ, but not perfectly informed. This seems to correspond closely to the definition of standard costing found in textbooks.

Negotiated Transfer Pricing. Alternatively, HQ can let the managers negotiate over the transfer, which essentially yields a more symmetric allocation of bargaining power. I model the bargaining process as an *equal-split sealed-bid (double auction) mechanism*, following Chatterjee and Samuelson (1983). Both managers submit sealed bids, and trade occurs if and only if the buyer's bid, b, exceeds the seller's ask, s. In that case, the surplus is split equally: $t^n = \frac{1}{2}(b+s)$. If $b < s$, then $q^n = t^n = 0$.[4] A Bayesian–Nash equilibrium in linear bidding strategies then is the solution to the following simultaneous optimization problems:

$$s(\theta_1, I) \in \arg\max_s \mathbb{E}_{\theta_2}\left[\left(\frac{s + b(\theta_2, I)}{2} - \theta_1 + I_1\right) q^n(\theta, I)\right], \tag{5.7}$$

$$b(\theta_2, I) \in \arg\max_b \mathbb{E}_{\theta_1}\left[\left(\theta_2 + I_2 - \frac{s(\theta_1, I) + b}{2}\right) q^n(\theta, I)\right], \tag{5.8}$$

where

$$q^n(\theta, I) = 1 \quad \text{if and only if} \quad b(\theta_2, I) \geq s(\theta_1, I). \tag{5.9}$$

Put yourself in the shoes of the selling manager, facing the optimization problem in Equation (5.7). By increasing your ask s by \$1, the transfer price realized if the transaction takes place increases by \$0.5, because $t^n = \frac{1}{2}[s+b]$. At the same time, however, the probability that the buyer's bid b exceeds s decreases, i.e., a successful transaction becomes less likely. In the optimal solution, the two effects just balance each other. The resulting Bayesian–Nash equilibrium in linear bidding strategies with upfront investments I can be described as follows:[5]

Lemma 5.1. Suppose divisional costs and revenues are as in Equation (5.1) for given investments I, and $\theta_i \sim U[\underline{\theta}_i, \bar{\theta}_i]$. Then the equal-split sealed-bid mechanism yields the following linear equilibrium

[4] See Baiman and Rajan (2002) for an alternative approach to bargaining with asymmetric information and specific investments.

[5] Leininger et al. (1989) derive other equilibria under this bargaining mechanism but show that the linear equilibrium that corresponds to that described in Lemma 5.1 is the unique polynomial one. Experimental evidence suggests that this equilibrium describes participants' bargaining behavior rather well (Radner and Schotter, 1989).

bidding strategies:

$$s(\theta_1, I) = \begin{cases} \tilde{b}(\underline{\theta}_2, I) = \frac{\bar{\theta}_2 + 3\underline{\theta}_1 - 3I_1 + 9I_2 + 8\underline{\theta}_2}{12}, & \text{if } \tilde{s}(\theta_1, I) < \tilde{b}(\underline{\theta}_2, I) \\ \tilde{s}(\theta_1, I) = \frac{3\bar{\theta}_2 + \underline{\theta}_1 - 9I_1 + 3I_2 + 8\theta_1}{12}, & \text{if } \tilde{s}(\theta_1, I) \in [\tilde{b}(\underline{\theta}_2, I), \tilde{b}(\bar{\theta}_2, I)] \\ \theta_1 - I_1, & \text{if } \tilde{s}(\theta_1, I) > \tilde{b}(\bar{\theta}_2, I) \end{cases} \quad (5.10)$$

$$b(\theta_2, I) = \begin{cases} \tilde{s}(\bar{\theta}_1, I) = \frac{3\bar{\theta}_2 + \underline{\theta}_1 - 9I_1 + 3I_2 + 8\bar{\theta}_1}{12}, & \text{if } \tilde{b}(\theta_2, I) > \tilde{s}(\bar{\theta}_1, I) \\ \tilde{b}(\theta_2, I) = \frac{\bar{\theta}_2 + 3\underline{\theta}_1 - 3I_1 + 9I_2 + 8\theta_2}{12}, & \text{if } \tilde{b}(\theta_2, I) \in [\tilde{s}(\underline{\theta}_1, I), \tilde{s}(\bar{\theta}_1, I)] \\ \theta_2 + I_2, & \text{if } \tilde{b}(\theta_2, I) < \tilde{s}(\underline{\theta}_1, I) \end{cases} \quad (5.11)$$

Let us examine this bidding equilibrium in more detail. The top branches of the respective strategies describe the situation where trade occurs with certainty. Manager 1, say, upon observing a favorable (low) cost realization θ_1 will never bid less than the lowest equilibrium bid of Manager 2, as that would only reduce the transfer price without further raising the probability of trade. The middle branches of Equations (5.10) and (5.11) apply to "moderate" realizations of θ_i, conditional on which the probability of the transfer taking place is strictly positive but less than one. In that case, the double auction-type tradeoff described above arises. The bottom branches of Equations (5.10) and (5.11) describe situations where the conditional probability of trade is zero because of very unfavorable type realizations (high- θ_1 or low- θ_2).[6]

While the following analysis of negotiated pricing is confined to the simultaneous-move sealed-bid mechanism, there are alternative

[6] In this asymmetric information model, standard-cost pricing can be reinterpreted as yet another sealed-bid mechanism, where Manager 2 is "pathologically honest" and Manager 1 optimizes accordingly. More specifically, both regimes can be nested in a more general mechanism in which the seller asks s, the buyer bids b, and the price equals $t = zs + (1 - z)b$, if $s \leq b$. While negotiations represent the case of $z = 1/2$, we would obtain standard-cost pricing for $z = 1$. The weight, z, attached to the seller's ask, can be viewed as a measure of the seller's bargaining power. In the symmetric information setting of Section 2, in contrast, standard-cost pricing was not a limit case of negotiations. There, the seller had the power to make a take-it-or-leave-it offer but was restricted to uniform pricing, which for continuous transfer quantities gave rise to distortions. No such distortions were present under negotiations where Coasian bargaining realized all gains from trade.

approaches to modeling bargaining under asymmetric information; see Ausubel et al. (2002) for details. Adopting a mechanism design approach, Myerson and Satterthwaite (1983) show that no fully efficient, incentive compatible, individually rational, and budget-balanced mechanism exist for nontrivial trading decisions (their model does not allow for *ex ante* investments, and hence the effective type supports are exogenous). That is, the first-best trading rule in Equation (5.2) cannot be implemented in a delegated fashion. Instead, the *second-best* solution maximizes the overall surplus while observing individual rationality and incentive compatibility. The equal-split sealed-bid mechanism achieves this second-best benchmark performance, but only in the special case where the relevant costs and revenues are uniformly distributed over identical supports. With relevant costs and revenues endogenously determined by divisional investments, the sealed-bid mechanism will be second-best only in knife-edge cases. At the same time, however, the mechanism design approach is ill-suited for the study of specific investments, because it characterizes implementable allocations and total surplus, while leaving *individual* divisional payoffs indeterminate — and it is the latter that determine *ex ante* investment incentives.[7]

The internal pricing mechanism again has to deal with the dual incentive problems of eliciting *ex ante* investments and *ex post* efficient transfer decisions. Holding investments fixed for now, I first focus on the latter. Unlike in the symmetric information setting, the *ex post* performance comparison is nontrivial as now both mechanisms suffer from undertrade. To avoid tedious case discussions, the following assumption will be useful

---

Assumption 1. $\bar{\theta}_2 \geq \bar{\theta}_1 \geq \frac{1}{4}[3\bar{\theta}_2 + \underline{\theta}_1 + 3(\bar{I}_1 + \bar{I}_2)]$ and $\underline{\theta}_1 \leq \underline{\theta}_2 \leq \frac{1}{4}[\bar{\theta}_2 + 3\underline{\theta}_1 - 3(\bar{I}_1 + \bar{I}_2)]$. Moreover, the total firmwide investment $\bar{I}_1 + \bar{I}_2$ is sufficiently small so that $[\underline{\theta}_1 - I_1, \bar{\theta}_1 - I_1]$ and $[\underline{\theta}_2 + I_2, \bar{\theta}_2 + I_2]$ intersect for all (I_1, I_2).

---

[7] Ausubel et al. (2002) also provide an overview of attempts at modeling alternating-offers bargaining with bilateral private information. Due to the complexity associated with belief revisions under sequential bargaining, the multiple-equilibria problem is vastly exacerbated as compared with the (static) sealed-bid mechanism.

Assumption 1 ensures that there do not exist any cost and revenue realizations for which a successful transfer is ensured. It generalizes the setting in Myerson and Satterthwaite (1983) who have assumed that $\bar{I}_i = 0$, $\bar{\theta}_1 = \bar{\theta}_2$, and $\underline{\theta}_1 = \underline{\theta}_2$.[8]

An evaluation of the respective efficiency losses under the two-candidate pricing regimes amounts to comparing the respective expected firmwide contribution margins for given investments, denoted by $M^k(I) = E_\theta[(\theta_2 - \theta_1 + I_1 + I_2)q^k(\theta, I)]$, for $k = n, s$. Under negotiations,

$$M^n(I) = \int_{\underline{\theta}_1}^{\theta_1^n(I)} \int_{\theta_2^n(\theta_1|I)}^{\bar{\theta}_2} [\theta_2 - \theta_1 + I_1 + I_2] dF_2(\theta_2) dF_1(\theta_1),$$

with $\theta_2^n(\theta_1 \mid I)$ as the revenue cutoff implicitly defined by $s(\theta_1, I) \equiv b(\theta_2^n(\theta_1 \mid I), I)$ for given I, and $\theta_1^n(I)$ defined such that $\theta_2^n(\theta_1^n(I) \mid I) \equiv \bar{\theta}_2$. Thus, if the selling division observes a θ_1-realization in excess of $\theta_1^n(I)$, it knows with certainty that the transfer will not take place. The corresponding expected firmwide surplus under standard-cost pricing equals

$$M^s(I) = \int_{\Theta_1} \int_{t^s(\theta_1, I) - I_2}^{\bar{\theta}_2} [\theta_2 - \theta_1 + I_1 + I_2] dF_2(\theta_2) dF_1(\theta_1).$$

Comparing these two expressions yields (see the Appendix for details):

Lemma 5.2. Suppose Assumption 1 holds and divisional costs and revenues are as described in Equation (5.1). Taking as given I, $M^s(I) < M^n(I) < M^*(I)$.

As noted above, Myerson and Satterthwaite (1983) have shown that the sealed-bid mechanism achieves second-best performance in the special case where $\underline{\theta}_1 - I_1 = \underline{\theta}_2 + I_2$ and $\bar{\theta}_1 - I_1 = \bar{\theta}_2 + I_2$. In our setting with endogenous cost and revenue supports, bargaining falls short of this benchmark. Yet as Lemma 5.2 demonstrates, this mechanism still ensures *more* efficient trade than a system that gives one

[8] More technically, Assumption 1 implies that under negotiations only the lower two branches of Equations (5.10) and (5.11), respectively, arise with positive probability. It corresponds to the case of "Bilateral Uncertainty (BU)" in Baldenius (2000).

division — here: the seller — the power to make a take-it-or-leave-it offer. The reason is that the latter system is *too rigid* in that the transfer price t^s does not reflect different revenue realizations at all. That is, even though θ_2 enters the negotiated price t^n in a biased way, due to Manager 2 shading his bid, the firm still wants this information to affect the resource allocation.

I now add relationship-specific investments to the picture, but confine attention to unilateral investments for reasons of tractability. With only Division i investing, denote expected firmwide profit under transfer pricing policy $k,\ k = n, s$, by $\pi^k \equiv \pi^k(I_i^k)$, where $\pi^k(I_i) = M^k(I_i) - w_i(I_i)$ with I_i^k as the equilibrium investment under this policy.

5.2 Investment by the Supplying Division

For now suppose only the supplying division has an investment opportunity, i.e., $I_1 \geq \bar{I}_2 = 0$. The relevant cost and revenues then are $C(\theta_1, I_1) = \theta_1 - I_1$ and $R(\theta_2) = \theta_2$. As in the symmetric information setting of Section 2, divisional investment incentives will be determined by (i) the expected firmwide investment returns and (ii) how these returns are split between the two divisions at the margin. As to (i), Lemma 5.2 suggests that the firmwide investment returns may be higher under negotiations, as trade is more efficient.[9] As regards (ii), however, upstream investments under negotiations are again subject to a holdup problem. To see this more specifically, note that by Equation (5.11) the buying manager will bid *more aggressively*, i.e., lower his bid b, upon observing higher seller investments. With the transfer price being the average of the seller and buyer bids, this translates into a lower t for any ask s submitted by the seller. Under standard-cost pricing, on the other hand, all pricing power rests with the selling division, which is therefore protected from any holdup. As in Section 2, it is thus unclear *a priori* which of the two mechanisms generates better upstream investment incentives.

[9] This statement is not entirely precise, however, since Lemma 5.2 refers to the expected *contribution margin* under regime k, $M^k = \int_{\Theta_1}\int_{\Theta_2}[\theta_2 - \theta_1 + I_1]q^k(\theta, I_1)dF_2(\theta_2)dF_1(\theta_1)$, whereas expected firmwide marginal investment returns are given by the expected *trading probability*, $\int_{\Theta_1}\int_{\Theta_2} q^k(\theta, I_1)dF_2(\theta_2)dF_1(\theta_1)$.

Under the pricing regime $k = s, n$, the seller invests according to:

$$I_1^k \in \arg\max_{I_1} \left\{ \mathbb{E}_\theta[(t^k(\theta_1, I_1) - \theta_1 + I_1)q^k(\theta, I_1)] - w_1(I_1) \right\}. \quad (5.12)$$

Consider first standard-cost transfer pricing. At Date 3, $t^s(\cdot)$ is chosen by the seller in such a way that it maximizes the bracketed term in Equation (5.12) in expectation over θ_2, for any cost realization θ_1. By the envelope theorem, therefore, the necessary first-order condition is:

$$w_1'(I_1^s) = \text{Prob}\{q^s(\theta, I_1^s) = 1\}. \quad (5.13)$$

The seller invests up to the point where the marginal investment cost equals the expected marginal cost savings. Comparing Equation (5.13) with the first-best investment condition in Equation (5.4), however, shows that these cost savings will not always be realized as $\text{Prob}\{q^s(\theta, I_1) = 1\} \leq \text{Prob}\{q^*(\theta, I_1) = 1\}$, for all I_1. Anticipated trade distortions hence again diminish upfront investment incentives.

The divisions' bidding strategies under negotiations are obtained by setting $I_2 = 0$ in Lemma 5.1. Adapting Equation (5.12) to negotiations and plugging in the relevant *ex post* trading rule as stated in Equation (5.9), the seller will choose I_1 so as to maximize

$$\int_{\underline{\theta}_1}^{\theta_1^n(I_1)} \int_{\theta_2^n(\theta_1|I_1)}^{\bar{\theta}_2} \left[\frac{b(\theta_2, I_1) + s(\theta_1, I_1)}{2} - \theta_1 + I_1 \right] \times dF_2(\theta_2) dF_1(\theta_1) \; - w_1(I_1).$$

As noted above, Manager 2's bidding strategy, $b : \Theta_2 \times [0, \bar{I}_1] \to \mathbb{R}_+$, is contingent on seller investments in such a way that greater I_1 will translate into more aggressive bidding by the buyer. More specifically, by Equation (5.11), $\partial b / \partial I_1 = -\frac{1}{4}$, for each θ_2, I_1, such that the conditional probability of trade is strictly positive. Using this fact, Manager 1's investment choice can be described by the necessary first-order condition (see Appendix, proof of Proposition 5.3, for details):

$$w_1'(I_1^n) = \frac{3}{4} \text{Prob}\{q^n(\theta, I_1^n) = 1\}. \quad (5.14)$$

Ranking the seller's equilibrium investments under the candidate pricing regimes thus requires evaluating the right-hand sides of

Equations (5.13) and (5.14). As shown in the Appendix, for any I_1,

$$\text{Prob}\{q^n(\theta, I_1^n) = 1\} > \text{Prob}\{q^s(\theta, I_1^n) = 1\} > \frac{3}{4}\text{Prob}\{q^n(\theta, I_1^n) = 1\}. \tag{5.15}$$

As a consequence, holdup problems are important enough to outweigh the higher trading probability under negotiations, with the result that $I_1^s \geq I_1^n$. Together with Lemma 5.2, this implies that HQ faces a tradeoff when choosing the pricing regime:

Proposition 5.3. Suppose Assumption 1 holds, divisional costs and revenues are as in Equation (5.1), and only the supplying division invests. Then the seller invests more *ex ante* under standard-cost transfer pricing, but trade is conditionally more efficient *ex post* under negotiations. Thus, the profit comparison is indeterminate in general.[10]

Without investment opportunities for the buying division, the performance comparison of the candidate regimes ultimately depends on the specific functional form of the seller's fixed costs function, $w_1(\cdot)$. If seller investments are essential, then standard-cost pricing will be preferred, and vice versa.[11]

This result illustrates once more the qualitative difference between the symmetric and asymmetric information settings. If only the seller invests, then assigning more bargaining power to the seller (i.e., raising η in Section 2) was unambiguously efficiency-enhancing; it alleviated the underinvestment problem and, meanwhile, negotiations were fully efficient *ex post* for any η. The optimal allocation of bargaining power in this case would have been to set $\eta = 1$, i.e., to make the seller residual claimant, say by granting him the right to charge a two-part tariff. With asymmetric information, in contrast, the allocation of bargaining

[10] In Baldenius (2000, Proposition 4), I evaluate the tradeoff between investment and trading efficiency for the simple case of a quadratic investment cost function and low cost uncertainty (i.e., $\delta\theta_1$ is small). The trade effect can then be shown analytically to dominate the investment effect, so that negotiations come out as the preferred regime.

[11] Suppose $w_j'(0) = 0$. If $w_j'(\cdot)$ increases in I_j at a slow rate, investments are "essential" in that the equilibrium investment will be high (subject to the bound in Assumption 1). On the other hand if $\lim_{I_j \to 0} w_j'(I_j) = \infty$, then investment by Division j is negligible.

power affects not just the *split* of the available surplus but also the *size* of the surplus itself, and by Lemma 5.2 the equal-split mechanism ensures greater trading efficiency than a lopsided mechanism.

5.3 Investment by the Buying Division

Now suppose that there is no scope for seller investments, but the buying division may raise its net revenues by investing, i.e., $I_2 \geq \bar{I}_1 = 0$. Accordingly, I will refer to the induced "relevant revenue distribution" defined over $\Theta_2 \times [0, \bar{I}_2]$ with support $[\underline{\theta}_2 + I_2, \bar{\theta}_2 + I_2]$ for given I_2.

Under either of the two pricing mechanisms, the non-investing party (the seller) has some, or all, bargaining power at the trading stage. Put differently, the seller's bidding/cost reporting strategy will always depend on I_2, which gives rise to holdup problems under *either* mechanism. (Recall that the buyer's acceptance decision when faced with a take-it-or-leave-it price under standard-cost pricing, in contrast, was independent of the seller's investment. Hence, seller investments were protected from holdup under this regime.) In fact, as the seller holds all pricing power under standard-cost pricing, one would expect this regime to fare particularly poorly for buyer investments, essentially for the same reasons as in Section 2.3 above.

Manager 1's reporting strategy under standard-cost pricing is found by setting $I_1 = 0$ in Equation (5.5). Specifically, Manager 1 will submit a cost quote of $t^s(\theta_1, I_2) = \frac{1}{2}(\bar{\theta}_2 + \theta_1 + I_2)$ whenever the conditional probability of trading is greater than zero but less than one. Thus, $\partial t^s / \partial I_2 = \frac{1}{2}$. In light of this severe holdup problem, it is easy to show that the buyer's investment satisfies the first-order condition:

$$w_2'(I_2^s) \;=\; \frac{1}{2}\,\mathrm{Prob}\{q^s(\theta, I_2^s) = 1\}. \tag{5.16}$$

Under negotiations, the divisions' bidding strategies are obtained by setting $I_1 = 0$ in Lemma 5.1. Due to the symmetry of the equal-split sealed-bid mechanism, the buying division's investment incentives under this policy are determined in a similar fashion as those of the selling division in Equation (5.14), i.e., $w_2'(I_2^n) = \frac{3}{4}\,\mathrm{Prob}\{q^n(\theta, I_2^n) = 1\}$. Comparing this with Equation (5.16), making use of the ranking of the conditional transfer probabilities in Equation (5.15), implies that

$I_2^n > I_2^s$. Together with Lemma 5.2, this implies that negotiations generate higher investments and more efficient trade *for any given level of investment*. To complete the profit comparison, thus, only a bit more structure on the profit function under negotiations is required.

Assumption 2. The function

$$\begin{aligned}\pi^n(I_2) &\equiv M^n(I_2) - w_2(I_2)\\ &= \int_{\Theta_1}\int_{\Theta_2}[\theta_2 + I_2 - \theta_1]q^n(\theta, I_2)\,dF_1(\theta_1)dF_2(\theta_2) - w_2(I_2)\end{aligned}$$

is single-peaked in I_2 with an interior maximizer $\hat{I}_2^n \in (0, \bar{I}_2)$.[12]

Assumption 2 ensures that higher *ex ante* investment efficiency and higher *ex post* trading efficiency together translate into greater expected profit under negotiations.

Proposition 5.4. Suppose Assumptions 1 and 2 hold, divisional costs and revenues are as in Equation (5.1), and only the buying division invests. Then $\pi^n > \pi^s$.

Comparing Propositions 5.3 and 5.4 confirms for the case of privately informed divisional managers the intuition from Section 2 that negotiations perform better than standard-cost pricing if the buyer has essential investment opportunities. The driving force behind this logic is the fact that negotiations defer more bargaining power to the buyer and, at the same time, remain more efficient as a mechanism to aggregate dispersed information into internal prices.

5.4 Limit Cases of One-Sided Uncertainty

The preceding analysis was based on the assumption of non-trivial cost and revenue uncertainty. While this may describe many situations in

[12] A sufficient condition for Assumption 2 to be satisfied is that $w_2''(I_2) > \frac{27}{32}(\bar{\theta}_2 + I_2 - \underline{\theta}_1)/(\delta_1\delta_2)$, for all I_2. Notice that single-peakedness of $\pi^n(I_2)$ is not implied by the maintained assumption that first-best profits $\pi^*(I_j) \equiv M^*(I_j) - w_j(I_j)$ are concave in I_j, for any j.

practice, at times information may be distributed asymmetrically. Take the example of a car manufacturer whose downstream (marketing) division introduces innovative car financing policies for the existing product line. In such a setting, the marketing division may have reasonably precise upstream cost information due to past negotiations with the upstream (manufacturing) division. At the same time, the manufacturing division may know little about the marketing division's sales forecast given the new financing policies. I will refer to this scenario as (one-sided) revenue uncertainty.

Conversely, (one-sided) cost uncertainty is a plausible description of a setting where the manufacturing division has undertaken a major overhaul of its processes for a given car model. The marketing division will then be uncertain about upstream costs whereas the manufacturing division may be able to assess downstream revenues reasonably precisely based on past negotiations. In Baldenius (2000) I provide formal definitions and analysis of these one-sided private information cases; here I only summarize the findings in a nontechnical manner.

Consider first the case of one-sided cost uncertainty, i.e., Manager 1 has precise knowledge of downstream revenues in the sense that $\theta_2 \sim U[\theta_2^o, \theta_2^o + \delta\theta_2]$ and $\delta\theta_2 \to 0$.[13] Allocating all bargaining power to the selling division by adopting standard-cost pricing then achieves conditionally efficient transfers *ex post*, because by pricing the input at $t^s = \max\{\theta_1 - I_1, \theta_2^o + I_2\}$, the seller can extract the entire surplus. Since the seller is fully protected from holdup under this regime, first-best performance will be attained, if only the seller can invest. The buyer, on the other hand, has no incentives to invest at all under this regime. Comparing this extreme no-investment outcome with the buyer's investment condition in Equation (5.16) illustrates that private information on the part of the investing party can provide partial protection from holdup, as the non-investing party has to use caution so as not to price (or bid) too aggressively. To see this, note that, conditional on $q^s = 1$, $\partial t^s/\partial I_2 = \frac{1}{2}$ if the seller is uncertain about the buyer's revenues, whereas $\partial t^s/\partial I_2 = 1$ if $\delta\theta_2 \to 0$.

---

[13] To ensure a nontrivial trade decision exists, assume feasible investment amounts are bounded such hat $\theta_2^o + I_2 \in [\underline{\theta}_1 - I_1, \overline{\theta}_1 - I_1]$ for any I.

Now consider negotiations, still under the assumption of one-sided cost uncertainty (due to the symmetry of the bargaining protocol, similar logic applies to one-sided revenue uncertainty). For given investments and θ, the bidding strategies employed by the managers under negotiations as of Lemma 5.1 converge to

$$b(\theta_2^o, I) = \frac{3}{4}\theta_2^o - \frac{1}{4}I_1 + \frac{3}{4}I_2 \quad \text{and} \quad s(\theta_1, I) = \max\{b(\theta_2^o, I), \theta_1 - I_1\}, \tag{5.17}$$

as $\delta_2 \to 0$. Both managers then know the downstream revenues, yet some gains from trade are lost. Thus, for any I, $M^n(I) < M^s(I) \to M^*(I)$ in the limit, i.e., negotiations are dominated by standard-cost pricing in terms of trading efficiency. Essentially, the buyer's bargaining power under negotiations results in the managers settling on the buyer's reservation price less a discount, hence the buyer earns a strictly positive contribution margin, $\theta_2^o + I_2 - b(\theta_2^o, I_1)$, if $q = 1$. As shown in Baldenius (2000), for any I, the seller's investment incentives under negotiations are described by the same first-order condition as in Equation (5.14) above, whereas buyer investments are subject to an exacerbated holdup problem as expressed by the first-order condition $w_2'(I_2^n) = \frac{1}{2}\text{Prob}\{q^n(\theta, I_2^n) = 1\}$. Again, a comparison with Equation (5.14) illustrates the potential for private information to serve as an investment shield.[14]

The results imply the following qualitative "rules of thumb" for HQ's choice of an internal pricing regime:

(1) To minimize *ex post* trade distortions, HQ should choose the pricing regime that confers most bargaining power to the division that has the most information. This resonates well with (agency-theoretic) adverse selection models.
(2) If the divisions are equally uncertain about each other's costs and revenues, then firmwide holdup problems are minimized by allocating bargaining power to the division that has the most significant investment opportunities. This is in line with the key theme of property rights theory.

[14] The reader is referred to Baldenius (2000) for settings involving one-sided revenue uncertainty and bilateral investments.

(3) If the divisions have equally important investment opportunities, then firmwide holdup problems are minimized by allocating bargaining power to the division that has less information about the respective other division's cost/revenue. This reflects the above insight that holdup problems will be exacerbated, if the investing party's type is known by the other division (and that other division has some or all bargaining power).

Note that Points 1. and 3. have somewhat conflicting implications for divisional investment incentives. Suppose the setting is one of one-sided cost uncertainty, i.e., $\delta\theta_2 \to 0$. The selling division thus has better transfer-relevant information than the buying division. Suppose further that HQ changes the pricing regime and thereby transfers bargaining power from, say, the buying to the selling division. By Point 1., then, transfers become more efficient conditional on investments undertaken, in that the contribution margin $M(I)$ and $\mathrm{Prob}\{q(\cdot) = 1\}$ both increase. Anticipating this, the divisions should have *stronger* incentives to invest. At the same time, firmwide (aggregate) holdup problems will be *exacerbated*, by Point 3., because bargaining power is shifted from the buyer, who is very vulnerable to holdup because the seller knows his revenues, to the seller, whose investments are partially protected by his private cost information. As noted in Baldenius (2000), in general the holdup effect (Point 3.) dominates the trade efficiency effect (Point 1.) when it comes to determining divisional investments. That is, $I_1^s > I_1^n$ and $I_2^s < I_2^n$ for any informational scenario (bilateral, or one-sided cost or revenue uncertainty).

5.5 Centralized Transfer Pricing

The analysis in Section 5 was built on the assumption that HQ in general has less information about Division i's operations (on a disaggregated product line level) than has Division j, which may have a history of directly trading with Division i. When choosing a transfer pricing scheme, HQ was assumed not to have precise knowledge of the divisions' *ex ante* type supports Θ_1 and Θ_2. While this seems descriptive for large diversified companies, it is nonetheless interesting to ask what would change if HQ knew the divisions' type supports, Θ_i. In particular, should HQ intervene in the standard setting process in a centralized fashion, instead of relying on inflated cost/revenue reports?

The answer to this question is, in general, yes, as shown by Baldenius (2000) and Pfeiffer et al. (2009). Centralized intervention by HQ in form of setting the transfer price equal to the expected variable cost of the seller — anticipating his equilibrium investment and including a markup — can be shown to outperform decentralized standard-cost pricing with seller discretion when there is non-trivial cost and revenue uncertainty. That is, the benefits associated with avoiding the monopoly distortions outweigh the flexibility loss that arises because the centralized price, by construction, cannot react to the seller's cost realization, θ_1. The performance comparison between centralized pricing and negotiations, however, remains ambiguous: centralization tends to generate higher investments, but negotiations are more efficient *ex post*.

The optimal markup under centralization has a very different interpretation than the markup in the symmetric information model of Section 2.5. There, the markup was over *actual* cost and it was designed to trade off the attendant trade distortions with improved investment incentives for the seller; that is, without any scope for seller investments, the optimal markup should be zero. In the asymmetric information model studied here, in contrast, the optimal markup over *expected* cost trades off the risks of inefficient refusal by the two divisions and actually enhances trade efficiency. Thus, the markup should be nonzero even in the absence of investments.[15]

[15] Pfeiffer et al. (2009) establish a qualitatively different result in a continuous quantity model. They show that the optimal centralized transfer price should just equal the expected upstream cost anticipating the equilibrium seller investment without any markup. That difference in prediction stems from the different sources of trade inefficiency: double marginalization in Pfeiffer et al. (2009) versus trading off the risks of refusal by the two privately informed managers for a fixed-quantity transfer in Baldenius (2000).

5.6 Incentive Contracts

An understudied issue in the organization economics literature is the interplay between division managers' incentive contracts and their investment and trading incentives. Early incomplete contracting studies have taken the players to be owner-managed firms and thereby blurred the distinction between firms and managers; see the critique in Holmstrom (1999). Clearly, holdup problems could be avoided altogether by compensating managers based on firmwide income (the same holds for bargaining inefficiencies of the kind in Proposition 5.2). Yet, profit sharing of this kind seems rarely used in practice below the top management level, mostly for reasons related to risk.

Starting from the observation that holdup is an externality problem, Holmstrom and Tirole (1991) and Anctil and Dutta (1999) derive the optimal weight on firmwide versus divisional profit in division managers' compensation. Since divisions are usually involved in many transactions with different parties (inside and outside the firm), the more the firm uses profit sharing, the more it holds Manager i responsible for risk affecting *all* of Division j's transactions.[16] The optimal contract trades off compensation risk premia with reduced holdup problems. Note that while profit sharing improves *ex ante* investment incentives, the efficiency of *ex-post* transfers under symmetric information negotiations remains unaffected (first-best), by the Coase theorem.[17] Yet, a recurring theme in the management literature is that managers who face high-powered incentives tend to haggle too aggressively, thereby sacrificing valuable opportunities for the firm (e.g., Argyres, 1995). The sealed-bid bargaining model can be augmented to address this issue. In

[16] In this context it is important to differentiate between investments made from divisional accounts (as assumed in this monograph) and those that are *personally* costly, i.e., effort. Holmstrom and Tirole (1991) consider both types of investment. Anctil and Dutta (1999) show that optimal risk sharing may in fact result in negative weights attached to cross-divisional income as negotiations endogenously create positively correlated divisional income metrics.

[17] A conceptual problem arises when modeling bargaining between two managers, each of whom is paid a bonus coefficient of α_j times his divisional profit. To fix ideas, suppose $\alpha_1 > \alpha_2$. When the managers bargain under symmetric information, they have incentives to collude, because for every unit traded at the maximum admissible transfer price, say t^{max}, collectively they will earn "arbitrage" profits of $(\alpha_1 - \alpha_2)t^{max}$. This dysfunctional behavior is similar to that afflicting dual transfer pricing.

the following, I sketch a simplified version of the model in Baiman and Baldenius (2009, henceforth BB).[18]

Suppose divisional managers are compensated based on divisional income, π_j, and the *nonfinancial performance measure* $q \in \{0,1\}$, which indicates whether the transaction has been carried out. Paying discrete bonuses for completion of major projects (milestones) is in line with firms' use of "Balanced Scorecards" (Ittner et al. 1997, p. 239). Assuming linear contracts to be descriptive, Manager j's compensation then is described by

$$S_j = \alpha_j + \beta_j \pi_j + \gamma_j q, \quad j = 1,2.$$

I confine attention to negotiated transfer pricing based on the same equal-split sealed-bid mechanism described above. Thus, instead of choosing a transfer pricing method at Date 0 (in Figure 5.1), HQ now offers contracts, $\mathcal{S} = \{\alpha_j, \beta_j, \gamma_j\}_{j=1,2}$. For the sake of presentation, I take the contract "slope", $\beta_j \in (0,1)$, to be determined by exogenous factors, whereas in BB these slopes were endogenously determined by agency problems at the divisional level. Ignoring any such agency problems here, and assuming the managers can fully commit to the contract before learning their respective θ_i, the fixed salaries α_j can be adjusted so as to extract the entire surplus, i.e., $\alpha_j = -\mathbb{E}_\theta[\beta_j \pi_j + \gamma_j q_j]$. With β_j exogenously given and α_j thus pinned down, the remaining question is, how should HQ set the γ_j (henceforth, the "bonuses") to be paid out upon completion of the transaction?[19]

To address this question, consider the case where only Manager 1 can invest, i.e., $\bar{I}_2 = 0$, and each type support is the unit interval, i.e., $\theta_j \sim U[0,1]$, $j = 1,2$. At Date 3, the bargaining strategies in Equations (5.7) and (5.8) now have to be amended for the bonuses γ_j. A Bayesian–Nash bidding equilibrium in linear strategies then is a

[18] Baldenius (2006) uses a related model to show that vertically integrated firms optimally choose lower-powered incentives and thereby elicit more conciliatory bargaining and higher investments from their managers as compared with vertically related firms that trade under separate ownership.

[19] Note that lack of verifiability of transaction-related costs and revenues does not mean that the completion of transaction itself is nonverifiable. Take the example of an R&D project. The *filing* of a patent is a verifiable event, yet the *value* of the patent may be known only to the divisions.

solution to the following two simultaneous equations:

$$s(\theta_1, I_1 \mid \mathcal{S}) \\ \in \arg\max_s \mathbb{E}_{\theta_2}\left[\beta_1\left(\frac{s + b(\theta_2, I_1 \mid \mathcal{S})}{2} - \theta_1 + I_1 + \frac{\gamma_1}{\beta_1}\right) q(\theta, I_1 \mid \mathcal{S})\right],$$

$$b(\theta_2, I_1 \mid \mathcal{S}) \\ \in \arg\max_b \mathbb{E}_{\theta_1}\left[\beta_2\left(\theta_2 - \frac{s(\theta_1, I_1 \mid \mathcal{S}) + b}{2} + \frac{\gamma_2}{\beta_2}\right) q(\theta, I_1 \mid \mathcal{S})\right].$$

As before, $q(\cdot) = 1$ if and only if $b(\cdot) \geq s(\cdot)$, in which case $t(\theta, I_1 \mid \mathcal{S}) = \frac{1}{2}[s(\cdot) + b(\cdot)]$; otherwise $q(\cdot) = t(\cdot) = 0$. Note that the slopes β_j are simply scaling variables and, as such, affect the divisions' bidding strategies only through $\frac{\gamma_j}{\beta_j}$. It is therefore useful to conduct a change of variables and henceforth refer to the *scaled bonuses*, $\Gamma_j \equiv \frac{\gamma_j}{\beta_j}$, $j = 1, 2$, as HQ's choice variable to mediate the managers' incentives. Solving the simultaneous optimization problems yields the following linear equilibrium (stated here are only the "interior" solutions where the conditional trading probability is strictly positive but less than one; the corner solutions are analogous to Equations (5.10) and (5.11)):

$$s(\theta_1, I_1 \mid \mathcal{S}) = \frac{3 - 9(I_1 + \Gamma_1) + 3\Gamma_2 + 8\theta_1}{12},$$
$$b(\theta_2, I_1 \mid \mathcal{S}) = \frac{1 - 3(I_1 + \Gamma_1) + 9\Gamma_2 + 8\theta_2}{12}.$$

At Date 1, the selling manager's incentives to invest are described as follows[20]:

$$I_1(\mathcal{S}) \in \arg\max_{I_1} \mathbb{E}_\theta \{[t(\theta, I_1 \mid \mathcal{S}) - \theta_1 + I_1 + \Gamma_1]\, q(\theta, I_1 \mid \mathcal{S}) - w_1(I_1)\}.$$

The analysis is greatly simplified by the following *Unraveling Property*: as shown in BB, only the *total* scaled bonus, $\Gamma \equiv \Gamma_1 + \Gamma_2$, matters for investment and trading incentives and overall efficiency, not the individual Γ_j. More specifically, Manager 1's equilibrium investment $I_1(\Gamma)$ and the resulting trading outcome — as measured by the trading

[20] Note that the bonus coefficient β_1 applies also to the divisional fixed investment costs $w_1(I_1)$, and thus will not affect the first-order investment condition directly.

probability, $\text{Prob}\{q(I_1(\Gamma), \Gamma) = 1\}$, or the expected contribution margin, $M(I_1(\Gamma), \Gamma)$ — each are sufficiently described by Γ. To illustrate the intuition underlying the Unraveling Property, suppose HQ raises Manager 1's bonus γ_1, while lowering γ_2 by just enough to keep Γ constant. Manager 1 then becomes more eager to reach an agreement and hence will bid *less aggressively*. At the same time, Manager 2 will bid *more aggressively* in equilibrium, because (i) his implementation bonus has decreased and (ii) he takes advantage of the anticipated "softening" of Manager 1's bidding behavior. In equilibrium, these changes in the respective strategies cancel each other precisely. Hence, HQ's design problem at Date 0 boils down to choosing Γ so as to maximize overall efficiency (recall that HQ can extract the expected surplus by adjusting the α_j's):

$$\pi(\Gamma) \equiv \mathbb{E}_\theta \{[\theta_2 - \theta_1 + I_1(\Gamma)]q(\theta, I_1(\Gamma) \mid \Gamma)\} - w_1(I_1(\Gamma)). \tag{5.18}$$

Note that for $\Gamma_j = 0$ the model will collapse to that of negotiated transfer pricing in Section 5.2. There we found that firmwide performance suffers from both *ex post* undertrade and *ex ante* underinvestment. Introducing bonuses tied to project completion alleviates *both* these inefficiencies, as the next result shows:

Proposition 5.5. Starting from $\Gamma = 0$, introducing some positive Γ increases firmwide efficiency, i.e., $\pi'(\Gamma = 0) > 0$.

Raising Γ is akin to "inflating the pie" over which the division managers will bargain. This results in an increased probability of the transfer taking place and (at least for Γ not too high) greater expected gains from trade. The trading inefficiency identified in Proposition 5.2 thus will be mitigated. Furthermore, Manager 1 anticipates the probability of trade to go up, and hence is willing to invest more, i.e., $I_1(\Gamma)$ is an increasing function. Thus, starting from $\Gamma = 0$, an increase in Γ alleviates both investment and trading distortions. Myerson and Satterthwaite (1983) have shown that introducing a budget breaker into a bargaining problem between asymmetrically informed parties increases the *ex post* gains from trade.[21] Proposition 5.5 demonstrates that this logic extends *a fortiori* to settings where managers can undertake upfront investments: budget breaking then has the additional benefit of improving investment efficiency.

If the total (scaled) bonus Γ is set very high, however, the managers will carry out some transactions that may have negative value to the firm. There is some anecdotal evidence of excessive project implementation in response to firms introducing Balanced Scorecards (e.g., Ittner and Larcker, 2003), which is usually interpreted as managers opportunistically gaming the system. In our setting, however, excessive transfers at Date 3 are the price HQ may have to pay to remedy the underinvestment problem at Date 1. More precisely, BB show that the *optimal* Γ^* that maximizes (Equation (5.18)) from HQ's point of view trades off *ex ante* underinvestment with excessive transfers *ex post*. Suppose HQ were to set Γ so that, anticipating the equilibrium investments, *ex post* transfers would be conditionally efficient. The divisions would still underinvest then, because of the holdup problem (at the margin the other division will change its bidding strategy upon observing higher investment). Increasing Γ further therefore only causes a second-order loss *ex post,* but a first-order gain on *ex ante* investments. By the same logic, the optimal total bonus Γ will be higher with relationship-specific investments than in pure exchange settings such as the one studied in Myerson and Satterthwaite (1983).[22]

Bonuses tied to nonfinancial performance measures are thus of use not just to combat managerial time horizon problems (the "leading indicator" motive emphasized in the practitioners' literature), but also to foster cross-divisional cooperation. This is in line with recent empirical evidence, documented in Bouwens and van Lent (2007), that firms' use of nonfinancial metrics is positively correlated with proxies for the scope of intrafirm cooperation.

[21] Myerson and Satterthwaite (1983) demonstrate the welfare-enhancing effects of budget-breaking using a more general mechanism design approach.

[22] BB also consider a dynamic version of the model with learning-by-doing. They show that Γ^* displays features of "fast-tracking" in that, conditional on early implementation bonuses being earned, future bonuses will be higher (i.e., bonuses are positively serially correlated).

6

Conclusion

This monograph has aimed to describe a unified framework for the study of intrafirm pricing. The model presented here in various versions is based on the incomplete contracting paradigm, which has proven useful and versatile in a variety of economic environments. The shift toward this paradigm in the theory of internal pricing, and away from mechanism design theory, has helped generate a host of results that are of practical value to managers in choosing between alternative pricing mechanisms. Critics of the incomplete contracting paradigm have held that it rests on shaky game-theoretical foundations if players are fully rational.[1] Furthermore, incomplete contracting models invariably require judgment calls by researchers. For instance, what precisely is the nature of the limitations to contracting, i.e., what is assumed contractible, and what is not? How should one model standard-cost pricing — is the cost standard determined by HQ, or *de facto* by the better informed selling division? Etc.

On the other hand, the potential rewards appear to outweigh these problems. Mechanism design theory has been extremely successful

[1] See the exchange between Maskin and Tirole (1999) and Hart and Moore (1999).

conceptually in understanding the benchmark performance of trading mechanisms in a world where arbitrarily complex contracts can be written. The contractual solutions thus derived, however, often bear little resemblance with firm practice, as firms tend to use rather simple mechanisms. Incomplete contracting models that are tailored to capture the key stylized facts of real-world mechanisms can help practitioners pick the appropriate pricing regime and implement it, e.g., with or without cost-plus markups or internal discounts relative to market prices. Having said this, researchers should aim to assess not just which among predefined candidate regimes performs best, but also search for environments in which a particular mechanism does well in *absolute* terms, say, by replicating first-best performance (e.g., Propositions 3.1 and Corollary 3.2, above).

This study has omitted a number of important aspects of the transfer pricing problem. As argued in Section 5.6, the role of divisional managers, as distinct from owner–managers, deserves more attention. Division managers are not residual claimants, yet most of the incomplete contracting literature has treated them as such by following the tradition of property rights theory modeling. Yet, as Holmstrom (1999, p. 100) argues, property rights theory "says very little about the firm ... [because] ... there are really no firms in these models, just representative entrepreneurs." Division managers are employed by principals, and these principals can design incentive contracts to regulate trade among these managers. Understanding the interplay between compensation contracts and investment and trading incentives appears crucial to further our understanding of the economics of divisionalized firms.

There are a number of additional promising avenues for future research in the area of internal pricing. The issue of taxation is a vexing one for multinationals, in particular as it relates to hard-to-value intangible assets. An interesting first step in this agenda has been made by Johnson (2006). Another limitation of the model(s) presented here is their static nature, whereas divisions interact dynamically over time. A recent paper that applies insights from the theory of capacity investments (e.g., Rogerson, 2008) to internal pricing is Dutta and Reichelstein (2009). Lastly, the competitive environment in which the firm operates has been modeled somewhat rudimentarily in this study. In particular, it was assumed here that parties outside the firm are passive and do not (or cannot) react in any way to investments made by the divisions. The models presented here have also been silent on the issue of downstream competition. Arya et al. (2008) develop a model that couches the firm more explicitly in the competitive context. It would be desirable if future research could develop a comprehensive theory of how alternative pricing mechanisms perform in settings that account for various external market environments.

A

Appendix

Proof of Proposition 2.1. First note that a necessary condition for standard-cost transfer pricing to dominate negotiated transfer pricing is that $I_1^s > I_1^n$. Suppose that $I_1^s \leq I_1^n$, yet $\pi^s \geq \pi^n$. This yields a contradiction, because the expected firm profit under standard-cost transfer pricing is:

$$\begin{aligned}
\pi^s &= \mathbb{E}_\theta \left[R(q^s(\theta, I_1^s), \theta) - C(q^s(\theta, I_1^s), \theta, I_1^s)\right] - w_1(I_1^s) \\
&< \mathbb{E}_\theta \left[R(q^*(\theta, I_1^s), \theta) - C(q^*(\theta, I_1^s), \theta, I_1^s)\right] - w_1(I_1^s) \\
&\leq \mathbb{E}_\theta \left[R(q^*(\theta, I_1^n), \theta) - C(q^*(\theta, I_1^n), \theta, I_1^n)\right] - w_1(I_1^n) \\
&= \pi^n.
\end{aligned}$$

The first inequality follows by definition of $q^*(\cdot)$, which by assumption is unique and interior. The second inequality follows from the assumption that $\pi^*(I_1)$ is monotone increasing on $[0, I_1^*]$ and that, by assertion, $I_1^s \leq I_1^n$. Overinvestment under either transfer pricing scheme can be ruled out by the revealed preference arguments made below.

Next, I show that:

$$q^s(\theta, I) \leq \frac{1}{2} q^*(\theta, I), \text{ if and only if } r'(q, \theta) \text{ is convex in } q. \tag{A.1}$$

The first-best quantity $q^*(\theta, I)$ maximizes $\{r(q,\theta) - C(q,\theta,I_1)\}$ and hence

$$r'(q^*(\theta,I_1),\theta) - [c(\theta) - I_1] = 0. \tag{A.2}$$

At the same time, $q^s(\theta,I)$ maximizes $\{r'(q,\theta)q - C(q,\theta,I_1)\}$, since the buyer's marginal revenue curve represents his demand function for the input, and thus

$$r''(q^s(\theta,I_1),\theta)q^s(\theta,I_1) + r'(q^s(\theta,I_1),\theta) - [c(\theta) - I_1] = 0. \tag{A.3}$$

Subtracting and rearranging the first-order conditions (A.2) and (A.3), we have

$$\int_{q^s(\theta,I_1)}^{q^*(\theta,I_1)} r''(u,\theta)\,du = r''(q^s(\theta,I_1),\theta)q^s(\theta,I_1). \tag{A.4}$$

Convexity of $r'(q,\theta)$ means that $r''(q,\theta)$ is increasing in q, which provides a lower bound on the left-hand side of the preceding equality:

$$\begin{aligned}\int_{q^s(\theta,I_1)}^{q^*(\theta,I_1)} r''(u,\theta)\,du &> \int_{q^s(\theta,I_1)}^{q^*(\theta,I_1)} r''(q^s(\theta,I_1),\theta)\,du \\ &= r''(q^s(\theta,I_1)[q^*(\theta,I_1) - q^s(\theta,I_1)].\end{aligned}$$

Using this for Equation (A.4), and dividing by $r''(q^s(\theta,I_1) < 0$, then confirms Equation (A.1).

It remains to be shown that, if $\eta \geq \frac{1}{2}$ and $r'(q,\theta)$ is convex in q, then $I_1^n \geq I_1^s$ indeed holds. Let $M_1^n(I_1) = \eta\mathbb{E}_\theta[M^*(\theta,I_1)]$ denote the expected contribution margin accruing to Division 1 under negotiated transfer pricing, given it has invested an amount I_1. Similarly, under standard-cost transfer pricing:

$$M_1^s(I_1) = \mathbb{E}_\theta[t^s(\theta,I_1)q(t^s(\theta,I_1),\theta) - [c(\theta) - I_1]q(t^s(\theta,I_1),\theta)].$$

Using the envelope theorem, the respective marginal investment returns are $M_1^{n'}(I_1) = \eta\mathbb{E}_\theta[q^*(\theta,I_1)]$ and $M_1^{s'}(I_1) = \mathbb{E}_\theta[q^s(\theta,I_1)]$. If $r'(q,\theta)$ is convex in q, then, by Equation (A.1), we have $M_1^{n'}(I_1) \geq M_1^{s'}(I_1)$, for all I_1. At the same time, by revealed preference,

$$\begin{aligned}M_1^n(I_1^n) - w_1(I_1^n) &\geq M_1^n(I_1^s) - w_1(I_1^s),\\ M_1^s(I_1^s) - w_1(I_1^s) &\geq M_1^s(I_1^n) - w_1(I_1^n).\end{aligned}$$

Adding and rearranging yields

$$\int_{I_1^s}^{I_1^n} M_1^{n'}(I_1)\,dI_1 \quad \geq \quad \int_{I_1^s}^{I_1^n} M_1^{s'}(I_1)\,dI_1.$$

If now $I_1^n < I_1^s$ were to hold, then there must exists some value I_1 such that $M_1^{s'}(I_1) > M_1^{n'}(I_1)$. But above I have shown that $M_1^{n'}(I_1) \geq M_1^{s'}(I_1)$, for all I_1, a contradiction. Thus we can conclude that $I_1^n \geq I_1^s$. This completes the proof of Proposition 2.1 □

Proof of Proposition 2.2. For the most part, this proof proceeds analogously to that of Proposition 2.1, so here I only highlight one key difference. The marginal investment return for the buying division under negotiations equals $M_2^{n'}(I_2) = (1-\eta)\mathbb{E}_\theta[q^*(\theta, I_2)]$, while under standard-cost transfer pricing, according to Equation (2.10), we have $M_2^{s'}(I_2) = \mathbb{E}_\theta\left[\left(1 - \frac{\partial t^s(\theta, I_2)}{\partial I_2}\right) q^s(\theta, I_2)\right]$. The result then follows from the observation (see BRS for details) that, provided $\eta \leq \frac{1}{2}$, $M_2^{n'}(I_2) \geq M_2^{s'}(I_2)$ holds for all I_2 in one of two mutually exclusive cases: (i) If $r'(q,\theta)$ is concave in q so that $\partial t^s(\cdot)/\partial I_2 \geq \frac{1}{2}$, for all θ; or (ii) if $r'(q,\theta)$ is convex (so that $q^s(\cdot) \leq \frac{1}{2} q^*(\cdot)$) but log-concave in q (as then $\partial t^s(\cdot)/\partial I_2 \geq 0$). □

Proof of Proposition 2.3. The conditions stated in Footnote 8 of Section 2 are sufficient for: (i) existence of a unique pure-strategy Nash equilibrium $I^n = (I_1^n, I_2^n)$ under negotiated transfer pricing; and for (ii) $I^n \leq I^*$ (see BRS for a formal proof). Under the conditions stated in Proposition 2.3, the revealed preference arguments employed in the proofs of Propositions 2.1 and 2.2 apply again to show that, for given investment I_j, Division i invests weakly more under negotiation than under standard-cost transfer pricing. (The one-sided investment scenarios constitute special cases of the bilateral model where one division chooses zero investments.) A sufficient condition thus for Proposition 2.3 to hold is that $I_j^s \leq I_j^n,\ j = 1,2$.

Suppose that this is not the case, and, say, $I_1^s > I_1^n$ would hold. Then, we know that: (i) $(I_1^n)^{-1}(I_1^s) > I_2^n(I_1^s)$, by uniqueness of the Nash equilibrium under negotiations; and (ii) $I_1^n(I_2^s) \geq I_1^s$ and $I_2^n(I_1^s) \geq I_2^s$, by revealed preference. Using (i) and (ii), we obtain a contradiction:

$$I_1^s \equiv I_1^n((I_1^n)^{-1}(I_1^s)) > I_1^n(I_2^n(I_1^s)) \geq I_1^n(I_2^s) \geq I_1^s.$$

A similar chain of inequalities proves by contradiction that $I_2^s \leq I_2^n$. □

Proof of Proposition 3.1

Part (i). By the assumed binding capacity constraint, Division 1's contribution margin function is decreasing in p at $p(0,\theta)$. By continuity, this also holds for small values of Δ. For small Δ, Division 1 will thus set the external price $p(\Delta,\theta)$ so that

$$Q_i(\hat{p}(\Delta,\theta),\theta) + Q_e(p(\Delta,\theta),\theta) = Q_i(p(0,\theta),\theta) + Q_e(p(0,\theta),\theta) = K(\theta),$$

where $\hat{p}(\Delta,\theta) \equiv p(\Delta,\theta) - \Delta$. This implies that $\hat{p}'(0,\theta) < 0$, $p'(0,\theta) > 0$ and $\frac{dQ_i}{d\Delta} = -\frac{dQ_e}{d\Delta} > 0$ at $\Delta = 0$. Given binding upstream capacity, firmwide contribution margin for any θ and any small Δ equals:

$$\begin{aligned} M(\Delta,\theta) \equiv\ & R_i(Q_i(\hat{p}(\Delta,\theta),\theta),\theta) + R_e(K(\theta) \\ & -Q_i(\hat{\phi}(\Delta,\theta),\theta),\theta) - c(\theta) \cdot K(\theta), \end{aligned}$$

and

$$\left.\frac{d}{d\Delta}M(\Delta,\theta)\right|_{\Delta=0} = [R_i'(Q_i(p(0,\theta),\theta),\theta) - R_e'(Q_e(p(0,\theta),\theta),\theta)]\frac{dQ_i}{d\Delta},$$

which is indeed positive since $\frac{dQ_i}{d\Delta} > 0$ at $\Delta = 0$ and $R_i'(Q_i(p(0,\theta),\theta),\theta) = p(0,\theta) > R_e'(Q_e(p(0,\theta),\theta),\theta)$, as argued in the text.

Part (ii). The upstream division chooses p so as to maximize $M_1(\cdot)$ as in Program $\mathcal{P}_{4.2}$. Suppose first the upstream division sets the market price at $p^*(\theta) \equiv P_e(q_e^*(\theta),\theta)$, where $q_e^*(\theta)$ denotes the first-best external quantity in the presence of the capacity constraint. By definition, $R_e'(q_e^*(\theta),\theta) = p^*(\theta)\left(1 - \frac{1}{\varepsilon^*(\theta)}\right)$. Given the transfer pricing rule $TP(p) = p\left(1 - \frac{1}{f(p)}\right)$ and the external market price $p^*(\theta)$, the downstream division chooses $\hat{q}_i$ so that:

$$R_i'(\hat{q}_i,\theta) = t(p^*(\theta)) = \left[1 - \frac{1}{\varepsilon^*(\theta)}\right]p^*(\theta).$$

Thus, $R_i'(\hat{q}_i,\theta) = R_e'(q_e^*(\theta),\theta)$. Because the internal and external marginal revenues are equal only at the first-best quantities $(q_i^*(\theta),q_e^*(\theta))$, it follows that $\hat{q}_i = q_i^*(\theta)$.

It remains to be shown that the upstream division does not want to choose a price which exceeds $p^*(\theta) \equiv P_e(q_e^*(\theta),\theta)$. As assumed in the text, both the external and internal monopoly pricing problems are concave p. Thus, since $q_e^*(\theta) < q_e^m(\theta)$, any price $p > p^*(\theta)$ lowers the contribution margin of the upstream division externally. With regard to internal sales, we have:

$$t(p^*(\theta)) = R_i'(q_i^*(\theta),\theta) > R_i'(q_i^m(\theta),\theta) \equiv p_i^m(\theta), \tag{A.5}$$

because $R_i'(Q_i(p,\theta),\theta) \equiv p$. The inequality in Equation (A.5) relies on the stated condition that $q_i^*(\theta) < q_i^m(\theta)$. Therefore $t(p^*(\theta))$ exceeds the internal monopoly price. Because $t(\cdot)$ is increasing in p and by concavity of the internal monopoly pricing problem, it follows that any price above $p^*(\theta)$ would also lower the upstream division's contribution margin from internal sales. $\square$

Proof of Proposition 3.3.

Part (i). Follows from the main text.

Part (ii). For any given $\theta \in U$, the transfer pricing rule $t(\cdot)$ induces the upstream division to set an external market price $p(\theta)$ that maximizes (I am suppressing the dependence of $p(\cdot)$ on $t(\cdot)$ to minimize clutter):

$$M_1(p,\theta \mid t(\cdot)) = [t(p) - c(\theta)]Q_i(t(p),\theta) + [p - c(\theta)]Q_e(p,\theta) \tag{A.6}$$

The corresponding first-order condition is:

$$\begin{aligned} t'(p(\theta))\{Q_i(t(p(\theta)),\theta) + [t(p(\theta)) - c(\theta)]Q_i'(t(p(\theta)),\theta)\} \\ + [p(\theta) - c(\theta)]Q_e'(p(\theta),\theta) + Q_e(p(\theta),\theta) = 0. \end{aligned} \tag{A.7}$$

Efficiency requires that $p(\theta) = p_e^m(\theta)$ and $t(p(\theta)) = c(\theta)$. Therefore the first-order condition in Equation (A.7) reduces to:

$$t'(p_e^m(\theta))Q_i(t(p_e^m(\theta)),\theta) = 0,$$

which implies $t'(p_e^m(\theta)) = 0$. But that contradicts the maintained assumption that $t(p_e^m(\theta)) = c(\theta)$ and $\nabla c(\theta) \neq 0$. $\square$

Proof of Proposition 3.4. The selling division's pricing problem at Date 3 for given Δ and θ reads as follows:

$$\max_p \; M_1(p,\theta \mid \Delta) \equiv [p - c(\theta)]Q_e(p,\theta) + [p - \Delta - c(\theta)]Q_i(p - \Delta,\theta).$$

Given the linear scenario in Equation (3.5), the solution to this problem is given by

$$p(\Delta, \theta) = p(0, \theta) + \nu(\theta)\Delta,$$

where $p(0, \theta) = \frac{1}{2}\left[\frac{\alpha_e(\theta)+\alpha_i(\theta)}{\beta_e(\theta)+\beta_i(\theta)} + c(\theta)\right]$ is the seller's optimal price for $\Delta = 0$, and $p'(\Delta, \theta) = \nu(\theta) \equiv \frac{\beta_i(\theta)}{\beta_e(\theta)+\beta_i(\theta)}$ is the rate at which this price increases as Δ increases. Using Equation (3.5) once more, it is easy to show that $dQ_e/d\Delta = -\beta_e(\theta)\nu(\theta) \equiv -dQ_i/d\Delta = [\nu(\theta) - 1]\beta_i(\theta) < 0$. Hence, the internal discount again just redirects some units of upstream output freon external sales to internal transfers. The remaining arguments are identical to those in Proposition 3.1 and thus omitted. □

Proof of Proposition 4.1.

Claim A.1. Given the conditions stated in Proposition 4.1, each division underinvests for $\Delta = 0$, i.e., $I_j(\Delta = 0) < I_j^*,\ j = i, e$.

Proof: Omitted. It follows from similar revealed preference arguments as the earlier results, together with the fact that investments are strategic complements.

Claim A.2. The Nash equilibrium $I(\Delta)$ is monotonically increasing in Δ.

Proof: Using the arguments in Milgrom and Roberts (1990), the Nash equilibrium in investments will be monotonically increasing in Δ, if the divisional objective functions in Equation (4.2) have increasing differences in Δ, I_1 and I_2. Now:

$$\begin{aligned}
\pi_1(\Delta, I) = \mathbb{E}_\theta\{&[p(\Delta, \theta, I) - c(\theta) + I_1]Q_e(p(\Delta, \theta, I), \theta) \\
&+[p(\Delta, \theta, I) - \Delta - c(\theta) + I_1] \\
&\times Q_i(p(\Delta, \theta, I) - \Delta, \theta, I_2)\} - w_1(I_1), \\
\pi_2(\Delta, I) = \mathbb{E}_\theta\{&R_i(Q_i(p(\Delta, \theta, I) - \Delta, \theta, I_2), \theta, I_2) \\
&-[p(\Delta, \theta, I) - \Delta]Q_i(p(\Delta, \theta, I) - \Delta, \theta, I_2)\} - w_2(I_2),
\end{aligned}$$

with $p(\Delta,\theta,I)$ as given by Equation (4.4). First consider the selling division's objective function: by the envelope theorem,

$$\begin{aligned}\frac{\partial \pi_1}{\partial \Delta} &= \mathbb{E}_\theta\{-Q_i(p(\Delta,\theta,I)-\Delta,\theta,I_2)\\ &\qquad -[p(\Delta,\theta,I)-\Delta-c(\theta)+I_1]Q_i'(\cdot)\}.\end{aligned}$$

Differentiating once more:

$$\begin{aligned}\frac{\partial^2 \pi_1}{\partial \Delta \partial I_1} &= \mathbb{E}_\theta\left[-\left(1+\frac{\partial p}{\partial I_1}\right)Q_i'(\cdot)\right] = \frac{1}{2}\mathbb{E}_\theta\left[\beta_i(\theta)\right] > 0,\\ \frac{\partial^2 \pi_1}{\partial \Delta \partial I_2} &= \mathbb{E}_\theta\left[-\frac{\partial p}{\partial I_2}Q_i'(\cdot)\right] = \mathbb{E}_\theta\left[\frac{\nu(\theta)}{2}\beta_i(\theta)\right] > 0\end{aligned}$$

using Equation (4.4). Also, by the envelope theorem,

$$\frac{\partial \pi_1}{\partial I_1} = \mathbb{E}_\theta\{Q_i(p(\Delta,\theta,I)-\Delta,\theta,I_2)+Q_e(p(\Delta,\theta,I),\theta)\} - w_1'(I_1), \tag{A.8}$$

$$\begin{aligned}\frac{\partial^2 \pi_1}{\partial I_1 \partial I_2} &= \mathbb{E}_\theta\left[\frac{\partial p}{\partial I_2}[Q_i'(\cdot)+Q_e'(\cdot)] + \frac{\partial}{\partial I_2}Q_i(p(\Delta,\theta,I)-\Delta,\theta,I_2)\right]\\ &= \mathbb{E}_\theta\left[-\beta_i(\theta)\frac{\nu(\theta)}{2}+\beta_i(\theta)-\beta_e(\theta)\frac{\nu(\theta)}{2}\right]\\ &= \frac{1}{2}\mathbb{E}_\theta[\beta_i(\theta)] > 0.\end{aligned} \tag{A.9}$$

As for the buying division's objective function:

$$\begin{aligned}\frac{\partial \pi_2}{\partial \Delta} &= \mathbb{E}_\theta\left[\left(1-\frac{\partial p}{\partial \Delta}\right)Q_i(p(\Delta,\theta,I)-\Delta,\theta,I_2)\right],\\ \frac{\partial^2 \pi_2}{\partial \Delta \partial I_2} &= \mathbb{E}_\theta\left[(1-\nu(\theta))\left(\frac{\partial p}{\partial I_2}Q_i'(\cdot)+\frac{\partial}{\partial I_2}Q_i(p(\Delta,\theta,I)-\Delta,\theta,I_2)\right)\right]\\ &= \mathbb{E}_\theta\left[(1-\nu(\theta))\left(1-\frac{\nu(\theta)}{2}\right)\beta_i(\theta)\right] > 0,\\ \frac{\partial^2 \pi_2}{\partial \Delta \partial I_1} &= \mathbb{E}_\theta\left[(1-\nu(\theta))\frac{\partial p}{\partial I_1}Q_i'(\cdot)\right]\\ &= \frac{1}{2}\mathbb{E}_\theta[(1-\nu(\theta))\beta_i(\theta)] > 0,\end{aligned}$$

and

$$\frac{\partial \pi_2}{\partial I_2} = \mathbb{E}_\theta \left[\left(1 - \frac{\partial p}{\partial I_2}\right) Q_i(p(\Delta, \theta, I) - \Delta, \theta, I_2) \right] - w_2'(I_2), \quad \text{(A.10)}$$

$$\frac{\partial^2 \pi_2}{\partial I_1 \partial I_2} = \mathbb{E}_\theta \left[\left(1 - \frac{\nu(\theta)}{2}\right) \frac{\partial p}{\partial I_1} Q_i'(\cdot) \right] \quad \text{(A.11)}$$

$$= \mathbb{E}_\theta \left[\left(1 - \frac{\nu(\theta)}{2}\right) \frac{\beta_i(\theta)}{2} \right] > 0. \quad \text{(A.12)}$$

Thus, both objective functions have increasing differences in Δ, I_1, I_2, which completes the proof of Claim A.2.

Equipped with Claims A.1 and A.2, I can now prove the main result of Proposition 4.1. Firm-wide expected profit as a function of Δ equals

$$\begin{aligned}\pi(\Delta) &= \pi_1(\Delta, I(\Delta)) + \pi_2(\Delta, I(\Delta)) \\ &= \mathbb{E}_\theta \{ R_i(Q_i(p(\Delta, \theta, I(\Delta)) - \Delta, \theta, I_2(\Delta)), \theta, I_2(\Delta)) \\ &\quad + R_e(Q_e(p(\Delta, \theta, I(\Delta)), \theta), \theta) - [c(\theta) - I_1(\Delta)] \cdot [Q_i(\cdot) + Q_e(\cdot)] \} \\ &\quad - \sum_{j=1}^{2} w_j(I_j(\Delta)).\end{aligned}$$

For given (Δ, θ, I, p), the buying division chooses its demand so that

$$R_i'(Q_i(p - \Delta, \theta, I_2), \theta, I_2) = p - \Delta. \quad \text{(A.13)}$$

At Date 3, the seller's pricing problem for given Δ, θ, I, is described by Equation (4.1). At Date 1, for given Δ, the Nash equilibrium of divisional investments are described by Equation (4.2). Differentiating $\pi(\Delta)$ and using Equation (4.1), Equation (A.13) yields:

$$\begin{aligned}\pi'(\Delta) = \mathbb{E}_\theta \Bigg[&\left(\frac{\partial R_i}{\partial I_2} - \frac{\partial p}{\partial I_2} - w_2'(I_2(\Delta)) \right) \frac{dI_2}{d\Delta} \\ &+ \left[Q_i(\cdot) + Q_e(\cdot) - w_1'(I_1(\Delta)) \right] \frac{dI_1}{d\Delta} - \left(\frac{\partial p}{\partial \Delta} + \frac{\partial p}{\partial I_1} \frac{dI_1}{d\Delta} \right) Q_i(\cdot) \\ &+ [p(\cdot) - \Delta - c(\theta) + I_1(\Delta)] \left(-Q_i'(\cdot) + \frac{\partial Q_i}{\partial I_2} \frac{dI_2}{d\Delta} \right) \Bigg]\end{aligned}$$

The two terms in the first row are zero each, by the first-order conditions associated with the investment choices in Equations (A.8) and (A.10). Evaluating $\pi'(\Delta)$ at $\Delta = 0$ gives:

$$\pi'(0) = \underbrace{\begin{array}{l}[R_i'(Q_i(p(0,\theta,I(0)),\theta,I_2(0)),\theta,I_2(0)) \\ \qquad -R_e'(Q_e(p(0,\theta,I(0)),\theta),\theta)]\,\frac{\partial p}{\partial \Delta}(-Q_e'(\cdot))\end{array}}_{(a)}$$
$$+\underbrace{\left(-\frac{\partial p}{\partial I_1}\right)\frac{dI_1}{d\Delta}Q_i(\cdot)}_{(b)}+\underbrace{[p(0,\theta,I(0)) - c(\theta) + I_1]\frac{\partial Q_i}{\partial I_2}\frac{dI_2}{d\Delta}}_{(c)}$$

Using Claims A.1 and A.2, it is easy to show that all three terms are positive. Term (a) is positive by virtue of the double-marginalization problem (together with $\partial p/\partial\Delta > 0$ and $Q_e' < 0$); (b) captures the first-order gain to the buyer arising from the reduction in the price $p(\cdot)$ due to the seller's investing more; and (c) captures the first-order gain to the seller due to the buyer requesting more units as a result of I_2 being increasing in Δ. Note that (a) is the sole effect driving Proposition 3.4, part (i), above; (b) and (c) express the additional rationale for imposing internal discounts if divisions can invest specifically in the transaction. □

Proof of Lemma 5.1. I follow Chatterjee and Samuelson (1983) in deriving the linear-strategy Bayesian–Nash equilibrium under negotiated transfer pricing. For given I, let $\tau_1 \equiv \theta_1 - I_1$, and $\tau_2 \equiv \theta_2 + I_2$ denote the relevant divisional costs and revenues. The corresponding support boundaries are $\underline{\tau}_1 \equiv \underline{\theta}_1 - I_1$, $\bar{\tau}_1 \equiv \bar{\theta}_1 - I_1$, $\underline{\tau}_2 \equiv \underline{\theta}_2 - I_2$, and $\bar{\tau}_2 \equiv \bar{\theta}_2 - I_2$. This induces (uniform) distributions $\hat{F}_i(\tau_i)$, defined over $[\underline{\tau}_i, \bar{\tau}_i]$. Now restate the maximization problems as given in Equations (5.7) and (5.8):

$$\hat{s}(\tau_1, I) = \arg\max_s \int_s^{\hat{b}(\bar{\tau}_2, I)} \left[\frac{1}{2}(b+s) - \tau_1\right] dG_2(b, I),$$
$$\hat{b}(\tau_2, I) = \arg\max_b \int_{\hat{s}(\underline{\tau}_1, I)}^{b} \left[\tau_2 - \frac{1}{2}(b+s)\right] dG_1(s, I).$$

The bid distribution functions G_i are induced by (i) the underlying type distributions $\hat{F}_i(\tau_i)$, and by (ii) the bidding strategies $\hat{s}$ and $\hat{b}$, where $G_1(\xi, I) = \hat{F}_1(\hat{s}^{-1}(\xi, I))$ and $G_2(\xi, I) = \hat{F}_2(\hat{b}^{-1}(\xi, I))$. The first-order condition for the seller is

$$\frac{1}{2}[1 - G_2(\hat{s}(\cdot), I)] - [\hat{s}(\cdot) - \tau_1] g_2(\hat{s}(\cdot), I) = 0,$$

with g_i denoting the density function to G_i. Defining $x \equiv \hat{b}^{-1}(\hat{s}, I)$, we have $g_2(\hat{s}, I) = f_2(x)/\hat{b}'(x, I)$, $\tau_1 = \hat{s}^{-1}(\hat{b}(x, I), I)$, and $G_2(\hat{s}, I) = F_2(x)$, and the first-order condition can be restated as

$$\hat{s}^{-1}(\hat{b}(x, I), I) = \hat{b}(x, I) - \frac{1}{2}\hat{b}'(x, I)\frac{1 - \hat{F}_2(x)}{\hat{f}_2(x)}.$$

Proceeding in a similar fashion for the buyer yields

$$\hat{b}^{-1}(\hat{s}(y, I), I) = \hat{s}(y, I) + \frac{1}{2}\hat{s}'(y, I)\frac{\hat{F}_1(y)}{\hat{f}_1(y)},$$

where $y \equiv \hat{s}^{-1}(\hat{b}, I)$. A Bayesian–Nash equilibrium now is a solution to these two linked differential equations. Restricting attention to linear bidding strategies $\hat{s}(\tau_1, I) = \alpha_1(I) + \beta_1(I)\tau_1$ and $\hat{b}(\tau_2, I) = \alpha_2(I) + \beta_2(I)\tau_2$ gives

$$\begin{aligned} \hat{s}^{-1}(\hat{b}(\tau_2, I), I) &= \hat{b}(\tau_2, I) - \frac{1}{2}\beta_2(I)[\bar{\tau}_2 - \tau_2], \\ \hat{b}^{-1}(\hat{s}(\tau_1, I), I) &= \hat{s}(\tau_1, I) + \frac{1}{2}\beta_1(I)[\tau_1 - \underline{\tau}_1]. \end{aligned}$$

Differentiating the latter determines the slopes of the strategies, which turn out to be independent of I: $\beta_1 = \beta_2 = \frac{2}{3}$. It follows that the intercept terms are $\alpha_1 = \frac{1}{4}\bar{\tau}_2 + \frac{1}{12}\underline{\tau}_1$ and $\alpha_2 = \frac{1}{12}\bar{\tau}_2 + \frac{1}{4}\underline{\tau}_1$. Re-scaling both divisions' valuations in terms of $\theta_i = \tau_i \pm I_i$ yields the interior solutions, i.e., the middle branches of Equations (5.10) and (5.11). For an extensive discussion of the boundary conditions, the reader is referred to Chatterjee and Samuelson (1983). □

Proof of Lemma 5.2. Recall that trade under negotiated transfer pricing occurs if and only if $s(\theta_1, I) \leq b(\theta_2, I)$. Given Assumption 1, it is easy to show that the conditional trading probability for any given (unilateral) type realization θ_1 is strictly less than one, and zero for the nonempty set of high-θ_1 and low-θ_2 types.

Denote by $\theta_2^n(\theta_1 \mid I)$ the (unique) revenue type-cutoff that is implicitly defined by $s(\theta_1, I) \equiv b(\theta_2^n(\theta_1 \mid I), I)$, for given I, and define $\theta_1^n(I)$ such that $\theta_2^n(\theta_1^n(I) \mid I) \equiv \overline{\theta}_2$. That is, for any cost realization $\theta_1 > \theta_1^n(I)$, the conditional probability of trade equals zero, as the set $\{\theta_2 \in \Theta_2 \mid s(\theta_1^n(I), I) \leq b(\theta_2, I)\}$ is empty. The expected contribution margins for given I under the respective two schemes are:

$$\begin{aligned} M^n(I) &= \int_{\underline{\theta}_1}^{\theta_1^n(I)} \int_{\theta_2^n(\theta_1|I)}^{\overline{\theta}_2} [\theta_2 - \theta_1 + I_1 + I_2] dF_2(\theta_2) dF_1(\theta_1) \\ &= \frac{9}{64\delta_1\delta_2}(\overline{\theta}_2 - \underline{\theta}_1 + I_1 + I_2)^3 \end{aligned} \tag{A.14}$$

$$\begin{aligned} M^s(I) &= \int_{\Theta_1} \int_{t^s(\theta_1, I) - I_2}^{\bar{\theta}_2} [\theta_2 - \theta_1 + I_1 + I_2] dF_2(\theta_2) dF_1(\theta_1), \\ &= \frac{1}{8\delta_1\delta_2}\big[(\overline{\theta}_2 - \underline{\theta}_1 + I_1 + I_2)^3 \\ &\quad - (\overline{\theta}_2 - \overline{\theta}_1 + I_1 + I_2)^3\big]. \end{aligned} \tag{A.15}$$

By Assumption 1, $\bar{\theta}_2 - \bar{\theta}_1 \geq 0$. Hence, $M^n(I) > M^s(I)$, for any I. $\square$

Proof of Proposition 5.3. It remains to be shown that $I_1^s > I_1^n$. By revealed preference, a sufficient condition for this to hold is that the seller's marginal investment returns are uniformly greater under cost-based pricing. Under negotiations the seller chooses I_1^n so as to maximize:

$$\begin{aligned} &\int_{\underline{\theta}_1}^{\theta_1^n(I_1)} \int_{\theta_2^n(\theta_1|I_1)}^{\bar{\theta}_2} \left[\frac{b(\theta_2, I_1) + s(\theta_1, I_1)}{2} - \theta_1 + I_1 \right] \\ &\qquad \times dF_2(\theta_2) dF_1(\theta_1) \; - w_1(I_1), \end{aligned}$$

where $\theta_2^n(\theta_1 \mid I_1) = \theta_1 + \frac{1}{4}(\bar{\theta}_2 - \underline{\theta}_1 - 3I_1)$, while $\theta_1^n(I_1)$ is as defined in the proof of Lemma 5.2. Clearly, the transfer price, computed as the average of the two bids, is a function of I_1. The seller's marginal investment return equals

$$M_1^{n'}(I_1) = \left[1 + \frac{1}{2}\left(\frac{\partial b(\cdot)}{\partial I_1} + \frac{\partial s(\cdot)}{\partial I_1}\right)\right](\text{Prob}\{q^n(\theta, I_1) = 1\})$$
$$- \int_{\underline{\theta}_1}^{\theta_1^n(I_1)} \left(\frac{\partial \theta_2(\theta_1 \mid I_1)}{\partial I_1}\right)\left(\frac{s(\theta_1, I_1) - (\theta_1 - I_1)}{\delta_1 \delta_2}\right) d\theta_1. \tag{A.16}$$

The necessary first-order condition for an optimum, I_1^n, requires that $w_1'(I_1^n) = M_1^{n'}(I_1^n)$, which, after some simplifications, becomes

$$M_1^{n'}(I_1) = \frac{3}{4}\text{Prob}\{q^n(\theta, I_1^n) = 1\}. \tag{A.17}$$

Under standard-cost transfer pricing, on the other hand, the seller's investment is described by Equation (2.8). Using Equations (2.8) and (A.17) the seller's marginal investment returns under the respective regimes can be rewritten as:

$$M_1^{n'}(I_1) = \frac{3}{4}\int_{\underline{\theta}_1}^{\theta_1^n(I_1)} \int_{\theta_2^n(\theta_1|I_1)}^{\bar{\theta}_2} dF_2(\theta_2) dF_1(\theta_1)$$
$$= \frac{27(\bar{\theta}_2 - \underline{\theta}_1 + I_1)^2}{128\ \delta\theta_1 \delta\theta_2}, \tag{A.18}$$

$$M_1^{s'}(I_1) = \int_{\Theta_1} \int_{t^s(\theta_1, I_1)}^{\bar{\theta}_2} dF_2(\theta_2) dF_1(\theta_1)$$
$$= \frac{(\bar{\theta}_2 - \underline{\theta}_1 + I_1)^2 - (\bar{\theta}_2 - \bar{\theta}_1 + I_1)^2}{4\ \delta\theta_1 \delta\theta_2}.- \tag{A.19}$$

Suppose now that, contrary to the claim, there exists some value I_1 such that $M_1^{n'}(I_1) \geq M_1^{s'}(I_1)$. Then, by Equations (A.18) and (A.19), $(\bar{\theta}_2 - \bar{\theta}_1 + I_1)^2 \geq (1 - \frac{27}{32})(\bar{\theta}_2 - \underline{\theta}_1 + I_1)^2$. This, however, contradicts Assumption 1, specifically the part that prescribes $\bar{\theta}_1 \geq \frac{1}{4}(3\bar{\theta}_2 + \underline{\theta}_1 + 3I_1)$. Thus, $I_1^s > I_1^n$. □

Proof of Proposition 5.4. The result follows immediately from the respective first-order conditions: $M_2^{n'}(I_2) = \frac{3}{4}\text{Prob}\{q^n(\theta, I_2) = 1\}$ under negotiations and $M_2^{s'}(I_2) = \frac{1}{2}\text{Prob}\{q^s(\theta, I_2) = 1\}$ under standard-cost pricing. Since $\text{Prob}\{q^n(\theta, I_2) = 1\} > \text{Prob}\{q^s(\theta, I_2) = 1\}$ by a comparison of Equations (A.18) and (A.19), it follows that $I_2^n > I_2^s$. Dominance of negotiations then follows from Assumption 2. $\square$

Proof of Proposition 5.5. Firmwide expected profit as a function of Γ is:

$$\pi(\Gamma) = M(\Gamma, I_1(\Gamma)) - w_1(I_1(\Gamma))$$

subject to

$$I_1(\Gamma) \in \arg\max_{I_1} \; M_1(\Gamma, I_1) - w_1(I_1),$$

where

$$\begin{aligned} M(\Gamma, I_1) &= \mathbb{E}_\theta\{[\theta_2 - \theta_1 + I_1(\Gamma)]q(\theta, \Gamma, I_1(\Gamma))\} \\ M_1(\Gamma, I_1) &= \mathbb{E}_\theta\{[t(\cdot) - \theta_1 + I_1(\Gamma)]q(\theta, \Gamma, I_1(\Gamma))\} \end{aligned}$$

are the contributions for the firm as a whole and for Division 1, respectively. Taking the derivative of $\pi(\Gamma)$ yields:

$$\pi'(\Gamma) = \frac{\partial}{\partial \Gamma} M(\Gamma, I_1(\Gamma)) + \left[\frac{\partial}{\partial I_1} M(\Gamma, I_1(\Gamma)) - w_1'(I_1(\Gamma))\right] \frac{dI_1}{d\Gamma}. \quad \text{(A.20)}$$

It is a matter of straightforward algebra to show that (i) $(\partial/\partial\Gamma)M(\Gamma, I_1(\Gamma))|_{\Gamma=0} > 0$ while (ii) $I_1(\Gamma)$ is increasing in Γ (see BB for details). It remains to be shown that the term in square brackets in Equation (A.20) is positive, i.e., that the induced increase in upstream investment indeed makes the firm better off. This, however, follows directly form the underinvestment result at $\Gamma = 0$, as the first-order condition describing Manager 1's investment choice is

$$\frac{\partial}{\partial I_1} M_1(\Gamma, I_1(\Gamma)) - w_1'(I_1(\Gamma)) = 0,$$

Lastly, it is straightforward to show that the marginal firmwide investment return exceeds the marginal investment return for Division 1, i.e.,

$$\left.\frac{\partial}{\partial I_1} M_1(\Gamma, I_1(\Gamma))\right|_{\Gamma=0} < \left.\frac{\partial}{\partial I_1} M(\Gamma, I_1(\Gamma))\right|_{\Gamma=0}.$$

This completes the proof of Proposition 5.5. $\square$

References

Alles, M. and S. Datar (1998), 'Strategic transfer pricing'. *Management Science* **44**, 451–461.

Anctil, R. and S. Dutta (1999), 'Negotiated transfer pricing and divisional vs. firm-wide performance evaluation'. *Accounting Review* **74**, 87–104.

Argyres, N. (1995), 'Technology strategy, governance structure and interdivisional coordination'. *Journal of Economic Behavior and Organization* **28**, 337–358.

Arya, A., B. Mittendorf, and D.-H. Yoon (2008), 'Friction in related party trade when a rival is also a customer'. *Management Science* **54**(11), 1850–1860.

Ausubel, L., P. Cramton, and R. Deneckere (2002), 'Bargaining with incomplete information'. In: R. Aumann and S. Hart (eds.): *Handbook of Game Theory*, Vol. 3.

Baiman, S. and T. Baldenius (2009), 'Non-financial performance measures as coordination devices'. *Accounting Review* **84**(2), 299–330.

Baiman, S. and M. Rajan (2002), 'The role of information and opportunism in the choice of buyer-supplier relationships'. *Journal of Accounting Research* **40**(2), 247–278.

Baldenius, T. (2000), 'Comparative statics of monopoly pricing'. *Economic Theory* **16**, 465–469.

Baldenius, T. (2006), 'Ownership, incentives and the holdup problem'. *The RAND Journal of Economics* **37**(2), 276–299.

Baldenius, T., N. Melumad, and S. Reichelstein (2004), 'Integrating managerial and tax objectives in transfer pricing'. *Accounting Review* **79**, 591–615.

Baldenius, T. and S. Reichelstein (2006), 'External and internal pricing in multidivisional firms'. *Journal of Accounting Research* **44**(1), 1–28.

Baldenius, T., S. Reichelstein, and S. Sahay (1999), 'Negotiated versus cost-based transfer pricing'. *Review of Accounting Studies* **4**, 67–91.

Bockem, S. and U. Schiller (2004), 'Transfer pricing and holdups in supply chains'. *German Economic Review* **5**, 211–230.

Bouwens, J. and L. van Lent (2007), 'Assessing the performance of business unit managers'. *Journal of Accounting Research* **45**(4), 667–697.

Bushman, R., R. Indjejikian, and A. Smith (1995), 'Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies'. *Journal of Accounting Research* **33**, 101–128.

Chatterjee, K. and W. Samuelson (1983), 'Bargaining under incomplete information'. *Operations Research* **31**, 835–851.

Che, Y.-K. and J. Sakovics (2004), 'A dynamic theory of holdup'. *Econometrica* **72**, 1063–1103.

Dutta, S. and S. Reichelstein (2009), 'Capacity management, decentralization, and internal pricing'. Working paper, U.C. Berkeley.

Eccles, R. (1985), *The Transfer Pricing Problem: A Theory for Practice*. Lexington: Lexington Books.

Eccles, R. and H. White (1988), 'Price and authority in inter-profit center transactions'. *American Journal of Sociology* **94**(Supplement), 17–51.

Edlin, A. and S. Reichelstein (1995), 'Specific investment under negotiated transfer pricing: An efficiency result'. *Accounting Review* **70**, 275–291.

Ernst and Young (2008), *Precision Under Pressure Global: Transfer Pricing Survey*. 2007-2008, `www.ey.com`.

Gox, R. (2000), 'Strategic transfer pricing, absorption costing, and observability'. *Management Accounting Research* **11**, 327–348.

Gox, R. and U. Schiller (2007), 'An economic perspective on transfer pricing'. In: C. S. Chapman, A. G. Hopwood, and M. D. Shields (eds.): *Handbook of Management Accounting Research*.

Halperin, R. and B. Srinidhi (1991), 'U.S. income tax transfer-pricing rules and resource allocation: The case of decentralized multinational firms'. *The Accounting Review* **66**, 141–157.

Harris, M., C. Kriebel, and A. Raviv (1982), 'Asymmetric information, incentives, and intrafirm resource allocation'. *Management Science* **28**, 604–620.

Hart, O. and J. Moore (1999), 'Foundations of incomplete contracts'. *Review of Economic Studies* **66**, 115–138.

Hart, O. and J. Moore (2008), 'Contracts as reference points'. *Quarterly Journal of Economics* **123**(1), 1–48.

Hirshleifer, J. (1956), 'On the economics of transfer pricing'. *Journal of Business* **29**, 172–189.

Holmstrom, B. (1999), 'The firm as a subeconomy'. *Journal of Law, Economics and Organization* **15**, 74–102.

Holmstrom, B. and J. Tirole (1991), 'Transfer pricing and organizational form'. *Journal of Law, Economics, and Organization* **7**, 201–228.

Horngren, C., S. Datar, and G. Foster (2007), *Cost Accounting: A Managerial Emphasis*. Prentice-Hall, 12th edition.

Hyde, C. and C. Choe (2005), 'Keeping two sets of books: The relationship between tax and incentive transfer prices'. *Journal of Economics and Management Strategy* **14**(1), 165–186.

Johnson, N. (2006), 'Divisional performance measurement and transfer pricing for intangible assets'. *Review of Accounting Studies* **11**, 339–365.

Joskow, P. (2005), 'Vertical integration'. In: C. Menard and M. Shirley (eds.): *Handbook of New Institutional Economics*. Springer.

Kanodia, C. (1991), 'Transfer payment mechanisms and intrafirm coordination'. Working paper, University of Minnesota.

Leininger, W., P. Linhart, and R. Radner (1989), 'Equilibria of the sealed-bid mechanism for bargaining with incomplete information'. *Journal of Economic Theory* **48**, 63–106.

Maskin, E. and J. Riley (1984), 'Monopoly with incomplete information'. *Rand Journal of Economics* **15**(2), 171–196.

Maskin, E. and J. Tirole (1999), 'Unforeseen contingencies and incomplete contracts'. *Review of Economic Studies* **66**, 83–114.

Merchant, K. (1989), *Rewarding Results: Motivating Profit Center Managers*. Boston: Harvard Business School Press.

Milgrom, P. and J. Roberts (1990), 'Rationalizability, learning and equilibrium in games with strategic complementarities'. *Econometrica* **58**, 1255–1278.

Myerson, R. and M. Satterthwaite (1983), 'Efficient mechanisms for bilateral trading'. *Journal of Economic Theory* **29**, 265–281.

Narayanan, V. G. and M. Smith (2000), 'Impact of competition and taxes on responsibility center organization and transfer prices'. *Contemporary Accounting Research* **17**(3), 497–529.

Pfeiffer, T., S. Lengsfeld, U. Schiller, and J. Wagner (2009), 'Cost based transfer pricing'. Working paper, University of Vienna.

Radner, R. and A. Schotter (1989), 'The sealed-bid mechanism: An experimental study'. *Journal of Economic Theory* **48**, 179–220.

Rogerson, W. (1992), 'Contractual solutions to the hold-up problem'. *Review of Economic Studies* **59**, 777–793.

Rogerson, W. (2008), 'Intertemporal cost allocation and investment decisions'. *Journal of Political Economy* **116**(5), 931–950.

Sahay, S. (2003), 'Transfer pricing based on actual costs'. *Journal of Management Accounting Research* **15**, 177–192.

Shor, M. and H. Chen (2009), 'Decentralization, transfer pricing, and tacit collusion'. *Contemporary Accounting Review* **26**, 581–604.

Tirole, J. (1988), *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.

Vaysman, I. (1996), 'A model of cost-based transfer pricing'. *Review of Accounting Studies* **1**, 73–108.

Vaysman, I. (1998), 'A model of negotiated transfer pricing'. *Journal of Accounting and Economics* **25**, 349–384.

Wagenhofer, A. (1994), 'Transfer pricing under asymmetric information: An evaluation of alternative methods'. *European Accounting Review* **1**, 71–104.

Wielenberg, S. (2000), 'Negotiated transfer pricing, specific investment, and optimal capacity choice'. *Review of Accounting Studies* **5**(3), 197–216.

Williamson, O. (1985), *The Economic Institutions of Capitalism*. New York: Free Press.

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