External and Internal Pricing in Multidivisional Firms

TIM BALDENIUS* AND STEFAN REICHELSTEIN†

Received 30 July 2004; accepted 2 August 2005

ABSTRACT

Multidivisional firms frequently rely on external market prices in order to value internal transactions across profit centers. This paper examines market-based transfer pricing when an upstream division has monopoly power in selling a proprietary component both to a downstream division within the same firm and to external customers. When internal transfers are valued at the prevailing market price, the resulting transactions are distorted by double marginalization. The imposition of intracompany discounts will always improve overall firm profits provided the supplying division is capacity constrained. Under certain conditions it is then possible to design discount rules so that the resulting prices and sales quantities are efficient from the corporate perspective. In contrast, the impact of intracompany discounts remains ambiguous when the capacity of the selling division is essentially unlimited. It is then generally impossible to achieve fully efficient outcomes by means of market-based transfer pricing unless the external market for the component is sufficiently large relative to the internal market.

1. Introduction

The determination of internal transfer prices remains a contentious issue in many divisionalized firms. To value intermediate products or services

* Columbia University; † Stanford University. This paper supersedes an earlier working paper titled “Market-Based Transfer Pricing and Intracompany Discounts.” We are grateful to Nicole Bastian, Severin Borenstein, Thomas Bruns, Aaron Edlin, Joseph Harrington, Bill Rogerson and Alfred Wagenhofer for many helpful comments and suggestions. We also thank seminar participants at the Universities of Bonn, British Columbia, Chicago, Cologne, Columbia, CUNY–Baruch College, INSEAD, Madrid, Minnesota, Simon Fraser, Stanford and Zurich.

Copyright ©, University of Chicago on behalf of the Institute of Professional Accounting, 2006
that are transacted across divisions (profit centers), firms rely on a range of transfer pricing methods, most of which can be classified as either market based, cost based or negotiated. Beginning with the classical papers of Cook [1955] and Hirshleifer [1956], the management literature has advocated that interdivisional sales be valued at the prevailing market price provided the external market for the intermediate product is competitive. There is virtual agreement in the literature that under those conditions market-based transfer pricing leads to an efficient decentralization of decision making. The competitive market scenario, however, is arguably only a conceptual benchmark. In many industries the very rationale for vertical integration appears to be that intermediate products and services are specialized or even proprietary.1

Even though the scenario of perfectly competitive external markets does not seem descriptive of most intermediate products, surveys have documented consistently that market-based transfer pricing is widely used in practice. Between 30% and 45% of the responding firms indicate that they adopt a market-based approach to internal pricing.2 These observations make it essential to analyze the extent to which the efficiency properties attributed to market-based transfer pricing under conditions of a competitive market also extend to imperfectly competitive markets. A better theoretical understanding of this issue will be helpful in predicting which of the commonly observed transfer pricing methods a particular firm is likely to adopt.3

Under market-based transfer pricing the external market price is frequently viewed as a starting point from which the internal transfer price is calculated by means of adjustments, or so-called intracompany discounts. The stated objective of these discounts is to account for (1) possible cost differences between internal and external transactions, (2) differences in product characteristics or quality, and (3) other market “imperfections.”4 Unfortunately, there does not seem to be much systematic evidence

1 See, for instance, Joskow [2004], Kaplan and Atkinson [1998], Brickley, Smith, and Zimmerman [1995], Milgrom and Roberts [1992], and Eccles and White [1988].

2 See Horngren, Datar, and Foster [2005], Tang [2002], Ernst and Young [1999], Borkowski [1999], and Price Waterhouse [1984].

3 One commonly cited advantage of market-based transfer pricing is its administrative simplicity. In contrast to other pricing mechanisms, which rely on extensive reporting and verification requirements, a market-based approach principally only requires an assessment of the price at which the intermediate product is sold to third parties.

4 Kaplan and Atkinson [1998, p. 454] explain adjustments to the market price as an attempt to “reflect cost savings and to encourage internal transfers.” The Price Waterhouse [1984] study refers to the “adjusted” market price as “simply the prevailing market price adjusted for market imperfections that are avoided by selling internally.” According to Merchant [2000, p. 618] “Many firms use market-based transfer pricing where competition is not perfect by allowing for deviations from the observed market prices.” In connection with imperfectly competitive external markets, Maher, Lanen, and Rajan [2005, p. 439] state “a common variation on this approach (transfers at market price) is to establish a policy that provides the buying division a discount for items produced internally.”
We depart from the standard competitive market scenario by assuming that the upstream division has effective monopoly power for the intermediate product, possibly because the product is proprietary due to intellectual property rights. As a consequence, the buying division cannot obtain the intermediate product in question from external sources, yet the supplying division sells to external and internal customers. Under such conditions, the external market price generally exceeds the marginal cost of supplying the intermediate product to the downstream division. Valuation of internal sales at the prevailing market price will therefore result in double marginalization, a problem commonly associated with vertical supply chains. This basic deficiency of valuing transfers at the external market price leads us to examine whether (1) internal transactions should be subjected to intracompany discounts and (2) discount rules can be designed so that the resulting prices and sales quantities maximize the firm’s overall profit.

For the competitive market scenario considered in the earlier literature, a maintained assumption is that firms are capacity constrained. In our scenario of price-setting firms, it is natural to assume that capacity investments have been chosen so that, depending on the stages of the business cycle, capacity at the upstream division can either be constrained or not. With unconstrained capacity, a policy of setting the transfer price equal to the external market price will induce the upstream division to set a price that is an average of the internal and external monopoly prices. One would then expect intracompany discounts to be effective in providing an internal valuation closer to marginal cost as determined by the variable production cost. This intuition is

---

5 The study by Price Waterhouse [1984] reports that 17% of the respondents value transfers at the prevailing market price, while 27% rely on adjusted market prices. For the companies surveyed by Eccles [1985] the discounts range from 5% to 40%, with apparently higher discounts for proprietary components like microelectronic parts. Without quantifying the magnitudes of discounts, Schnell [1995] describes categories of production cost and expenses that the German chemical company BASF subtracts from the external market price to arrive at the adjusted internal price.

6 Double marginalization and the attendant output distortions are the central issue in Arya and Mittendorf [2005] and in the case study by Baker and Monsler [1994]. Summarizing his interviews with more than 100 divisional and financial managers, Eccles [1985, p. 93] states: “The greatest concern about market-based transfers expressed by managers in this study was that these transfers would result in uncompetitive pricing of the final good if the selling profit center attempted to obtain the same gross margin percentages as it could get by marking up full cost.”

7 This prediction is consistent with some of the field study evidence on market-based transfer pricing. For example, Eccles [1985, p. 199] quotes a manager of a downstream division within a chemical company who complains that the upstream division “was not soliciting additional external business, which would lower the external market price and thereby lower the transfer price.” Similarly, in the case study by Bastian and Reichelstein [2004] management of the buying division expresses concern that the external market price set by the selling division for particular types of specialty steel is partly driven by the firm’s market-based transfer pricing system.
incomplete, however, because the upstream division will generally raise the external market price in response to mandated internal discounts. Specifically, the external price will move even further away from the external monopoly price if that price was below the internal monopoly price to begin with. Any such losses from external sales have to be netted against gains from internal trade and therefore the resulting net effect of intracompany discounts remains generally ambiguous.\(^8\)

With constrained production capacity, in contrast, intracompany discounts will always improve the firm’s total profit. If the internal transfer price is set equal to the market price, the resulting transfer quantity will be such that the marginal revenue of the downstream division equals the (transfer) price, which always exceeds the external marginal revenue with a downward sloping demand curve. Thus, double marginalization results in a biased allocation of the firm’s scarce production capacity such that the external sales quantity is too high. By mandating intracompany discounts, the firm’s central office effectively reallocates the available capacity so as to increase the firm’s overall revenue. Our results therefore provide an alternative to the usual justification for intracompany discounts as an adjustment that accounts for cost differences between external and internal sales.

To achieve fully efficient outcomes that maximize a firm’s overall corporate profit, the transfer price faced by the internal buyer must be equal to the marginal cost of internal transfers. With constrained capacity, this marginal cost is given by the external marginal revenue, evaluated at the efficient external sales quantity. Therefore the intracompany discount must be set as the inverse of the external price elasticity of demand at the optimal price. Even though in our model the central office does not have sufficient information to determine the efficient price and quantity decisions, we identify conditions under which it can design a suitable intracompany discount rule. Such a rule serves as an effective decentralization device in that the better informed upstream division will in equilibrium be induced to implement the efficient solution.

The possibility of achieving efficient outcomes through market-based transfer pricing hinges critically on the presence of a capacity constraint. Such constraints ensure that the upstream division cannot do better from its own perspective than to choose the efficient external price because it is “wedged in” between the discount rule and the capacity constraint. With unlimited capacity, in contrast, efficiency would require that the external monopoly price be charged to outside buyers while the internal price be equal to the marginal cost of the intermediate product. Yet, for any market-based transfer pricing rule, the upstream division would always find it desirable to raise the price somewhat since the loss in external profits at the

\(^8\) In a different context, Borenstein [1996] describes a related tradeoff in connection with coupon settlements for antitrust violations. When discounts (coupons) are given as compensation to a group of customers, a monopolist will raise the base price for all customers such that the resulting losses for the nondiscount group may exceed the gains for the discount group.
monopoly price would be of second order compared with a first-order gain from internal transactions. It is only in settings in which the external market is large relative to the internal one that suitably chosen discount rules can approximate efficient outcomes. The upstream division will then find it too costly to increase the market price substantially above the external monopoly price in order to obtain a transfer price exceeding its unit variable cost.

Our model framework differs from other recent theoretical studies on transfer pricing along several dimensions. First, we do not adopt a second-best contracting approach in which resource allocations and compensation payments emerge as the outcome of an optimal mechanism design problem. Optimal (second-best) contracting mechanisms are often complex and highly sensitive to the specification of preferences, beliefs, and the information available to the parties. In contrast, we take several organizational features as given, including that the firm is vertically integrated, yet run in a decentralized fashion with division managers seeking to maximize divisional profits. An additional restriction is that transfer prices are derived from external market prices. Toward the end of our analysis, we provide some observations and conjectures about extending our results to second-best contracting frameworks.

In most of the recent theoretical work, intrafirm transactions are unrelated to any external market transactions. In contrast, our study examines whether a firm’s central office can rely on the observed transactions between some of its divisions and external third parties in order to solve internal resource allocation problems. Our approach thus complements earlier studies on “strategic” internal pricing in which a central office distorts internal valuations in order to affect the outcome of oligopolistic competition in the final product market.

The remainder of the paper is organized as follows. Section 2 presents the model. In Section 3 we ask whether the expected corporate profit will be improved by imposing an intracompany discount rather than valuing internal transfers at the market price. Section 4 analyzes the possibility of achieving fully efficient outcomes by way of suitably chosen discount rules. Section 5 summarizes our findings and concludes.

9 For instance, Harris, Kriebel, and Raviv [1982], Kanodia [1993], Wagenhofer [1994], Vaysman [1996, 1998], and Dikolli and Vaysman [2005] emphasize managerial rents due to private information and the attendant distortions in transfer quantities. In contrast, risk sharing considerations are central to the work of Christensen and Demski [1998].

10 This perspective also underlies the transfer pricing studies that have focused on relationship-specific (e.g., cost-reducing) investments. The papers by Edlin and Reichelstein [1995], Ancil and Dutta [1999], and Sahay [2003] fall into this category.

11 See, for instance, the work of Alles and Datar [1998], Hughes and Kao [1997], and Narayanan and Smith [2000]. Recent work by Arya and Mittendorf [2004] allows for oligopolistic competition in the market for the intermediate product. They demonstrate that it may be beneficial to set the internal transfer price as a function of the observed market price as a means of creating a strategic commitment device in the presence of Cournot competition.
2. Model Description

We study the interactions of two divisions (profit centers) within a firm. An upstream division (division 1) sells its output to external customers and to a downstream division (division 2), which further processes it and sells a final product. Initially, internal and external sales are assumed to have the same costs. The divisional revenue and cost functions are parameterized by \( \theta = (\theta_1, \theta_2) \in \Theta \). This (multidimensional) state variable comprises \( \theta_1 \), which determines division 1’s cost and revenue environment, and \( \theta_2 \), which determines the net revenues of division 2. Each component \( \theta_i \) itself is potentially multidimensional and drawn from a compact support. Division 1’s production costs are assumed to be linear in the total quantity with unit variable cost of \( c(\theta_1) \); for an external market price, \( p \), division 1 faces a demand of \( Q_e(p, \theta_1) \) units for its product from external customers. Given an internal transfer of \( q_i \) units of the intermediate product, division 2 will earn a net revenue of \( R_i(q_i, \theta_2) \). This net revenue reflects the gross sales revenue of the downstream division’s final product less applicable costs other than those for the intermediate product.\(^{12}\)

The realization of the state variable \( \theta \) is unobservable to the firm’s central office, but this information is collectively available to the division managers prior to any decision making. Specifically, we assume that the manager of the upstream division observes all elements of \( \theta \), while the manager of the downstream division only observes \( \theta_2 \). Thus, the manager of the supplying division is assumed to know production costs of the intermediate product as well as internal and external revenues, whereas the manager of the downstream division is assumed to know only the net revenue attainable from the final product.\(^{13}\) To streamline the notation in the following derivations, we generally suppress the subscripts of the components of \( \theta \) and simply write \( c(\theta) \), \( Q_e(p, \theta) \), and \( R_i(q_i, \theta) \).

The firm’s total corporate profit associated with an external price \( p \) and an internal transfer quantity of \( q_i \) becomes:

\[
\pi = p \cdot Q_e(p, \theta) + R_i(q_i, \theta) - c(\theta) \cdot [Q_e(p, \theta) + q_i].
\]

This profit is split between the two divisions by the transfer price \( TP \) such that:

\[
\pi_1 = p \cdot Q_e(p, \theta) + TP \cdot q_i - c(\theta) \cdot [Q_e(p, \theta) + q_i]
\]

\(^{12}\)As discussed in the introduction, the specification that the internal buyer cannot obtain the intermediate product from outside sellers may reflect the proprietary nature of the product. In connection with the Paine Chemical Company, Eccles and White [1988, p. S35] cite managers who allege high transaction cost for switching to a potential outside supplier.

\(^{13}\)Since the only decision made by the downstream division is to choose the quantity \( q_i \) for a given transfer price and \( \theta_2 \), our results would be unaffected if the downstream manager also were to observe the entire vector \( \theta \). In that case, however, market-based transfer pricing would be dominated by a different mechanism, in which the downstream division is given authority to make a take-it-or-leave-it offer to the upstream division.
and

\[ \pi_2 = R_i(q_i, \theta) - TP \cdot q_i. \]

Pricing and production decisions are made in the following sequence. Initially, the central office specifies a transfer pricing rule \( TP(p) \), which determines the transfer price as a function of the external market price.\(^{14}\) Having observed the actual state \( \theta \), the upstream division chooses the external sales price \( p \), which results in external demand \( Q_e(p, \theta) \). The downstream division is given authority to choose the internal transfer quantity \( Q_i(TP, \theta) \). For internal accounting purposes, the transfer quantity is valued at \( TP(p) \) per unit.

For a given transfer pricing rule, \( TP(\cdot) \), and a current state, \( \theta \), the selling division will choose the external sales price, \( p(\theta) \), so as to maximize its divisional profit:

\[
\pi_1(p, \theta | TP(\cdot)) = [p - c(\theta)] \cdot Q_e(p, \theta) + [TP(p) - c(\theta)] \cdot Q_i(TP(p), \theta),
\]

subject to:

\[ Q_i(TP, \theta) \in \arg\max_{q_i} \{R_i(q_i, \theta) - TP \cdot q_i\}, \]

and

\[ Q_i(TP(p), \theta) + Q_e(p, \theta) \leq K(\theta), \]

where \( K(\theta) \) denotes a capacity constraint at the upstream division. Thus, the available capacity may vary with the underlying state \( \theta \).\(^{15}\) In managing its capacity, the upstream division must give priority to the internal buyer before selling to outside parties. Throughout our analysis the demand functions are assumed to be (almost everywhere) differentiable in the market price \( p \) or in the transfer price \( TP \), respectively. Denoting the external revenue function by \( R_e(q, \theta) \equiv q \cdot P_e(q, \theta) \), where \( P_e(\cdot, \theta) \) is the inverse of \( Q_e(p, \theta) \) for a given \( \theta \), we also require that both revenue functions are concave in their respective quantity arguments.

To summarize our model setup of market-based transfer pricing, the upstream division sets the external sales price, which then also determines the internal transfer price via the endogenously chosen rule \( TP(p) \). One possibility for the firm’s central office is not to adjust the external price for

\(^{14}\) In comparing alternative transfer pricing rules, the firm’s central office will generally use its beliefs regarding the state variable \( \theta \) to maximize the expected firm-wide profit. For much of our analysis, these beliefs do not play a role since our results hold pointwise, that is, for all states \( \theta \).

\(^{15}\) In terms of information, it is natural to assume that \( K(\cdot) \) is a function of \( \theta_1 \), the state variable pertaining to division 1’s environment. As before, we suppress the subscript and simply write \( K(\theta) \).
internal valuation purposes, in which case \( TP(p) \equiv p \). While discount rules amount to an administered form of price discrimination, our framework does not include the possibility of free (or third-degree) price discrimination, that is, it does not allow the selling division to set entirely separate prices internally and externally. Below we offer some observations on how such free price discrimination compares with our representation of market-based transfer pricing.

3. Intracompany Discounts

This section examines whether a policy of valuing internal transfers at the prevailing market price can be improved upon by imposing an intracompany discount. To that end, we focus first on discount rules that determine the transfer price in proportion to the external market price, that is, \( TP(p) = (1 - \gamma) \cdot p \). Given the discount factor \( \gamma \) and the state \( \theta \), the upstream division selects an external market price \( p(\gamma, \theta) \) to solve the optimization program in equation (2). We assume first that the firm’s central office has sufficient information (e.g., macroeconomic and/or industry variables) about the underlying environment to know that capacity will be scarce, in that in the absence of a discount the capacity constraint is binding for all states \( \theta \):

**Assumption 1.** The profit function \( \pi_1(\cdot, \theta | \gamma = 0) \) is decreasing in \( p \) at \( p(0, \theta) \) for all \( \theta \).

According to Assumption 1, the upstream division would do better by lowering the price below \( p(0, \theta) \), yet the capacity constraint prevents it from doing so. In particular, A1 implies that for all \( \theta \):

\[
Q_i(p(0, \theta), \theta) + Q_e(p(0, \theta), \theta) = K(\theta).
\]

**Proposition 1.** Suppose capacity is constrained such that Assumption 1 holds. A suitably chosen proportional discount then increases the firm’s corporate profit.\(^{17}\)

With constrained capacity, the corporate objective is to maximize total revenue. 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 the market price, this bias will result in double marginalization.\(^{18}\) As a consequence, the market (and transfer)

---

\(^{16}\) In the following analysis, we are making the implicit assumption that, absent a discount, both markets will be served, i.e., \( Q_i(p(0, \theta), \theta) > 0 \) and \( Q_e(p(0, \theta), \theta) > 0 \).

\(^{17}\) All proofs are provided in Appendix A.

\(^{18}\) This feature of our model is consistent with the observation in Eccles and White [1988, p. S31]: “Buying profit centers that pay market price commonly complain that the intermediate product is being marked up twice, once by the selling profit center and again by the buying profit center.”
price exceeds the marginal cost of delivering another unit of the intermediate product internally. That marginal (opportunity) cost is given by the external marginal revenue because the last unit could have been sold at the current price and, if sold internally, the external price on all previous units would have increased correspondingly.¹⁹ Formally,

\[ R'_i(Q_i(p(0, \theta), \theta), \theta) = p(0, \theta) > R'_e(Q_e(p(0, \theta), \theta), \theta). \]

In response to a (small) discount, γ, the upstream division will be induced to increase the market price. At the same time, the resulting new transfer price, \((1 - \gamma) \cdot p(\gamma, \theta)\), will be below \(p(0, \theta)\). Otherwise, some production capacity would be left idle. Thus, intracompany discounts counteract the original bias in favor of external sales with the consequence of higher total revenues across the two markets.

The intuition for Proposition 1 is particularly transparent when the unrestricted external price exceeds the unrestricted internal price (“unrestricted” here refers to the prices the upstream division would charge if it could freely price discriminate). In this scenario, a discount will induce the upstream division to move closer to the unrestricted solution, and at the same time, the downstream division obtains the intermediate product at a lower price, leaving both divisions better off.

Our model has so far ignored the possibility of cost differences between internal and external sales. As argued in the Introduction, the survey evidence on market-based transfer pricing suggests that cost differences (such as the absence of certain selling and administrative expenses) are cited as the primary rationale for imposing intracompany discounts. It is readily verified that the result of Proposition 1 extends 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. Accordingly, one would expect that optimal intracompany discounts are increasing in cost differences between external and internal sales. We verify this intuition in section 4 below.

The case for intracompany discounts seems intuitively even more compelling in settings in which production capacity at the upstream division is effectively unconstrained and therefore the marginal cost of internal transfers is \(c(\theta)\) and hence less than the external marginal revenue. 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 resulting outcome would be inefficient for two reasons. First, the firm would not take full advantage of its market power externally. Second,

¹⁹ In connection with their Padres Papers example, Maher, Lanen and Rajan [2005, Ch. 14] note that with a downward sloping demand curve the marginal cost of internal transfers is below the market price.
internal transactions would be inefficiently low since the unit cost perceived by the downstream division exceeds the actual marginal cost, $c(\theta)$.20

The ultimate impact of an intracompany discount depends on the shift in the external market price set by the upstream division. As argued above, when capacity is constrained, such price shifts always move the sales quantities in the right direction so as to improve total profit. The following example shows that Proposition 1 does not carry over to settings in which production capacity is effectively unlimited. In terms of notation, let $p^m_j(\theta)$ denote the internal and external monopoly prices, respectively, that the upstream division would charge if it could freely price discriminate between the two markets absent any capacity constraints. Formally, $p^m_j(\theta)$ maximizes:

$$[p - c(\theta)] \cdot Q_j(p, \theta), \quad j = i, e. \tag{3}$$

**EXAMPLE 1.** The downstream division seeks to buy a fixed quantity of the intermediate product, $q_i(\theta)$, possibly because it is faced with an external order of fixed size. As illustrated in Figure 1, internal demand therefore is inelastic up to a reservation price, $p_i(\theta)$, and drops to zero above that price:

$$Q_i(p, \theta) = \begin{cases} q_i(\theta) & \text{if } p \leq p_i(\theta), \\ 0 & \text{otherwise}. \end{cases}$$

Absent a discount, division 1 will set a price $p(\theta) \in [p^m_e(\theta), \hat{p}_i(\theta))$. If the central office stipulates a transfer price of $TP(p) = (1 - \gamma) \cdot p$, division 1 will increase the external market price, that is, $p(\gamma, \theta) \geq p(\theta)$. As a consequence, the external sales price is driven further away from the monopoly price, $p^m_e(\theta)$, without improving internal trade, since the internal quantity will not increase beyond $q_i(\theta)$. Hence, external profit decreases while the contribution to total profit from internal trade remains unaffected.

It is readily seen that in the above example the firm would benefit from imposing an intracompany markup on internal transactions. We 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. Finally, the reasoning in the above example extends to settings in which internal sales are $k$ dollars cheaper per unit. Thus, we conclude that cost differences between internal and external sales are neither necessary nor sufficient for the desirability of intracompany discounts.

The impact of intracompany discounts is unambiguously positive if the external monopoly price exceeds the internal monopoly price (i.e., $p^m_e(\theta) > p^m_i(\theta)$), the discount induces the supplying division to charge a higher external price, and at the same time the resulting internal price decreases. Total profit will then be higher as the supplying division is driven closer to the monopoly price externally and the buying division faces a price closer to the

20 This scenario captures several of the features identified by Eccles and White [1988, p. S32] in connection with the Locke Chemical Company.
firm’s marginal cost. Conversely, a discount will be harmful if $p_e^m(\theta) < p_i^m(\theta)$ and internal trade does not increase. The latter effect may occur because the external price increase in response to the discount is so sharp that the resulting internal transfer price also increases (alternatively, internal trade may not improve because demand is inelastic, as in Example 1). With proportional discounts, this “overreaction” to a discount would require that:

$$p(\gamma, \theta) > (1 - \gamma) \cdot p(\gamma, \theta) > p(0, \theta).$$  \hspace{1cm} (4)

We recall that with constrained capacity the second inequality in equation (4) could not arise or otherwise some capacity would have been left idle at the adjusted market price.

For further examination of the tradeoffs associated with intracompany discounts under conditions of unconstrained capacity, we now consider a setting in which both internal and external demand can be described by linear functions:

$$Q_e(p, \theta) = \alpha_e(\theta) - \beta_e(\theta) \cdot p \quad \text{and} \quad Q_i(TP, \theta) = \alpha_i(\theta) - \beta_i(\theta) \cdot TP.$$  \hspace{1cm} (5)

We denote the corresponding external willingness-to-pay curve by $P_e(q_e, \theta) = a_e(\theta) - b_e(\theta) \cdot q_e$, so that:

$$a_e(\theta) = \frac{\alpha_e(\theta)}{\beta_e(\theta)} \quad \text{and} \quad b_e(\theta) = \frac{1}{\beta_e(\theta)}.$$

In accordance with equation (5), the buying division’s net revenue from the intermediate product is given by $R_i(q_i, \theta) = [a_i(\theta) - \frac{1}{2} \cdot b_i(\theta) \cdot q_i] \cdot q_i$, so that:
\[ R_i(q_i, \theta) = a_i(\theta) - b_i(\theta) \cdot q_i, \quad a_i(\theta) = \frac{\alpha_i(\theta)}{\beta_i(\theta)}, \quad \text{and} \quad b_i(\theta) = \frac{1}{\beta_i(\theta)} \cdot \]

To avoid corner solutions, we restrict attention to parameter values such that, absent any discount, the seller finds it advantageous to charge a price that attracts both external and internal sales. Specifically, this requires that for all \( \theta \):

\[
p(0, \theta) = \frac{1}{2} \left[ \frac{\alpha_e(\theta) + \alpha_i(\theta)}{\beta_e(\theta) + \beta_i(\theta)} + c(\theta) \right] < \min \{ a_e(\theta), a_i(\theta) \}. \tag{6}
\]

The expression on the left-hand side of the above inequality is the average monopoly price when the demand curves in both markets are linear. Provided this price is below the intercept of both the internal and the external willingness-to-pay curves, the upstream division’s profit is maximized at \( p(0, \theta) \) if internal transfers are valued at the prevailing market price.

**Proposition 2.** Suppose capacity is unconstrained and both internal and external demand are given by linear functions such that inequality (6) holds. A suitably chosen proportional discount then increases the firm’s corporate profit.

Contrary to the message of Example 1, we find that discounts are unambiguously beneficial with linear demand curves regardless of the relative magnitudes of the two markets and the respective price elasticities. The key to Proposition 2 is that discounts increase not only the external but also the internal price for certain parameter configurations (i.e., inequality (4) may hold), yet this range of parameters is precisely the one in which \( p_m^e(\theta) > p_m^i(\theta) \). Discounts thus increase the profit from external trade and this effect dominates the loss associated with lower internal trade in this case. Conversely, when \( (1 - \gamma) \cdot p(\gamma, \theta) < p(0, \theta) \), the internal gains always outweigh potential losses in the external market with linear demand curves.

The intuition for our finding in Proposition 2 becomes particularly transparent if discounts are calculated in an additive fashion, that is, the transfer price is equal to the market price less a constant amount. Linearity of the demand curves then implies that the total quantity across the two markets is invariant to the magnitude of the discount.\(^{21}\) Therefore, total production costs are unaffected by discounts and, like in the constrained capacity setting, a small additive discount increases the overall corporate profit if it increases firm-wide revenues. Yet, as demonstrated in Proposition 1, double marginalization leads to sales quantities for which the internal marginal revenue always exceeds the external marginal revenue. A discount counteracts that bias and hence increases firm-wide profit.

To conclude this section, we note that our results also speak to a setting in which capacity is constrained for certain states \( \theta \in \Theta \) but not for others.

\(^{21}\) 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 “conservation property” to hold.
Provided the demand curves can be described by linear functions, Propositions 1 and 2 taken together show that discounts unambiguously improve overall corporate profits independent of whether capacity is constrained.\textsuperscript{22}

4. Implementing Efficient Outcomes

In this section, we turn to the more ambitious question of whether a multidivisional firm can achieve efficient decentralization by subjecting market prices to some suitable intracompany discount rule. With constrained capacity, efficiency requires that the upstream division be induced to set the market price so that the resulting allocations maximize the overall revenue attainable to the firm. In particular, the resulting transfer price must be equal to the external marginal revenue evaluated at the efficient external sales quantity.

With unconstrained capacity, in contrast, an efficient outcome requires the upstream division to charge the monopoly price externally, and transfer internally at marginal cost, $c(\theta)$. Stated differently, the internal discount should be the inverse of the optimal monopoly markup over cost in the external market. To illustrate, suppose the firm’s central office knows that the price elasticity of demand in the external market is constant and equal to some $\epsilon > 1$. In that case, a discount equal to $\gamma = \frac{1}{\epsilon}$ would result in a transfer price equal to the variable production cost $c(\theta)$, if the supplying division were to choose the external monopoly price. However, the upstream division would respond to this discount with a higher external price, since the loss of external profits at the monopoly price is of second order compared to a first-order gain from internal transactions. The following result formalizes this argument for general demand scenarios.

ASSUMPTION 2. The function $c(\theta)$ is differentiable such that $\nabla c(\theta) \neq 0$ for all $\theta$ in some neighborhood $U \subset \Theta$.

PROPOSITION 3. Given Assumption 2 and unconstrained capacity, there does not exist a transfer pricing rule $TP(\cdot)$ that induces efficient sales quantities.

Assumption 2 is essentially a minimal condition for the impossibility result in Proposition 3. If the firm’s central office knew the supplier’s marginal cost, that is, the unit variable cost did not vary with the underlying state, then the rule $TP(p) \equiv c$ would result in efficient outcomes.

Since Proposition 3 shows the impossibility of attaining fully efficient outcomes through market-based transfer pricing when internal transfers must be valued at the seller’s unit variable cost, it is instructive to examine how

\textsuperscript{22} Baldenius, Melumad and Reichelstein [2004] focus on linear demand curves to study the impact of income taxes on transfer prices. In their analysis, internal transfers are valued at the external market price for tax purposes (CUP method). An additional advantage of discounts then is that by inducing a higher external market price, the firm achieves tax savings provided the downstream division faces a higher income tax rate than the upstream division.
intracompany discounts should be set so as to balance the attendant trade-offs. This analysis is presented in Appendix B for linear demand functions and constant discounts. As noted in connection with Proposition 2, it then turns out that the total quantity, \( q_i + q_e \), is invariant to the discount and therefore optimal discounts have the property that internal and external marginal revenues are equal in expectation over all states \( \theta \). The analysis in Appendix B also shows that any cost differentials between internal and external transactions should not be passed on fully to the buyer via intracompany discounts. The reason essentially is that a lower cost of internal sales will already be partly reflected in the external market price chosen by the seller. Incorporating cost differentials fully into the discount would thus amount to double-counting.

As mentioned above, our notion of market-based transfer pricing is that the internal valuation is “pegged” to the price set externally by the upstream division. The impossibility result in Proposition 3 naturally raises the question whether a complete “laissez-faire” policy, according to which the upstream division is permitted to freely price discriminate across the two markets, could be more effective. Such an approach will frequently dominate a policy of valuing transfers at the prevailing market price (i.e., \( TP(p) \equiv p \)), provided \( p^m_e(\theta) > p^m_i(\theta) \) for all \( \theta \). The reason is that unadjusted market-based transfer pricing will result in a market (and transfer) price that is an average of \( p^m_e(\theta) \) and \( p^m_i(\theta) \). As a consequence, both external and internal profits will be distorted and, subject to certain regularity conditions, both these distortions could be mitigated by allowing for free price discrimination. It remains a question for future research to compare such a laissez-faire policy with a policy of valuing internal transfers at adjusted market prices.

We now turn to a setting in which production capacity is effectively constrained to examine whether suitably chosen discount rules can fully eliminate the double marginalization problem. Let \( q^*_i(\theta) \) and \( q^*_e(\theta) \) denote the quantities that maximize the firm’s overall profit subject to the capacity constraint, that is, \( q^*_i(\theta) + q^*_e(\theta) = K(\theta) \). As before, the monopoly prices the upstream division would set if it could freely price discriminate (absent any capacity constraint) are denoted by \( p^m_i(\theta) \). Let \( q^m_i(\theta) \) and \( q^m_e(\theta) \) denote the corresponding monopoly quantities, that is, \( q^m_i(\theta) \equiv Q_i(p^m_i(\theta), \theta) \) and \( q^m_e(\theta) \equiv Q_e(p^m_e(\theta), \theta) \).

ASSUMPTION 3. Both the internal and external monopoly pricing problem are single peaked in \( p \).

---

23 For reasons of analytical tractability, Appendix B is confined to an examination of additive discounts. It remains an open question which particular form of discounts—i.e., proportional or additive—performs better. Sahay [2003] compares the effectiveness of additive and multiplicative markups over cost in connection with cost-based transfer pricing.

24 We thank an anonymous reviewer for raising this issue.
ASSUMPTION 4. $q^*_i(\theta) < q^m_i(\theta)$ and $q^*_e(\theta) < q^m_e(\theta)$ for all $\theta$.

Since the monopoly quantities refer to a setting with unlimited capacity, Assumption 4 will be satisfied for any given demand curves provided the capacity constraint is sufficiently tight.

A necessary condition for eliminating the double marginalization problem is that the upstream division has an incentive to choose the optimal price $P_e(q^*_e(\theta), \theta)$ and the discount is set inversely to the price elasticity of demand at $P_e(q^*_e(\theta), \theta)$. The downstream division will then face the “correct” transfer price equal to the external marginal revenue at the optimal quantity $q^*_e(\theta)$. To formalize this notion, we denote the price elasticity of external demand by:

$$
\epsilon(p, \theta) = -\frac{dQ_e(p, \theta)}{dp} \cdot \frac{p}{Q_e(p, \theta)}
$$

and, for brevity, define $p^*(\theta) \equiv P_e(q^*_e(\theta), \theta)$ and $\epsilon^*(\theta) \equiv \epsilon(p^*(\theta), \theta)$.

PROPOSITION 4. Suppose Assumptions 3 and 4 hold. If there exists a function $f(\cdot)$ such that $f(p^*(\theta)) = \epsilon^*(\theta)$ for all $\theta$, the transfer pricing rule

$$
TP(p) = p \cdot \left[1 - \frac{1}{f(p)}\right]
$$

induces the upstream division to set a market price that implements the efficient sales quantities, provided $TP(p)$ is increasing in $p$.

Proposition 4 postulates an invertibility property requiring that for every external price $p^*(\theta)$ there is at most one corresponding price elasticity $\epsilon^*(\theta)$. To see why this is sufficient, suppose that the upstream division does indeed choose the optimal external market price $p^*(\theta)$. The corresponding external marginal revenue is, by definition, equal to $p^*(\theta) \cdot [1 - \frac{1}{\epsilon^*(\theta)}]$. By construction of the transfer pricing rule, the downstream division will demand the quantity $q^*_e(\theta)$ so as to equate its marginal revenue with the transfer price, which in turn is equal to:

$$
p^*(\theta) \cdot \left[1 - \frac{1}{f(p^*(\theta))}\right] = p^*(\theta) \cdot \left[1 - \frac{1}{\epsilon^*(\theta)}\right].
$$

Under Assumptions 3 and 4, the upstream division cannot do better by raising the price above $p^*(\theta)$. Externally, the upstream division would move further away from the monopoly price, $p^m_e(\theta)$, because $q^*_e(\theta) \leq q^m_e(\theta)$. The same argument applies to internal sales provided $TP(p)$ is indeed increasing in $p$. On the other hand, the capacity constraint prevents the upstream division from lowering the external market price below $p^*(\theta)$.

The sharp contrast between Propositions 3 and 4 underscores the centrality of capacity constraints for the performance of market-based transfer pricing. With constrained capacity, it is possible to “wedge in” the upstream division between the discount and the capacity constraint. 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. In contrast, such deviations only cause a second-order loss to an upstream division with idle capacity.

It is straightforward to construct examples in which the state variable \( \theta \) is of low dimensionality and the conditions of Proposition 4 are satisfied. \(^{25}\) When the state of the world is very complex (i.e., \( \theta \) is multidimensional), on the other hand, it will frequently be impossible to infer the price elasticity of demand from the price charged by the upstream division. However, an important application of Proposition 4 is the case in which the price elasticity of demand does not depend on \( \theta \) directly but only through the price \( p^*(\theta) \). This will be the case if the underlying state variable shifts the external demand function in a multiplicative fashion.

**Proposition 5.** Suppose Assumptions 3 and 4 hold. If \( Q_e(p, \theta) = a_e(\theta) \cdot b_e(p) \), the transfer pricing rule

\[
TP(p) = p \left[ 1 - \frac{1}{\epsilon(p)} \right],
\]

with \( \epsilon(p) \equiv -b'_e(p) \cdot p/b_e(p) \), induces the upstream division to set a market price that implements the efficient sales quantities.

As shown in the proof of Proposition 5, the transfer pricing function in equation (8) is indeed monotone increasing in \( p \) by virtue of the assumption that the external marginal revenue is decreasing in \( q_e \) for all \( \theta \). The curvature of the optimal transfer pricing function is given by the curvature of \( b_e(p)/b'_e(p) \). For the special case in which \( b_e(p) = \exp(-v \cdot p) \), for some \( v \geq 0 \), we conclude that the optimal discount is an additive constant so that \( TP(p) = p - \frac{1}{v} \). Simple proportional discounts, on the other hand, are optimal if the external price elasticity is constant and independent of the underlying state \( \theta \).

**Corollary 1.** If the external market demand exhibits constant price elasticity, that is, \( Q_e(p, \theta) = a_e(\theta) \cdot p^{-\epsilon} \) with \( \epsilon > 1 \), a proportional discount of \( \gamma^o = \frac{1}{\epsilon} \) achieves efficient outcomes.

In the industrial organization literature, the price elasticity of demand is generally viewed as a measure of the distortion associated with monopoly pricing. Our finding here is consistent with this perspective in the sense that transferring at market price becomes approximately efficient as the external demand becomes more elastic, that is, as \( \epsilon \) approaches one. While the above

\(^{25}\) To illustrate, suppose that the only variable that depends on \( \theta \), and therefore is unknown to the firm’s central office, is the capacity constraint. Without loss of generality, let \( K(\theta) = K + \theta \) for \( \theta \) in some interval \([\theta, \bar{\theta}]\). It follows that \( q_e^*(\theta) \) is increasing in \( \theta \), while both \( p^*(\theta) \) and \( R_e'(q_e^*(\theta)) \) are decreasing in \( \theta \). For the transfer pricing rule identified in Proposition 4, we find that \( TP(p^*(\theta)) = R_e'(q_e^*(\theta)) \), for all \( \theta \). Because both \( p^*(\theta) \) and \( R_e'(q_e^*(\theta)) \) are monotone decreasing in \( \theta \), this implies that \( TP(p) \) is monotone increasing in \( p \), as postulated in Proposition 4.
findings identify optimal discounts with identical costs, it is instructive to ask how optimal discounts are affected by cost differences between internal and external sales. Like in Section 3, suppose that selling the intermediate product internally is $k$ dollars cheaper per unit than external sales, that is, $c_i(\theta) = c(\theta) - k$. Such cost differences effectively raise the revenues from internal sales by $k$ per unit and, therefore, an optimal allocation of the firm’s capacity requires that the external marginal revenue exceed the internal marginal revenue by $k$.

**COROLLARY 2.** Suppose the external market demand exhibits constant price elasticity, $\epsilon$, and the unit variable cost of internal transfers is $c_i(\theta) = c(\theta) - k$. The transfer pricing rule

$$TP(p) = p \cdot \left[1 - \frac{1}{\epsilon}\right] - k$$

then achieves efficient outcomes.

According to equation (9), the optimal transfer pricing rule is affine, involving both a constant and a proportional component.\footnote{Corollary 2 requires that Assumptions 3 and 4 apply to the functions $[p - c(\theta) - k] \cdot Q_j(p, \theta)$.} Equation (9) may prove useful in future empirical work seeking to relate the magnitude of observed discounts to both cost differences and the intermediate product’s external price elasticity of demand.

The above results extend without change to firms in which multiple internal divisions seek to buy the intermediate product. Assuming that these buyers do not compete directly with each other in selling their final products (and ignoring their further processing costs), the firm seeks an allocation of its scarce capacity that will maximize the sum of the divisional revenues. This allocation problem can effectively be decentralized by letting the upstream division set the external price and instructing the internal divisions to demand quantities that maximize their respective divisional profits. Provided all internal divisions face the same transfer price, which is obtained by applying the discount identified above to the external market price, the resulting sales quantities will maximize total firm profit.\footnote{We thank Joseph Harrington for suggesting this extension.}

Corollary 1 naturally raises the question of how a proportional discount equal to $\gamma^o = \frac{1}{\epsilon}$ fares if the external price elasticity, $\epsilon$, is constant and capacity is unconstrained. If the upstream division was to charge the external monopoly price $P^m_i(\theta)$, transfers would indeed be valued at the unit variable cost, since the discount factor is precisely the inverse of the monopoly markup, that is:

$$P^m_i(\theta) \cdot \left[1 - \frac{1}{\epsilon}\right] = c(\theta).$$

As observed in connection with Proposition 3, however, the upstream division will raise the external price beyond $P^m_i(\theta)$ in order to make a profit
internally. In particular, the upstream division chooses \( p(\gamma^o, \theta) \) so as to solve:

\[
\max_p \left\{ \left[ p - c(\theta) \right] \cdot a_e(\theta) \cdot p^{-\epsilon} + \left[ (1 - \gamma^o) \cdot p - c(\theta) \right] \cdot Q_e((1 - \gamma^o) \cdot p, \theta) \right\}.
\]

We note that the resulting deviation from \( p_m(\theta) \) will be small if the external market is large relative to the internal one. In particular, with constant price elasticity, the parameter \( a_e(\theta) \) reflects the size of the external market, yet this factor does not affect the external monopoly price. It is readily seen that \( p(\gamma^o, \theta) \) will be close to \( p_m(\theta) \) as \( a_e(\theta) \) becomes large, because the external market then creates an effective “countervailing incentive” that makes it too costly for the upstream division to choose an external price that substantially exceeds the monopoly price. We conclude that for a constant external price elasticity of demand a discount equal to \( \frac{1}{\epsilon} \) works well in either one of two nonexclusive scenarios: the supplying division’s capacity is constrained or the external market is large compared to the internal one.\(^{28}\)

As noted in the Introduction, our model has taken it as given that divisional managers seek to maximize the profits of their own divisions. While this specification is descriptive of practice in most firms, we conclude this section by sketching two principal-agent models in which transfer pricing and the need for divisional performance evaluation emerge endogenously. First, suppose the managers choose “general purpose” efforts such that the cash flow of the downstream division is given by \( CFL_2 = x_2 + R_i(q_i, \theta) \), while the upstream division’s cash flow is \( CFL_1 = x_1 + p \cdot Q_e(p, \theta) - c(\theta) \cdot (Q_e(p, \theta) + q_i) \). Managers choose \( x_j \) at a personal cost of \( D_j(x_j, \mu_j) \), respectively, with \( \mu_j \) denoting a (one-dimensional) productivity parameter that represents the private information of manager \( j \). For contracting purposes, the firm’s central office can observe the divisional cash flows without being able to disentangle the components of these cash flows. In contrast to the state variable \( \theta \), which managers collectively learn after contracting, the productivity parameters \( \mu_j \) are pre-contractual private information and therefore become a source of informational rents for the managers.

The firm’s central office then faces a combined contracting and resource allocation problem that can be solved with standard techniques, as in Laffont and Tirole [1993].\(^{29}\) Provided the cost functions \( D_j(x_j, \mu_j) \) are multiplicatively separable, it can be shown that the results in Propositions 4 and 5 (and their corollaries) emerge as part of an optimal second-best mechanism with the following properties. Managers are offered a menu of contracts, each one of which is linear in their own division’s income (choosing from this menu reveals the private information \( \mu_j \)). By rewarding managers with a

\(^{28}\) For instance, in the case study of Bastian and Reichelstein [2004], the upstream steel division sells more than 80% of its output externally. Depending on the relative elasticities in the two markets, the transfer pricing rule \( TP(p) = p \cdot [1 - \frac{\epsilon}{\epsilon}] \) can then achieve up to 95% of the theoretically attainable overall profit, according to simulations omitted here.

\(^{29}\) Edlin and Reichelstein [1995] provide details of the solution to a contracting problem similar to the one sketched here.
share of their divisional incomes and designing internal discount rules as specified above, the central office can implement the second-best solution to the moral hazard problem and, at the same time, motivate managers to make the desired pricing and internal transfer decisions. An important advantage of this delegation scheme is that it requires less communication than a (centralized) revelation mechanism, since the pricing and internal resource allocation decisions are delegated to the divisional managers.30

While the preceding model rationalizes the use of divisional performance evaluation, it requires a high degree of separability between the moral hazard and the resource allocation problem. An alternative mechanism design problem, like in Vaysman [1996, 1998], is one in which managers choose their respective \(\theta_j\) values at a personal cost of \(D_j(\theta_j, \mu_j)\), with \(\mu_j\) again representing private pre-contractual information. The divisional income measures are \(\pi_1 = p \cdot Q_e(p) + TP(p) \cdot q_i - c(\theta_1) \cdot (Q_e(p) + q_i)\) for the upstream division and \(\pi_2 = R_i(q_i, \theta_2) - TP(p) \cdot q_i\) for the downstream division.31 Unlike the “general effort” scenario above, the downstream manager’s choice variable, the revenue parameter \(\theta_2\), now affects the price \(p(\cdot)\) set by the upstream manager. When \(\theta_2\) is a choice variable, it will generally be a function of the downstream manager’s (reported) type: More efficient internal buyers in equilibrium will be asked to implement a higher \(\theta_2\), resulting in higher internal demand for the intermediate good and ultimately in a higher price \(p(\cdot)\). The problem of truthfully eliciting \(\mu_2\) from the downstream manager therefore becomes more intricate because he now has a twofold incentive to misreport his type as being less efficient: (1) to elicit higher information rents from the principal and (2) to induce the selling manager to lower the external price \(p\) and thereby also the transfer price.

While an exhaustive characterization of the optimal revelation (and delegation) mechanism is beyond the scope of the present paper, we conjecture that the severity of the price manipulation incentive under delegation will depend on the relative slopes of the two demand functions for the intermediate good. If the external demand \(Q_e(\cdot)\) is very price sensitive, the resulting external price \(p\) will not vary much across different \(\mu_2\)-reports submitted by manager 2 (the selling division would otherwise sacrifice too much profit from external sales). In that case, our results obtained above are likely to generalize in the sense that a delegation scheme with market-based transfer

---

30 An alternative contracting arrangement would be to pay each manager a share of the total corporate profit. While this performance measure potentially sidesteps any issues of transfer pricing, an immediate drawback is that each manager is subjected to the total variability of cash flows. Even with risk-neutral agents, excessive variability of the performance metric would become an issue for contracting in the presence of limited liability constraints.

31 Assuming that the external demand function \(Q_e(\cdot)\) is deterministic (i.e., unaffected by \(\theta_1\)) restricts the generality of this setup but is essential for technical reasons. If manager 1 were to observe two private information parameters (i.e., \(\theta_1\) is two-dimensional with one component affecting costs, the other affecting external demand), then this would yield a multidimensional screening model. As discussed in Rochet and Stole [2003], such models are generally intractable except for special cases.
prices can replicate the optimal revelation mechanism, provided capacity is constrained. Such a replication result may no longer be obtained if $Q_e(\cdot)$ is relatively price insensitive, because the buyer’s price manipulation incentive will then become too severe.

5. Concluding Remarks

Multidivisional firms frequently make adjustments to external market prices in order to value internal transactions. In addition to possible cost differences between internal and external transactions, we find that imperfect competition for the intermediate product in question frequently makes intracompany discounts desirable. Cost differences are neither necessary nor sufficient for intracompany discounts to improve overall corporate profits. The impact of discounts and the overall efficiency of market-based transfer pricing depends critically on the presence of capacity constraints at the upstream division. Fully efficient outcomes can be attained via (properly adjusted) market-based transfer pricing only if production capacity is constrained or the external market is large relative to the internal one.

Our analysis suggests several directions for future research. One limitation of our model is that we have ignored the possibility that the downstream division and the external buyers may compete with each other in the final product market. One would expect that such downstream competition lends additional impetus to intracompany discounts. It remains to be seen, though, whether the presence of capacity constraints again plays a central role in affecting the magnitude of the desired discount.

The field studies by Eccles [1985] and others suggest that issues of “sourcing policy” arise when an intermediate product or service can either be supplied by the internal upstream division or by a limited number of outside suppliers. With imperfect competition for the intermediate product, the immediate question becomes whether the internal buyer should be free to source the product externally and whether internal transfers should be valued at the external price quotes. While such policies are in use at some companies, others insist on “mandated internal transfers” despite the presence of potential external suppliers.\(^{32}\)

For multinational firms, a major advantage of market-based transfer pricing is that this pricing method is also acceptable for statutory purposes. Most tax authorities allow the comparable uncontrolled price (CUP) method for arriving at “arm’s length” prices to value cross-border transactions between parent companies and foreign subsidiaries. If the prevailing market price is used for tax purposes, the firm still has the option to use adjusted market prices for internal profit measurement. Aside from the allocative function studied in this paper, intracompany discounts then also become

---

\(^{32}\) A policy of mandatory internal sourcing may reflect the importance of relationship-specific investments. Related issues have been explored in the industrial organization literature in connection with “exclusive dealing clauses;” see, for example, Segal and Whinston [2000].
an instrument for reducing the firm’s tax liability. For instance, if the upstream division is located in a low-tax jurisdiction, larger intracompany discounts may result in higher external market prices (and tax transfer prices), with the consequence of income being shifted to the low-tax jurisdiction. Such considerations naturally complicate the characterization of optimal discounts.33

Aside from market-based transfer pricing, most firms resort either to cost-based or negotiated transfer prices for internal valuation purposes. To date, the literature is lacking a comprehensive theory as to which of these alternative policies a particular firm is likely to adopt. In addition to external competition (in both the intermediate and the final goods market), one would expect the need for relationship-specific investments, the frequency of internal transactions, and the distribution of information about costs and revenues to jointly shape the tradeoff among these alternative policies. The results of this paper add several pieces to the transfer pricing puzzle by identifying settings in which market-based transfer pricing performs well despite the problems associated with imperfectly competitive external markets.

APPENDIX A

Proof of Proposition 1. By Assumption 1 division 1’s profit function is decreasing in \( p \) at \( p(0, \theta) \). By continuity, this also holds for small values of \( \gamma \). For small \( \gamma \), division 1 will thus set the external price \( p(\gamma, \theta) \) so that

\[
Q_i(\hat{p}(\gamma, \theta), \theta) + Q_e(p(\gamma, \theta), \theta) = Q_i(p(0, \theta), \theta) + Q_e(p(0, \theta), \theta) = K(\theta),
\]

where \( \hat{p}(\gamma, \theta) \equiv (1-\gamma) \cdot p(\gamma, \theta) \). This implies that \( \hat{p}'(0, \theta) < 0, \hat{p}'(0, \theta) > 0, \) and \( dQ_i/d\gamma > 0 \) at \( \gamma = 0 \).

Given Assumption 1, firmwide profit for any \( \theta \) and any small \( \gamma \) equals:

\[
\pi(\gamma, \theta) \equiv Ri(\hat{p}(\gamma, \theta), \theta) + Re(K(\theta) - Q_i(\hat{p}(\gamma, \theta), \theta), \theta) - c(\theta) \cdot K(\theta),
\]

and

\[
\frac{d\pi(\gamma, \theta)}{d\gamma} \bigg|_{\gamma=0} = [R'_i(Q_i(p(0, \theta), \theta), \theta) - R'_e(Q_e(p(0, \theta), \theta), \theta)] \cdot \frac{dQ_e}{d\gamma},
\]

which is indeed positive since \( dQ_e/d\gamma > 0 \) at \( \gamma = 0 \) 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. \( \Box \)

33 Recent work by Narayanan and Smith [2000], Smith [2002], Korn and Lengsfeld [2003], Baldenius, Melumad, and Reichelstein [2004], Hyde and Choe [2005], and others has begun to explore the link between divisional income taxes, arm’s length prices, and managerial transfer prices.
Proof of Proposition 2. We demonstrate that, for all $\theta$, total corporate profit $\pi(\gamma, \theta) \equiv \pi_1(p(\gamma, \theta), \theta | \gamma) + \pi_2((1 - \gamma) \cdot p(\gamma, \theta), \theta)$ is increasing in $\gamma$ at zero, that is:

$$\frac{d}{dy} \pi(y, \theta) \bigg|_{y=0} > 0. \quad (A1)$$

With a constant percentage discount, divisional profit of the upstream division is:

$$\pi_1(p, \theta | \gamma) = \pi_i((1 - \gamma) \cdot p, \theta) + \pi_e(p, \theta) = [(1 - \gamma) \cdot p - c(\theta)] \cdot Q_i((1 - \gamma) \cdot p, \theta) + [p - c(\theta)] \cdot Q_e(p, \theta),$$

while that of the downstream division is given by:

$$\pi_2((1 - \gamma) \cdot p, \theta) = R_i(Q_i((1 - \gamma) \cdot p, \theta), (1 - \gamma) \cdot p \cdot Q_i((1 - \gamma) \cdot p, \theta)).$$

Since $p(\gamma, \theta)$ maximizes $\pi_1$ and $Q_i$ is chosen so that the internal marginal revenue is equal to the transfer price, that is, $R_i'(Q_i(\cdot), \theta) = (1 - \gamma) \cdot p(\gamma, \theta)$, the Envelope Theorem implies:

$$\frac{d}{dy} \pi(y, \theta) \bigg|_{y=0} \equiv \pi'(0, \theta) = -[p(0, \theta) - c(\theta)] \cdot Q_i'(p(0, \theta), \theta) + Q_i(p(0, \theta), \theta) - [p'(0, \theta) - p(0, \theta)] \cdot Q_e(p(0, \theta), \theta).$$

To show that $\pi'(0, \theta) > 0$ for all $\theta$, it will be notationally convenient to suppress the dependence on $\theta$. Therefore, we write $Q_i = \alpha_i - \beta_i \cdot p$ and $Q_e = \alpha_e - \beta_e \cdot p$, where $\alpha_i \equiv a_i \cdot \beta_i$, $\beta_i \cdot \beta_i \equiv 1$, $\alpha_e \equiv a_e \cdot \beta_e$, and $\beta_e \cdot \beta_e \equiv 1$. Collecting terms, we obtain:

$$\pi'(0) = \beta_i \cdot [(p(0) - c) \cdot p(0) - p'(0) \cdot (a_i - p(0))]. \quad (A2)$$

The average monopoly price, $p(0)$, as given in equation (6), can be rewritten as:

$$p(0) = \frac{1}{2} \cdot [v \cdot a_i + (1 - v) \cdot a_e + c],$$

where $v = \beta_i/(\beta_i + \beta_e)$. Straightforward algebra yields:\[34\]

$$p'(0) = v \cdot \left[ \left( v - \frac{1}{2} \right) \cdot a_i + (1 - v) \cdot a_e + \frac{1}{2} \cdot c \right]. \quad (A3)$$

\[34\] We can validate here our claim in the text that a multiplicative discount can lead to over-compensation in the sense that the resulting transfer price exceeds the original market price. We note that $TP(\gamma) > 0$ at $y = 0$ if and only if $p'(0) > p(0)$. The last inequality holds, for instance, if $c = 0$ and $(v - 1/2)/v \cdot a_e > a_i$. \[34\]
To evaluate the expression on the right hand side of equation (A2), it will be convenient to define $\lambda \equiv a_e / a_i$ and $\mu \equiv c / a_i$. Thus, $\mu < 1$. Collecting terms, the right-hand side of $\pi'(0)$ is proportional to:

$$Z(\lambda, \mu, v) = \frac{1}{4} \cdot (T(\lambda) + \mu) \cdot (T(\lambda) - \mu) - v \left[ T(\lambda) + \frac{1}{2} (\mu - 1) \right]$$

$$\cdot \left[ 1 - \frac{1}{2} (T(\lambda) + \mu) \right],$$

where $T(\lambda) \equiv v + (1 - v) \cdot \lambda$.

**Case 1.** $a_e < a_i$ (and therefore $\lambda < 1$).

It is readily verified that the function $Z(\lambda, \cdot, v)$ is monotonically decreasing in $\mu$ for all $(\lambda, v)$. Inequality (6) determines the highest possible value $\mu$ so that both markets will be served, that is:

$$\mu \leq (1 + v) \cdot \lambda - v. \quad (A4)$$

Thus

$$Z(\lambda, \mu, v) > Z(\lambda, (1 + v) \cdot \lambda - v, v)$$

for all $\mu$ satisfying equation (A4). Finally:

$$Z(\lambda, (1 + v) \cdot \lambda - v, v) = \frac{1}{2} \cdot v \cdot (1 - v) \cdot (1 - \lambda)^2 > 0.$$

**Case 2.** $a_e \geq a_i$ (and therefore $\lambda \geq 1$).

Straightforward differentiation shows that $Z(\cdot, \mu, v)$ is monotonically increasing in $\lambda$. Hence, it suffices to evaluate $Z(\cdot, \mu, v)$ at $\lambda = 1$ and to verify that:

$$Z(1, \mu, v) = \frac{1}{4} \cdot (1 - v) \cdot (1 - \mu^2) > 0.$$

That completes the proof of Proposition 2. $QED$.

**Proof of Proposition 3.** For any given $\theta \in U$, the transfer pricing rule $TP(\cdot)$ induces the upstream division to set an external market price $p(\theta)$ that maximizes:

$$\pi_1(p, \theta | TP(\cdot)) = [TP(p) - c(\theta)] \cdot Q_i(TP(p), \theta) + [p - c(\theta)] \cdot Q_e(p, \theta). \quad (A5)$$

The corresponding first-order condition is:

$$TP'(p(\theta)) \cdot \{Q_i(TP(p(\theta)), \theta) + [TP(p(\theta)) - c(\theta)] \cdot Q'_i(TP(p(\theta)), \theta) \}$$

$$+ [p(\theta) - c(\theta)] \cdot Q'_e(p(\theta), \theta) + Q_e(p(\theta), \theta) = 0. \quad (A6)$$
Efficiency requires that \( p(\theta) = \hat{p}_c^*(\theta) \) and \( TP(p(\theta)) = c(\theta) \). Therefore the first order condition in equation (A6) reduces to:

\[
TP'(\hat{p}_c^*(\theta)) \cdot Q_\theta(TP(\hat{p}_c^*(\theta)), \theta) = 0,
\]

which implies \( TP'(\hat{p}_c^*(\theta)) = 0 \). That, however, leads to a contradiction with the requirements that \( TP(p_c^*(\theta)) = c(\theta) \) and \( \nabla c(\theta) \neq 0 \) for all \( \theta \) in the neighborhood \( U \).

**Proof of Proposition 4.** The upstream division chooses \( p \) so as to maximize:

\[
\pi_1 = Q_\epsilon(p, \theta) \cdot [p - c(\theta)] + [TP(p) - c(\theta)] \cdot Q_\epsilon(TP(p), \theta)
\]

subject to:

\[
Q_\epsilon(TP(p), \theta) + Q_\epsilon(p, \theta) \leq K(\theta).
\]

Suppose first that the upstream division sets the market price at \( \hat{p}_c^*(\theta) \equiv P_\epsilon(q_c^*(\theta), \theta) \), where \( q_c^*(\theta) \) denotes the first-best external quantity in the presence of the capacity constraint. By definition:

\[
R'_\epsilon(q_c^*(\theta), \theta) = \hat{p}_c^*(\theta) \cdot \left[1 - \frac{1}{\epsilon^*(\theta)}\right].
\]

Given the transfer pricing rule \( TP(p) = p \cdot [1 - 1/ f(p)] \) and the external market price \( \hat{p}_c^*(\theta) \), the downstream division chooses \( \hat{q}_i \) so that:

\[
R'_i(\hat{q}_i, \theta) = TP(\hat{p}_c^*(\theta)) = \left[1 - \frac{1}{\epsilon^*(\theta)}\right] \cdot \hat{p}_c^*(\theta).
\]

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

It remains to show that the upstream division does not want to choose a price that exceeds \( P_\epsilon(q_c^*(\theta), \theta) \). By Assumption 3, the external monopoly pricing problem is single peaked in \( p \) and, since \( q_c^*(\theta) < q_i^m(\theta) \), any price \( p > P_\epsilon(q_c^*(\theta), \theta) \) lowers the profit of the upstream division externally.

With regard to internal sales, we have:

\[
TP(p_c^*(\theta)) = R'_\epsilon(q_i^*(\theta), \theta) > R'_\epsilon(q_i^m(\theta), \theta) \equiv \hat{p}_i^m(\theta), \tag{A7}
\]

because \( R'_\epsilon(Q_\epsilon(p, \theta), \theta) \equiv p \). The inequality in equation (A7) relies on Assumption 4 requiring that \( q_i^*(\theta) < q_i^m(\theta) \). Therefore \( TP(p_c^*(\theta)) \) exceeds the internal monopoly price. Because \( TP(\cdot) \) is increasing in \( p \) and by Assumption 3, the internal monopoly pricing problem is single peaked, it follows that any price above \( p_c^*(\theta) \) would also lower the upstream division’s profits from internal sales.

**QED.**
Proof of Proposition 5. The claim follows essentially from the arguments given in Proposition 4. It only remains to establish that:

\[ TP(p) = p \cdot \left[ 1 - \frac{1}{\epsilon(p)} \right] \]

is increasing in \( p \), as postulated in Proposition 4. Straightforward algebra yields:

\[ R_e'(q_e(p, \theta), \theta) \equiv p + \frac{b_e(p)}{b_e'(p)} \]

for all \( p \) and all \( \theta \). Therefore \( TP(p) \) is monotone increasing in \( p \) provided \( R_e'(\cdot, \theta) \) is monotone decreasing in \( q_e \). \( \text{QED} \).

APPENDIX B

This appendix derives optimal discounts for the class of linear demand functions as given in equation (5). For tractability reasons, we confine attention to additive (absolute dollar amount) discounts, \( \Delta \). Thus, \( TP = p - \Delta \). For a given discount, \( \Delta \), and a state of the world, \( \theta \), the upstream division will set an external price \( p(\Delta, \theta) \) so as to maximize:

\[ \pi_1(p, \Delta, \theta) = [p - \Delta - c(\theta)] \cdot Q_e(p, \theta) + [p - c(\theta)] \cdot Q_e(p, \theta). \]

Solving for \( p(\Delta, \theta) \) yields a pricing function that is linear in the discount:

\[ p(\Delta, \theta) = \frac{1}{2} \left[ \frac{\alpha_i(\theta)}{\beta_e(\theta)} + \frac{\alpha_e(\theta)}{\beta_e(\theta)} + c(\theta) \right] + \frac{\beta_i(\theta)}{\beta_e(\theta) + \beta_i(\theta)} \cdot \Delta \equiv p(0, \theta) + v(\theta) \cdot \Delta, \]

where \( p'(\Delta, \theta) = v(\theta) = \frac{\beta_i(\theta)}{\beta_e(\theta) + \beta_i(\theta)}. \) Firm-wide profit equals:

\[ \pi(\Delta, \theta) = R_e(Q_e(p(\Delta, \theta) - \Delta, \theta), \theta) - c(\theta) \cdot Q_e(p(\Delta, \theta) - \Delta, \theta) + R_e(Q_e(p(\Delta, \theta), \theta) - c(\theta) \cdot Q_e(p(\Delta, \theta), \theta). \]

In this setting, \( \frac{\partial Q_e(p(\Delta, \theta) - \Delta, \theta)}{\partial \Delta} = Q_e' \cdot [p'(\Delta, \theta) - 1] + Q_e', p'(\Delta, \theta) \equiv 0, \) so that \( \pi'(\Delta, \theta) = \beta_e(\theta) \cdot v(\theta) \cdot [R_e(Q_e(p(\Delta, \theta) - \Delta, \theta), \theta) - R_e(Q_e(p(\Delta, \theta), \theta), \theta) - a_e(\theta) - 2p(\Delta, \theta), \theta), \theta) = -a_e(\theta) + 2p(\Delta, \theta), \theta) \] follows that:

\[ \pi'(\Delta, \theta) = \beta_e(\theta) \cdot v(\theta) \cdot \left[ a_e(\theta) - p(0, \theta) - [1 + v(\theta)] \cdot \Delta \right]. \]

The optimal discount \( \Delta^* \) satisfies \( E_\theta[\pi'(\Delta^*, \theta)] = 0, \) or equivalently,

\[ \Delta^* = \frac{E_\theta[\beta_e(\theta) \cdot v(\theta) \cdot [a_e(\theta) - p(0, \theta)]]}{E_\theta[\beta_e(\theta) \cdot v(\theta) \cdot [1 + v(\theta)]]}. \]  \( \text{(B1)} \)

To develop some intuition for the optimal \( \Delta^* \), consider the special case in which the slope coefficients are independent of the underlying state, that is, \( \beta_i(\theta) = \beta_i, \beta_e(\theta) = \beta_e, \) and hence \( v(\theta) = v = \beta_i/(\beta_i + \beta_e). \) Since the total sales quantity, \( Q_e(\cdot) + Q_e(\cdot), \) is invariant to \( \Delta, \) an optimal discount
has the feature that internal and external marginal revenue be equal in expectation over all $\theta$. Formally, the optimal $\Delta$ is such that $E_{\theta}[R'_e(\cdot)] = E_{\theta}[p(\Delta, \theta) - \Delta]$. Since $p(\Delta, \theta) = p(0, \theta) + v \cdot \Delta$, the optimal discount for a given $\theta$ would be:

$$\alpha_\theta(\theta) - \frac{p(0, \theta)}{1 + v}.$$  \hfill (B2)

Equation (B1) simply generalizes equation (B2), reflecting that the optimal discount must be chosen in expectation over all $\theta$.

The calculation of the optimal discount is readily extended to settings with cost differences between internal and external sales. As in Section 3, suppose that the cost of internal transfers is $c_i(\theta) = c(\theta) - k$. For state-independent slope parameters ($\beta_i, \beta_e$), the seller’s optimal external price becomes:

$$p(\Delta, \theta, k) = p(0, \theta, 0) + v - \frac{v}{2} \cdot k,$$ \hfill (B3)

and the optimal discount is equal to:

$$\Delta^*(k) = \Delta^* + \frac{1 + \frac{v}{2}}{1 + v} \cdot k.$$ \hfill (B4)

According to equation (B4), internal cost savings should not be reflected dollar-by-dollar in $\Delta^*$. Instead, the discount should be chosen so that, at the induced sales quantities, the expected internal marginal revenue exceeds the expected external marginal revenue by the cost difference $k$. By equation (B3), an increase in $k$ has a direct effect on the seller’s price in that $\partial p/\partial k = -v/2$. At the same time, the central office anticipates that raising $\Delta$ induces the seller to increase $p$, that is, $\partial p/\partial \Delta = v$. The net effect is that the optimal discount increases with internal cost savings at a rate between 75% and 100% depending on the relative slope parameters $\beta_i$ and $\beta_e$.

REFERENCES

[References are listed here.]


