Capacity Rights and Full Cost Transfer Pricing

Sunil Dutta

Haas School of Business

University of California, Berkeley

dutta@haas.berkeley.edu

and

Stefan Reichelstein[*]

School of Business, University of Mannheim

and

Graduate School of Business, Stanford University

reichelstein@uni-mannheim.de

August 2019

Abstract: Capacity Rights and Full Cost Transfer Pricing

This paper examines the theoretical properties of full cost transfer prices in multi-divisional firms. In our model, divisional managers are responsible for the initial acquisition of productive capacity and the utilization of that capacity in subsequent periods, once operational uncertainty has been resolved. We examine alternative variants of full cost transfer pricing with the property that the discounted sum of transfer payments is equal to the initial capacity acquisition cost and the present value of all subsequent variable costs of output supplied to a division. Our analysis identifies environments where particular variants of full cost transfer pricing induce efficiency in both the initial investments and the subsequent output levels. Our findings highlight the need for a proper integration of intracompany pricing rules and divisional control rights over capacity assets.

1 Introduction

The transfer of intermediate products and services across divisions of a firm is frequently valued at full cost. Surveys and textbooks consistently report that in contexts where a market-based approach is either infeasible or unreliable, cost-based transfer pricing is the most prevalent method for both internal managerial and tax reporting purposes.[1] Case studies and many managerial accounting textbooks have pointed out that full cost transfer pricing may result in sub-optimal resource allocations. One prominent issue is "double marginalization" which arises when an internal division is charged on the margin for costs that are fixed and sunk. An additional issue, featured in the HBS case study "Polysar Limited" (Simons, 2000), for instance, is that an internal buying division may be induced to reserve too much production capacity because demand for its product is uncertain and the full cost pricing rule charges the division only in proportion to the capacity actually utilized.

This paper identifies conditions pertaining to the firm's production environment that enable full-cost transfer pricing mechanisms to effectively align divisional and overall corporate interests. Our model considers settings in which the resource allocation decisions pertain to upfront capacity investments and the subsequent need to optimize the available capacity upfront capacity investments.[2] In particular, our model considers two divisions that sell a product each in separate markets. Due to technical expertise, the upstream division installs and maintains all productive capacity. It also produces the output sold by the downstream division. For performance evaluation purposes, the upstream division is therefore viewed as an investment center, while the downstream division, having no capital assets, is merely a profit center.

We characterize transfer pricing rules that charge the divisions for the full cost of services provided and induce divisional managers to make both efficient capacity

---

[1]See, for instance, Eccles and White (1988), Ernst & Young (1993), Tang (2002), Feinschreiber and Kent (2012), Datar and Rajan (2014), and Zimmerman (2016).

[2]The perspective in this paper is similar to that underlying the literature on the use of full cost measures for pricing and capacity expansion decisions. See, for example, Banker and Hughes (1994), Balachandran et al.(1997), Goex (2002), Balakrishanan and Sivaramakrishnan (2002), Gramlich and Ray (2016), and Reichelstein and Sahoo (2018). While these studies examine the role of full cost primarily from both a central planning and a divisional coordination perspective, our focus is on divisional incentives and management control.

investment and capacity utilization decisions. Our criterion for incentive compatibility follows the literature on goal congruent performance measures such as Rogerson (1997), Dutta and Reichelstein (2002), Baldenius et al.(2007), and Nezlobin et al.(2015). Accordingly, the divisional performance measures must in any particular time period be congruent with the objective of maximizing firm value. Put differently, regardless of the managers' planning horizons and intertemporal preferences, a goal congruent mechanism must induce (i) the efficient levels of capacity investments upfront, and (ii) the efficient production quantities in subsequent time periods after the resolution of revenue uncertainty in those periods.

Unlike the standard textbook definition of full cost which fails to properly account for the time value of money, our notion of full cost includes imputed interest charges for capital. Specifically, we refer to a transfer pricing rule as a full cost rule if the discounted sum of transfer payments is equal to the present value of cash outflows associated with the capacity assigned to the downstream division and all subsequent output services rendered to that division. In particular, a two-part pricing rule that charges in a lump sum fashion for capacity in each period in addition to variable charges, based on actual production volumes, will be considered a full-cost transfer price.

Our analysis distinguishes two alternative scenarios depending on whether the divisions' products can share the same capacity assets. In the *dedicated capacity* scenario, the products require different productive assets, and hence the capacity cannot be shared across the divisions. Private information at the divisional level then makes it natural to give each division unilateral capacity rights. We identify production and information environments where a suitable variant of full cost transfer pricing induces efficient outcomes. Common to these pricing rules is that the fixed cost charges for capacity must be equal to what earlier literature has referred to as the "user cost of capital"[3]

When both products in question can be produced with same capacity infrastruc-

---

[3]In contrast to our framework here, the derivation of the user cost of capital has been derived in models with overlapping investments in an infinite horizon setting, e.g., Arrow (1964), Carlton and Perloff (2005), Spear (2003), Rogerson (2008, 2011), Rajan and Reichelstein (2009) and Reichelstein and Sahoo (2018).

ture, it suggests itself to allow the divisions to negotiate ex-post over the utilization of the available capacity. In such fungible capacity settings, the cost-based transfer price defines the parties' status quo payoffs in the subsequent negotiations. If the capacity acquisition decision were to be delegated to the upstream division in its role as an investment center, the resulting outcome would generally entail under-investment. The upstream division would then anticipate not earning the full expected return on its investment because gains from the optimized total contribution margin would be shared in the negotiation between the two divisions, when the initial acquisition cost would already be sunk.[4] Under certain conditions, we find that the coordination and hold-up problem associated with the initial capacity choice can be resolved by giving both divisions the unilateral right to reserve capacity, charging the downstream division for capacity reserved with a fixed cost and allowing the divisions to negotiate the actual use of the available capacity in subsequent time periods.

A coordination mechanism that works in a broader class of environments is obtained in the fungible capacity scenario if the downstream division must obtain approval from the investment center manager for any capacity it wants to reserve for its own use. The upstream division then becomes essentially a gatekeeper that will agree to let the downstream division reserve capacity for itself in exchange for a stream of lump-sum payments determined through initial negotiation. The upstream division will thereafter have an incentive to invest in additional capacity on its own up to the efficient level. The resulting mechanism can be viewed as a hybrid between cost-based and negotiated transfer pricing rules such that the downstream division is charged the full cost of the total capacity acquired and total output produced.

Taken together, our results yield predictions regarding the environments in which particular variants of cost-based transfer pricing are likely to perform well. For both the dedicated and the fungible capacity scenario, full cost transfer pricing rules tend to work well if the divisional revenue opportunities are expected to be stable over time.

---

[4]Even though investments are verifiable in our model, the hold-up problem that arises when only the upstream division makes capacity investments is essentially the same as in earlier incomplete contracting literature. One branch of that literature has explored how transfer pricing can alleviate hold-up problems when investments are "soft" (unverifiable); see, for example, Baldenius et al.(1999), Edlin and Reichelstein (1995), Sahay (2000), Baldenius (2008), and Pfeiffer et al. (2009).

Furthermore, cost based transfer pricing tends to work better in environments where the capacity related costs are large relative to the variable periodic operating costs. If capacity is fungible across multiple divisions, our results predict that the efficiency of cost based transfer pricing is enhanced if the divisions do not have symmetric capacity rights, but instead one division has the effective authority of a gatekeeper.

Numerous theoretical and empirical studies have examined alternative cost measures that should apply to an internal buyer for receiving intracompany services.[5] Among these studies, Dutta and Reichelstein (2010) is structurally closest to the analysis in this paper. Their findings identify conditions under which full cost transfer pricing will lead to efficient outcomes. However, while capacity investments are costly, there are no subsequent variable costs associated with producing output in their model. Unlike our analysis in this paper where it may be efficient not to exhaust the available capacity in bad states of the world, capacity is always fully utilized in Dutta and Reichelstein (2010).

Our analysis also builds on the work of Reichelstein and Rohlfing-Bastian (2015), which examines the relevant full cost measure for capacity investments from a planning and product pricing perspective. Baldenius, Nezlobin and Vaysman (2016) is another precursor to the present paper insofar as they study managerial performance evaluation in a setting where capacity may remain idle in unfavorable states of the world. Their analysis is confined to a single division firm, and thus coordination and internal pricing issues do not arise. In the context of the dedicated capacity investments, our analysis can be viewed as providing further possibility results for performance measures that satisfy the requirement of sequential goal congruence.

We view the analysis in this paper as being both descriptive and prescriptive in nature. The descriptive aspect is that our results provide support for the ubiquitous practice of full cost transfer pricing. At the same time, our results show that the specifics of the full cost transfer pricing rule ought to be tailored to the characteristics of the intrafirm resource allocation problem. Aside from the need for two-part tariffs

---

[5]A partial list of references includes Eccles and White (1988), Vaysman (1996), Baldenius et al. (1999), Sahay (2002), Pfeiffer et al.(2009), Arya and Mittendorf (2011), Baldenius (2008), and Bouwens and Steens (2016).

and the inclusion of imputed interest charges in the measurement of capacity costs, our results point to the importance of assigning capacity rights and the attendant divisional charges for those capacity rights so as to effectively coordinate the initial investment decisions.

The remainder of the paper proceeds as follows. The basic model is described in Section 2. Section 3 examines a setting in which the divisions' products require different production facilities and therefore capacity is dedicated. Section 4 considers the alternative setting in which capacity is fungible and can be traded across divisions. Propositions 2 and 3 demonstrate the importance of allowing the downstream division to secure capacity rights for itself initially, even if the entire available capacity can be reallocated through negotiations in subsequent periods. While our initial results apply to divisional environments that are stationary in expectation, Section 5 identifies settings in which cost based transfer pricing rules can achieve goal congruence in non-stationary environments. We conclude in Section 6.

2   Model Description

Consider a vertically integrated firm comprised of two divisions and a central office. Both divisions sell a marketable product (possibly a service) in separate and unrelated markets. In order for either division to deliver its product in subsequent periods, the firm needs to make upfront capacity investments. Because of technical expertise, only the upstream division (Division 1) is in a position to install and maintain the productive capacity for both divisions. Division 1 also carries out the production for both divisions, and therefore incurs all periodic production costs.[6]

Our analysis considers two distinct scenarios of *dedicated* and *fungible* capacity. In the former scenario, the two products are sufficiently different so as to require separate production facilities. With fungible capacity, in contrast, both products can utilize the same capacity infrastructure. The upfront cash expenditure for one unit of capacity for Division i is v_i in the dedicated capacity setting. If Division i acquires

---

[6]It is readily verified that our findings would be unchanged if the upstream division were to transfer an intermediate product which the downstream division can then convert into a final product after incurring an incremental cost.

k_i units of capacity, it has the option to produce up to k_i units of output in each of the next T periods.[7] In case of fungible capacity, the cost of acquiring one unit of capacity is v, which allows either division to produce one unit of output in each of the next T periods.

The actual production levels for Division i in period t are denoted by q_{it}. We assume that sales in each period are equal to the amount of production in that period; i.e., the divisions do not carry any inventory. Aside from requisite capacity resources, the delivery of one unit of output for Division i requires a unit variable cost of w_{it} in period t. These unit variable costs are anticipated upfront by the divisional managers with certainty, though they may become known and verifiable to the firm's accounting system only when incurred in a particular period. The divisional contribution margins are given by

$$CM_{it}(q_{it}, \epsilon_{it}) = x_{it} \cdot R_i(q_{it}, \epsilon_{it}) - w_{it} \cdot q_{it}.$$

The first term above, $x_{it} \cdot R_i(q_{it}, \epsilon_{it})$, denotes Division i's revenues in period t with $x_{it} > 0$ representing intertemporal parameters that allow for the possibility of declining, or possibly growing, revenues over time. To avoid laborious checking of boundary cases, we assume throughout our analysis that the marginal revenue at zero always exceeds the unit variable cost of production, i.e.,

$$x_{it} \cdot R_i'(0, \epsilon_{it}) - w_{it} > 0$$

for all realizations of ϵ_{it}.

The periodic revenues vary with the production quantities q_{it}. they are also subject to one-dimensional transitory shocks ϵ_{it}. These random shocks are realized at the beginning of period t before the divisions choose their output levels for the current period, and prior to any capacity trades in the fungible capacity setting. We assume that the random shocks ϵ_{it} are distributed according to density functions $f_i(\cdot)$ with support on the interval $[\underline{\epsilon}_i, \bar{\epsilon}_i]$. The random variables $\{\epsilon_{it}\}$ are also assumed to be independently distributed across time; i.e., $Cov(\epsilon_{it}, \epsilon_{i\tau}) = 0$ for each $t \neq \tau$, though

---

[7]Studies on capacity investments generally allow for capacity to be subject to economic depreciation possibly in the form of physical delay (Rogerson, 2008, and Nezlobin, Rajan and Reichelstein, 2012). In Section 5, we consider the possibility that revenues obtained from the initial capacity may be subject to time-driven degradation.

they may be correlated across the two divisions; i.e., $Cov(\epsilon_{1t}, \epsilon_{2t})$ may be non-zero in any given period t.

Though the exact shapes of divisional revenue functions, $R_i(q_{it}, \epsilon_{it})$, are known to the divisional managers, the firm's central office does not have this information.[8] For this reason, the capacity investment and utilization decisions are delegated to the divisional managers. The divisional revenue functions are assumed to be increasing and concave in q_{it} for each i and each t. At the same time, the marginal revenue functions:

$$R_i'(q, \epsilon_{it}) \equiv \frac{\partial R_i(q, \epsilon_{it})}{\partial q}$$

are assumed to be increasing in ϵ_{it}.

In any given period, the actual production quantity for a division may differ from its initial capacity rights for two reasons. First, for an unfavorable realization of the revenue shock ϵ_{it}, a division may decide not to exhaust the entire available capacity because otherwise marginal revenues would not cover the incremental cost w_{it}. Second, in the case of fungible capacity, a division may want to yield some of its capacity rights to the other division if that division has a higher contribution margin.

We assume that each divisional manager is evaluated by a performance measure Π_{it} in each of the T time periods. The downstream division, which has only operational responsibilities for procuring and selling output, is treated as a profit center whose performance measure is measured by its divisional profit. In contrast, the upstream division, which also has control over capacity assets, is viewed as an investment center with *residual income* as its performance measure.[9] The remaining design variables of the internal managerial accounting system then consist of divisional capacity rights, depreciation schedules, and the transfer pricing rule. The default scenario in our analysis is that the capacity investment decisions are decentralized in the sense that each manager can unilaterally decide the amount of capacity that will be avail-

---

[8]In Section 3 below, our analysis allows for the divisional revenue functions to be private information of that division's management. For the fungible capacity setting in Section 4, the two managers are assumed to have the same information about divisional revenues. In either setting, the managers will have superior information about expected divisional revenues relative to the firm's central office.

[9]Earlier literature, including Reichelstein (1997), Dutta and Reichelstein (2002), and Baldenius et al.(2007), has argued that among a particular class of accounting based metrics only residual income can satisfy the requirement of goal congruence.

able to his/ her division in the future. For certain environments, it will turn out to be advantageous to consider a more asymmetric assignment of capacity rights.

Figure 1 illustrates the structure of the multi-divisional firm and its two constituent responsibility centers.
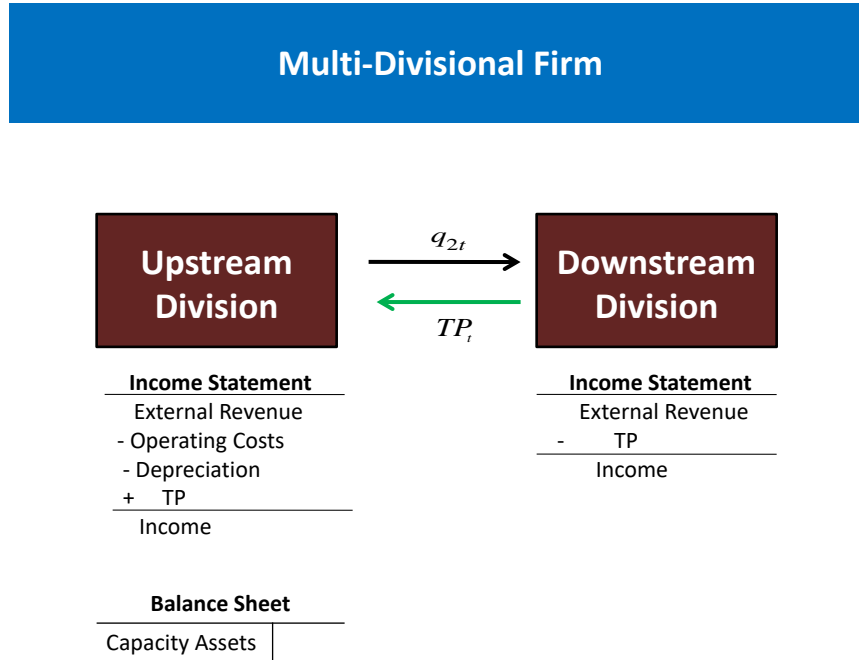
Multi-Divisional Firm

Upstream Division

q_{2t}

Downstream Division

TP_t

Income Statement
External Revenue
- Operating Costs
- Depreciation
+ TP
Income

Income Statement
External Revenue
- TP
Income

Balance Sheet
Capacity Assets

Figure 1: **Divisional Structure of the Firm**

The downstream division's performance measure (i.e., its operating income) in period t is given by

$$\Pi_{2t} = Inc_{2t} = x_{2t} \cdot R_2(q_{2t}, \epsilon_{2t}) - TP_t(k_2, q_{2t}), \tag{1}$$

where $TP_t(k_2, q_{2t})$ denotes the transfer payment to the upstream division in period t for securing k_2 units of capacity and obtaining q_{2t} units of output. Equation (1) reflects that Division 1 incurs the variable cost of producing output for Division 2. The transfer pricing rule $TP_t(k_2, q_{2t})$ specifies how Division 2 is charged for these variable costs.

The residual income measure for the upstream division is given by

$$\Pi_{1t} = Inc_{1t} - r \cdot BV_{t-1}, \qquad (2)$$

where BV_t denotes book value of capacity assets at the end of period t and r denotes the firm's cost of capital. The corresponding discount factor is denoted by $\gamma \equiv (1 + r)^{-1}$. The residual income measure in (2) depends on two accruals: the transfer price received from the downstream division and the depreciation charges corresponding to the initial capacity investments. Specifically,

$$Inc_{1t} = x_{1t} \cdot R_1(q_{1t}, \epsilon_{1t}) - w_{1t} \cdot q_{1t} - w_{2t} \cdot q_{2t} - D_t + TP_t(k_2, q_{2t}),$$

where D_t is the total depreciation expense in period t. When there are two distinct capacity assets, d_{it} denotes the depreciation charge in period t per dollar of initial capacity investment undertaken for Division i. Thus,

$$D_t = d_{1t} \cdot v_1 \cdot k_1 + d_{2t} \cdot v_2 \cdot k_2.$$

The depreciation schedules satisfy the usual tidiness requirement that $\sum_{\tau=1}^{T} d_{i\tau} = 1$; i.e, the depreciation charges sum up to an asset's historical acquisition cost over its useful life. Book values at time t then become: $BV_t = BV_{t-1} - D_t$, with $BV_0 = v_1 \cdot k_1 + v_2 \cdot k_2$ and $BV_T = 0$. In the fungible capacity scenario, $v_1 = v_2$, though the central office may still apply different depreciation schedules for the expenditures associated with the reserved capacity by each division.

Under the residual income measure, the overall *capital charge* imposed on the upstream division is the sum of depreciation charges plus imputed interest charges. Given the depreciation schedules $\{d_{it}\}_{t=1}^{T}$, the overall capital charge becomes:

$$D_t + r \cdot BV_{t-1} = z_{1t} \cdot v_1 \cdot k_1 + z_{2t} \cdot v_2 \cdot k_2, \qquad (3)$$

where $z_{it} \equiv d_{it} + r \cdot (1 - \sum_{\tau=1}^{t-1} d_{i\tau})$ for $t \geq 2$. It is well known from the general properties of the residual income metric that regardless of the depreciation schedule, the present value of the z_{it} is equal to one; that is, $\sum_{t=1}^{T} z_{it} \cdot \gamma^t = 1$ (Hotelling, 1925).

9

Our model is in the tradition of the earlier literature which does not model explicit agency problems and their associated costs, but instead focuses on identifying goal congruent performance measures. A performance measure is said to be goal congruent if in equilibrium it induces decisions that maximize the net present value of firm-wide future cash flows. Consistent with this earlier literature, we impose the criterion of *strong goal congruence*, which requires performance measures to be robust in the sense that they induce the managers to make efficient production and investment decisions even if the managers are less patient than the firm's owners (possibly because of higher discount rates or shorter planning horizons) and the compensation rules vary over time. Formally, suppose the manager of Division i attaches non-negative weights $\{u_{it}\}_{i=1}^{T}$ to her performance measure in different time periods. The weights $u_i = (u_{i1}, ..., u_{iT})$ reflect both the manager's discount factor as well as the bonus coefficients attached to the periodic performance measures. Manager i's objective function can thus be written as $\sum_{t=1}^{T} u_{it} \cdot E[\Pi_{it}]$. A combination of transfer pricing and depreciation rules is said to satisfy the requirement of strong goal congruence if the divisional managers are incentivized to make efficient quantity and capacity decisions for *any* configuration of the coefficients $u_{it} \geq 0$.

The goal congruence approach essentially separates the problems of choosing accounting rules and performance measures from that of choosing compensation rules for a particular agency problem. The perspective is that the accounting rules and performance measures are chosen before the specifics of the managerial contracting problem become known. Yet, the chosen rules and performance measures are robust enough to generate desirable incentives for managers with different time preferences.[10] A central implication of strong goal congruence requirement is that the managers must not face any intertemporal tradeoffs in making their decisions; i.e., value maximizing decisions may not lower the performance measure in any period. That is, each manager must have incentives to make efficient production and capacity decisions *even* if

---

[10]Arya et al (2009) study design of robust mechanisms in an auction setting with private values. Similar to our approach, they assume that the mechanism must be designed before agent characteristics are fully known. In contrast to our goal congruence framework, however, they adopt a Bayesian perspective in which the mechanism is designed by taking expectations over possible agent characteristics that will subsequently emerge.

that manager were solely focused on maximizing her performance measure Π_{it} in any particular period t.[11]

Managerial accounting textbooks usually define full cost as the sum of fixed and variable product costs, where the fixed costs include accruals like depreciation. Since we seek to identify full cost transfer pricing rules that can induce divisional managers to make efficient investment decisions, the measure of full cost must include imputed interest charges on the capital employed so as to reflect the time value of money. Accordingly, we classify a transfer pricing rule as full cost if the sum of the discounted cash outflows is equal to the sum of the discounted transfer payments.

Definition *A transfer pricing rule is said to be a full cost pricing rule if, in equilibrium:*

$$\sum_{t=1}^{T} TP_t(k_2, q_{2t}) \cdot \gamma^t = v_2 \cdot k_2 + \sum_{t=1}^{T} w_{2t} \cdot q_{2t} \cdot \gamma^t.$$

The qualifier "in equilibrium" in the preceding definition refers to the notion that the transfer payments must be balanced only for the investment and operating decisions that the divisional managers are supposed to make in equilibrium. The specific notion of equilibrium we refer to will vary with the particular environment considered.

Our analysis first focuses on stationary environments in which the economic fundamentals are, in expectation, identical across periods. Formally, an environment is said to be *stationary* if $x_{it} = 1$, $w_{it} = w_i$ and the $\{\epsilon_{it}\}$ are i.i.d. for each i. We assume that these stationarity conditions hold for the dedicated and fungible capacity settings analyzed in Sections 3 and 4, respectively. For these stationary environments, we drop subscript t from $CM_{it}(\cdot)$. Section 5 examines how our results generalize to non-stationary environments.

---

[11]The concept of goal congruence dates back to the early work of Solomons (1964). Dutta (2008) identifies settings in which the accrual accounting rules that emerge as goal congruent are also part of optimal contracting arrangements in agency problems.

11

3 Dedicated Capacity

We first investigate a setting in which the divisional products require different capacity infrastructures. The firm's capacity and production planning problem then separates across the two divisions. In some contexts, it may actually be possible for the two products to share the same capacity infrastructure, but long lead times for the setup required to switch to a different product may effectively preclude the firm from redeploying the available capacity after the random shocks $(\epsilon_{1t}, \epsilon_{2t})$ have been realized in period t.[12]

The divisional managers are assumed to have private information regarding their future revenue opportunities, and each division is therefore given the unilateral right to procure capacity for its own use. The analysis in this section focuses on identifying the depreciation schedules and transfer pricing rules that provide incentives for the divisional managers to choose efficient levels of capacity upfront and make optimal production decisions in subsequent periods. The following time line illustrates the sequence of events at the initial investment date and in a generic period t.
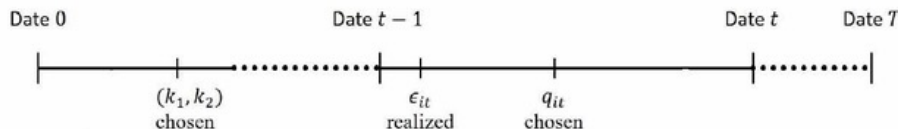

Figure 2: Sequence of Events in the Dedicated Capacity Scenario

If the central office had full information regarding future revenues, it would choose the optimal capacities (k_1, k_2) so as to maximize the net present value of the firm's expected future cash flows $NPV(k_1, k_2) = NPV_1(k_1) + NPV_2(k_2)$, where

$$NPV_i(k_i) = \sum_{t=1}^{T} E_{\epsilon_i}\left[CM_i(k_i|w_i, \tilde{\epsilon}_{it})\right] \cdot \gamma^t - v_i \cdot k_i, \tag{4}$$

and $CM_i(\cdot)$ denotes the maximized value of the expected future contribution margin in period t. That is,

---

[12]The case study on the Timken company illustrates this issue in connection with tapered roller bearings produced for automotive and industrial applications (Bastian and Reichelstein, 2004. In that industrial environment, the random shocks $(\tilde{\epsilon}_{1t}, \tilde{\epsilon}_{2t})$ tended to be positively correlated.

$$CM_i(k_i|w_i, \epsilon_{it}) \equiv R_i(q_i^o(k_i, \cdot), \epsilon_{it}) - w_i \cdot q_i^o(k_i, \cdot),$$

with

$$q_i^o(k_i, \cdot) = \underset{q_i \leq k_i}{\operatorname{argmax}}\{R_i(q_i, \epsilon_{it}) - w_i \cdot q_i\}.$$

The notation $q_i^o(k_i, \cdot)$ above is short-hand for the sequentially optimal quantity $q_{it}^o(k_i, w_i, \epsilon_{it})$ that maximizes the divisional contribution margin in period t, given the initial capacity choice k_i, the unit variable cost w_i, and the realization of the current revenue shock ϵ_{it}. The result below characterizes the efficient capacity levels, k_i^o, for the dedicated capacity setting.

Lemma 1 *Suppose capacity is dedicated and the divisional environments are station-ary. The optimal capacity level k_i^o is given by the unique solution to the equation:*

$$E_{\epsilon_i}\left[R_i'(q_i^o(k_i^o, w_i, \tilde{\epsilon}_{it}), \tilde{\epsilon}_{it})\right] = c_i + w_i, \qquad (5)$$

where

$$c_i = \frac{v_i}{\sum_{t=1}^{T} \gamma^t}. \qquad (6)$$

Proof: All proofs are in Appendix A.[13]

Earlier literature, including Rogerson (2008) and Rajan and Reichelstein (2009), refers to c_i as the user cost of capital or the unit cost of capacity. The user cost of capital c_i is obtained by annuitizing the unit cost of capacity v_i (i.e., dividing v_i by $\sum_{t=1}^{T} \gamma^t$, which is the present value of \$1 annuity over T periods). It is readily verified that c_i is the price that a hypothetical supplier would charge for renting out capacity for one period of time if the rental business breaks even.

---

[13]Lemma 1 implicitly assumes that $k_i^o > 0$. We note that $k_i^o = 0$ if the left-hand side of (5) is less than $c_i + w_i$ at $k_i = 0$.

Lemma 1 says that the optimal capacity level k_i^o is such that the expected marginal revenue at the *sequentially optimal* production levels, $q_i^o(k_i^o, \cdot)$ is equal to the sum of the unit cost of capacity c and the variable cost w_i. We shall subsequently refer to the sum, $c_i + w_i$, as the *full cost* per unit of output. This value will generally exceed the traditional measure of full cost in managerial accounting because the latter measure does not include the imputed interest charges for capital. For instance, if the depreciation charges are uniform, the traditional measure of full cost in each period is given by $\frac{v_i}{T} + w_i$, which is less than $\frac{v_i}{\sum_{t=1}^{T} \gamma^t} + w_i \equiv c_i + w_i$ (Reichelstein and Rohlfing-Bastian, 2015).

In the context of our model, one common representation of full cost transfer pricing is that the downstream division is charged in the following manner for intra-company transfers:

1. Division 2 has the unilateral right to reserve capacity at the initial date.

2. Division 2 can choose the quantity, q_{2t}, to be transferred in each period subject to the initial capacity limit.

3. In period t, Division 2 is charged the full cost of output delivered, that is:
$$TP_t(k_2, q_{2t}) = (w_2 + c_2) \cdot q_{2t}.$$

This variant of full cost transfer pricing is essentially the one featured in the Harvard case study "Polysar" (Simons, 2000). The downstream division is charged for capacity only to the extent that it actually utilizes that capacity. A key takeaway from the Polysar case study is that the buying division will tend to reserve too much capacity upfront in the face of uncertain demand for its product. Such a strategy preserves the division's option to meet market demand if it turns out to be strong, while it incurs no penalty for idling capacity if market conditions turn out to be unfavorable. Adding to this bias, the downstream division faces the classic double-marginalization problem.

With dedicated capacity, the requirement of strong goal congruence can be formalized as follows. A goal congruent mechanism must induce efficient production $q_i^o(k_i^o, \cdot)$ in each period given the optimal capacity level k_i^o; that is,

$$q_i^o(k_i^o, w_i, \epsilon_{it}) \in \underset{q_{it} \leq k_i^o}{\operatorname{argmax}}\{\Pi_{it}(q_{it}, k_i^o, w_i, \epsilon_{it})\} \tag{7}$$

for all ϵ_{it} in each period $1 \leq t \leq T$. Furthermore, strong goal congruence requires that the divisional managers have incentives to choose the optimal capacity levels at the outset and that these incentives hold not only over the entire planning horizon, but also on a period-by-period basis. This implies that k_i^o must be a maximizer of the expected performance measure in each period t; i.e.,

$$k_i^o = \underset{k_i}{\operatorname{argmax}}\{E_{\epsilon_i}[\Pi_{it}^*(k_i, w_i, \tilde{\epsilon}_{it})]\}, \tag{8}$$

for each t, where $\Pi_{it}^*(k_i, w_i, \epsilon_{it})$ denotes the maximized value of $\Pi_{it}(q_{it}, k_i, w_i, \epsilon_{it})$ with regard to q_{it}. The requirement that k_i^o must be Division i's preferred choice in each period follows directly from the initial definition of strong goal congruence if one holds the weight $u_{it} > 0$ but allows for all other weights $u_{i\tau}$, $\tau \neq t$, to go to zero.

To characterize the transfer pricing and depreciation rules that meet these requirements of strong goal congruence, we consider (i) transfer pricing rules $TP_t(k_2, q_{2t})$ that are differentiable in both variables, and (ii) proper depreciation rules that satisfy $\sum_{t=1}^T d_{it} = 1$.

Proposition 1 *Suppose capacity is dedicated and the environment is stationary. Strong goal congruence will be achieved if and only if*

$$TP_t(k_2, q_{2t}) = c_2 \cdot k_2 + w_2 \cdot q_{2t} \tag{9}$$

and Division 1's capacity assets are depreciated according to the annuity rule.

We note that the transfer pricing rule in (9) is not only unique in achieving strong goal congruence, but also meets the above definition of full cost pricing. By charging the downstream division separately for (i) the amount of capacity reserved and (ii) the variable cost of output procured actually in each period, this division will no

15

longer have a tendency to reserve too much capacity upfront nor will it be subject to any double marginalization. We note that the transfer prices in (9) can be rewritten as:

$$TP_t(k_2, q_{2t}) = (w_2 + c_2) \cdot q_{2t} - c_2 \cdot (k_2 - q_{2t}),$$

where the last term $c_2 \cdot (k_2 - q_{2t})$ can be interpreted as a fixed overhead volume variance (Datar and Rajan, 2014). The allocation of such variances among different departments has been discussed extensively in the cost accounting literature; see, for example, Kaplan (2006) and Martinez-Jerez (2007). In the context of our multi-divisional firm model, the conclusion is that a simplistic full cost transfer pricing rule of the form $TP_t(k_2, q_{2t}) = (w_2 + c_2) \cdot q_{2t}$ can generate efficient incentives, provided any fixed overhead volume variances are charged to the buying division that is given authority to choose both the initial capacity choice and the subsequent production levels.

The uniqueness of the additively separable full-cost transfer pricing rule in (9) reflects that our criterion of goal congruence requires the divisional managers to have the requisite incentives for any revenue function satisfying the regularity conditions specified in Section 2. That domain of possible revenue functions including the special case of quadratic functions. The proof in Appendix A shows that in case the underlying revenue functions do assume a quadratic form, the sequential goal congruence requirement in (7) implies that

$$\frac{\partial}{\partial q} TP_t(k_2, q_{2t}) \equiv w_2$$

for all k_2. Furthermore, the initial incentive compatibility condition in (8) requires that

$$\frac{\partial}{\partial k} TP_t(k_2, q_{2t}) \equiv c_2.$$

These two requirements on the partial derivatives imply that, up to an indeterminate constant, the transfer pricing rule must be equal to that in (9).[14] If Division 1's

---

[14]Our approach here is the same as in earlier studies on resource allocation mechanisms where the uniqueness of a particular mechanism emerged because certain performance requirements had to be met for a broad class of possible environments; see, for instance, Mount and Reiter (1974) and Marschak and Reichelstein (1998).

investment is depreciated according to the annuity rule, the effective charge for capacity and actual production in period t is $c_1 \cdot k_1 + w_1 \cdot q_{1t}$, ensuring goal congruence. Since there is a 1-1 relationship between depreciation schedules and intertemporal cost allocations under the residual income performance measure (Rogerson, 1997), the annuity rule emerges as the unique goal congruent solution.[15] The depreciation schedule for Division 2's operating assets remains indeterminate since the choice of depreciation schedule only affects the upstream division. One advantage of applying the annuity rule for both divisions would be that the performance measures for the two divisions become effectively independent of each other since Division 2's transactions is merely a pass-through from Division 1's perspective.

In contrast to our result in Proposition 1 which shows that it is necessary to use a two-part full cost transfer price, Dutta and Reichelstein (2010) argue that with dedicated capacity, a simplistic full cost transfer pricing rule of the form $TP_t(k_2, q_{2t}) = c_2 \cdot q_{2t}$ can result in efficient capacity investments. In their setting, however, the issue of capacity under-utilization never arises, irrespective of the initial capacity investment, because there are no variable costs of production (i.e., $w_i = 0$). Even though divisional managers may face uncertainty regarding the value of capacity, they will ex-post always prefer to exhaust the available capacity.

4   Fungible Capacity

In contrast to the scenario considered thus far, where the products or services provided by the two divisions required different production assets, we now consider the plausible alternative of fungible capacity. Accordingly, the production processes of the two divisions have enough commonalities and the demand shocks ϵ_t are realized sufficiently early in each period, so that the initial capacity choices can be reallocated across the

---

[15]The literature on value-based management frequently refers to the annuity rule as "sinking fund" depreciation (Young and O'Byrne, 2000). Though the annuity depreciation method is not generally admissible for financial reporting purposes, the new lease accounting standard, ASC 842, requires its use for a class of leased assets. Specifically, if a lease is classified as an operating lease, the leased asset must be effectively depreciated according to the annuity rule. The annuity rule is the only amortization rule with the property that total lease related expenses (i.e., the sum of depreciation and interest expenses) are identical across periods.

two divisions. The following analysis assumes that the two divisional managers have symmetric information regarding the expected cost and revenue structure of both divisions. In particular, they are assumed to negotiate the capacity rights under conditions of symmetric information.[16] This information is, of course, not known to the firm's central office that sets the transfer pricing and depreciation rules. The following time line illustrates the sequence of events at the initial investment date and in a generic period t.
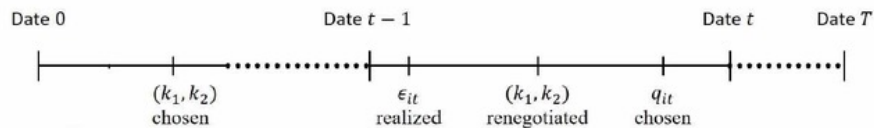

Figure 3: Sequence of Events in the Fungible Capacity Scenario

The analysis in this section again focuses on stationary environments. With fungible capacity, the optimal investment from a firm-wide perspective is the one that maximizes total expected future cash flows:

$$NPV(k) = \sum_{t=1}^{T} E_{\epsilon_t} \left[CM(k|w, \tilde{\epsilon}_t) \right] \cdot \gamma^t - v \cdot k, \tag{10}$$

where $w \equiv (w_1, w_2)$, $\epsilon_t \equiv (\epsilon_{1t}, \epsilon_{2t})$, and $CM(\cdot)$ denotes the maximized value of the aggregate contribution margin in period t. That is,

$$CM(k|w, \epsilon_t) \equiv \sum_{i=1}^{2} [R_i(q_i^*(k, \cdot), \epsilon_{it}) - w_i \cdot q_i^*(k, \cdot)],$$

where

$$(q_1^*(k, \cdot), q_2^*(k, \cdot)) = \underset{q_1 + q_2 \leq k}{\operatorname{argmax}} \{ \sum_{i=1}^{2} [R_i(q_i, \epsilon_{it})) - w_i \cdot q_i] \}.$$

As before, the notation $q_i^*(k, \cdot)$ is short-hand for $q_i^*(k, w, \epsilon_t)$.

Provided the optimal quantities $q_i^*(k, \cdot)$ are both positive, the first-order condition:

$$R_1'(q_1^*(k, \cdot), \epsilon_{1t}) - w_1 = R_2'(q_2^*(k, \cdot), \epsilon_{2t}) - w_2 \tag{11}$$

---

[16]In the dedicated setting above, it did not matter whether a divisional manager also had information pertaining to the other division's revenue opportunities.

18

must hold. Allowing for corner solutions, we define the *shadow price* of capacity in period t, given the available capacity k, as follows:

$$S(k|w, \epsilon_t) \equiv \max\{R_1'(q_1^*(k, \cdot), \epsilon_{1t}) - w_1, R_2'(q_2^*(k, \cdot), \epsilon_{2t}) - w_2\}. \tag{12}$$

The shadow price of capacity identifies the maximal change in periodic contribution margin that the firm can obtain from an extra unit of capacity.[17] We note that $S(\cdot)$ is increasing in ϵ_{it}, but decreasing in w_i and k.

Lemma 2 *Suppose capacity is fungible and the divisional environments are stationary. The optimal capacity level, k^*, is given by the unique solution to the equation:*[18]

$$E_\epsilon\left[S(k^*|w, \tilde{\epsilon}_t)\right] = c, \tag{13}$$

where

$$c = \frac{v}{\sum_{t=1}^T \gamma^t}. \tag{14}$$

Proposition 1 has shown that unilateral capacity rights combined with a two-part full cost transfer pricing rule can attain strong goal congruence when capacity is dedicated. An obvious question is whether there are circumstances under which a symmetric allocation of capacity decision rights can also solve the coordination problem when capacity is fungible across the two business segments. If the two divisions are given unilateral capacity rights at the outset, an ex-post efficient allocation of the available capacity will require the flexibility of allowing the divisions to renegotiate in each period as current revenue information is revealed.

Suppose that the central office has imposed the two-part full cost transfer pricing rule $TP_t(k_2, q_{2t}) = c \cdot k_2 + w_2 \cdot q_{2t}$, and capacity assets are depreciated according to the annuity rule. Suppose further that at the outset Division i has procured k_i units of capacity for its own use. Given symmetric information about each other's revenues and costs, the divisional managers can increase the firm-wide contribution margin by

---

[17]The assumption that $R_i'(0, \epsilon_{it}) > w_i$ for all ϵ_{it} ensures that the shadow price of capacity is always positive.

[18]As in Lemma 1, the statement of Lemma 2 makes the implicit assumption that $k^* \neq 0$.

reallocating the available capacity $k_1 + k_2$ at the beginning of each period after the relevant shock ϵ_t is realized. The resulting "trading surplus" of

$$TSP \equiv CM(k|w, \epsilon_t) - \sum_{i=1}^{2} CM_i(k_i|w_i, \epsilon_{it}) \tag{15}$$

can then be shared by the two divisions. Let $\delta \in [0, 1]$ denote the fraction of the total surplus that accrues to Division 1. Thus, the parameter δ measures the relative bargaining power of Division 1, with the case of $\delta = \frac{1}{2}$ corresponding to the familiar Nash bargaining outcome. The negotiated adjustment in the transfer payment, ΔTP_t, that implements the above sharing rule is given by

$$R_1(q_1^*(k, \cdot), \epsilon_{1t}) - w_1 \cdot q_1^*(k, \cdot) + \Delta TP_t = CM_1(k_1|w_1, \epsilon_{1t}) + \delta \cdot TSP,$$

where we recall that $q_1^*(k, \cdot)$ and $q_2^*(k, \cdot)$ are the divisional production choices that maximize the aggregate contribution margin. At the same time, Division 2 obtains:

$$R_2(q_2^*(k, \cdot), \epsilon_{2t}) - w_2 \cdot q_2^*(k, \cdot) - \Delta TP_t = CM_2(k_2|w, \epsilon_{2t}) + (1 - \delta) \cdot TSP.$$

These payoffs ignore any transfer payment that Division 2 makes in each period to account for its initial capacity choice, since this payment is viewed as sunk at the renegotiation stage. The total transfer payment made by Division 2 in return for the ex-post efficient quantity $q_2^*(k, \cdot)$ is then given $c \cdot k_2 + w_2 \cdot q_2^*(k, \cdot) + \Delta TP_t$.

After substituting for TSP from (15), the effective contribution margin to Division i can be expressed as follows:

$$CM_1^*(k_1|k_2, \epsilon_t) = (1 - \delta) \cdot CM_1(k_1|w_1, \epsilon_{1t}) + \delta \cdot [CM(k|w, \epsilon_t) - CM_2(k_2|w_2, \epsilon_{2t})]$$

and

$$CM_2^*(k_1|k_2, \epsilon_t) = \delta \cdot CM_2(k_2|w_2, \epsilon_{2t}) + (1 - \delta) \cdot [CM(k|w, \epsilon_t) - CM_1(k_1|w_1, \epsilon_{1t})].$$

We note that the expected value of the effective contribution margin, $E_\epsilon [CM_i^*(k_i|k_j, \epsilon_t)]$, is identical across periods for stationary environments. Combined with the annuity

depreciation rule for capacity assets, this implies that Division i will choose k_i to maximize:

$$E_\epsilon \left[CM_i^*(k_i | k_j, \tilde{\epsilon}_t) \right] - c \cdot k_i \tag{16}$$

taking division j's capacity request k_j as given. With fungible capacity, we interpret strong goal congruence to mean that, given the optimal k^*, there exist unilateral capacity choices (k_1^*, k_2^*) that constitute a Nash equilibrium at the initial capacity choice stage, and $k_1^* + k_2^* = k^*$.

The following possibility result concerning goal congruent mechanism relies on two structural assumptions. First, we assume that the revenue shocks ϵ_{it} exhibit limited volatility in the sense that capacity is fully utilized on the *equilibrium path* in the dedicated capacity scenario. Following Reichelstein and Rohlfing-Bastian (2015), the *limited volatility condition* is said to hold if $q_i^o(k_i^o, \cdot) = k_i^o$ for all realizations of ϵ_{it} where $q_i^o(\cdot, \cdot)$ and k_i^o denote the efficient quantity and capacity levels in the dedicated capacity scenario. Clearly, this condition will be met if and only if

$$R_i'(k_i^o, \epsilon_{it}) - w_i \geq 0$$

for *all* realizations of ϵ_{it}. Intuitively, the available capacity will always be exhausted in environments with relatively low volatility in terms of the range and impact of the $\tilde{\epsilon}_{it}$, or alternatively, if the unit variable cost, w_i, is small relative to the full cost, $w_i + c_i$.

The limited volatility condition is thus a joint condition on the range of ex-post uncertainty and the relative magnitude of the unit variable cost relative to the full cost. If the separability condition $R_i(q_i, \epsilon_{it}) = \epsilon_{it} \cdot \hat{R}_i(q_i)$ with $E(\tilde{\epsilon}_{it}) = 1$ is met, the limited volatility condition holds if and only if $\underline{\epsilon}_{it} \geq \frac{w_i}{w_i + c_i}$. We emphasize that the limited volatility condition does not trivialize the capacity utilization decision. The incentive compatibility conditions need to consider the possibility that the managers may opt for larger capacity levels than what is optimal for the firm, and if they do so it may then become sequentially rational not to exhaust the available capacity in unfavorable states of of the world. Such possibility may lead the parties away from the desired capacity choice, k_i^o.

The next result assume that the divisional revenue functions take the following quadratic form

$$R_i(q, \epsilon_{it}|\theta_i) = \theta_i \cdot \epsilon_{it} \cdot q - h_i \cdot q^2, \tag{17}$$

where θ_i is a positive constant. Though the quadratic functional form in (17) is common knowledge, the firm's central office does not have sufficient information about the divisional revenue functions because the parameters (θ_1, θ_2) are known only to the two divisional managers.

Proposition 2 *Suppose the divisional revenue functions take the quadratic form in (17) and the limited volatility condition is satisfied in the dedicated capacity setting. A system of decentralized initial capacity choices combined with the full cost transfer pricing rule*

$$TP_t(k_2, q_{2t}) = c \cdot k_2 + w_2 \cdot q_{2t}$$

achieves strong goal congruence, provided the divisions are free to renegotiate the initial capacity rights and capacity assets are depreciated according to the annuity rule.

Suppose (k_1, k_2) constitutes a Nash equilibrium of the divisional capacity choice game with $k_i > 0$ for each i. By the Envelope Theorem, the following first-order conditions are met:

$$E_\epsilon \left[(1 - \delta) \cdot CM_1'(k_1|w_1, \tilde{\epsilon}_{1t}) + \delta \cdot S(k_1 + k_2|w, \tilde{\epsilon}_t) \right] = c \tag{18}$$

and

$$E_\epsilon \left[\delta \cdot CM_2'(k_2|w_2, \tilde{\epsilon}_{2t}) + (1 - \delta) \cdot S(k_1 + k_2|w, \tilde{\epsilon}_t) \right] = c, \tag{19}$$

where $CM_i'(k_i|w_i, \epsilon_{it}) \equiv R_i'(q_i^o(k_i, \cdot), \epsilon_{it}) - w_i$ is the marginal contribution margin in the dedicated capacity scenario. It can be verified from the proofs of Lemmas 1 and 2 that $CM_i'(\cdot)$ and $S(\cdot)$ are decreasing functions of k_i, and hence each division's objective function is globally concave.

Similar to the arguments in Dutta and Reichelstein (2010), the above first-order conditions show that each division's incentives to acquire capacity stem both from the

unilateral "stand-alone" use of capacity as well as the prospect of trading capacity with the other division. The second term on the left-hand side of both (18) and (19) represents the firm's aggregate and *optimized* marginal contribution margin, given by the (expected) shadow price of capacity. Since the divisions individually only receive a share of the aggregate return (given by δ and $1 - \delta$, respectively), this part of the investment return entails a "classical" holdup problem.[19] Yet, the divisions also derive *direct* value from the capacity available to them, even if the overall capacity were not to be reallocated ex-post. The corresponding marginal revenues are given by the first terms on the left-hand side of equations (18) and (19), respectively.[20]

The proof of Proposition 2 shows that the quadratic form of divisional revenues in (17) has the property that the resulting shadow price function $S(k|\theta, w, \epsilon_t)$ is linear in ϵ_t. Combined with the limited volatility condition, linearity of the shadow price $S(\cdot)$ in ϵ_t implies that the efficient capacity in the fungible capacity scenario is the same as in the dedicated capacity setting; i.e., $k^* = k_1^o + k_2^o$. Furthermore, when the limited volatility condition holds, the stand-alone capacity levels (k_1^o, k_2^o) are the unique solution to the divisional first-order conditions in (18) and (19).

To ensure that the divisions produce efficient quantities, it is essential that the divisions are allowed to renegotiate initial capacity rights after the revenue shocks $(\epsilon_{1t}, \epsilon_{2t})$ are realized in each period. If the divisions were prohibited from such ex-post renegotiations, the two divisions would still choose the efficient capacity levels (k_1^o, k_2^o) provided the quadratic revenue functions are quadratic and the limited volatility conditions holds. However, the production choices would now be those that maximize the stand-alone divisional contribution margins (i.e., $q_i^o(k_i^o, \cdot)$), rather than the firm-wide contribution margin (i.e., $q_i^*(k_1^o + k_2^o, \cdot)$). Clearly, the efficiency loss from this inability to tailor divisional production plans to ex-post revenue information would be

---

[19]Earlier papers on transfer pricing that have examined this hold-up effect include Edlin and Reichelstein (1995), Baldenius et al.(1999), Anctil and Dutta (1999), Wielenberg (2000), Goex and Schiller (2007) and Pfeiffer et al. (2009).

[20]A similar convex combination of investment returns arises in the analysis of Edlin and Reichelstein (1995), where the parties sign a fixed quantity contract to trade some good at a later date. While the initial contract will almost always be renegotiated, its significance is to provide the divisions with a return on their relationship-specific investments, even if the status quo were to be implemented.

small when the revenue shocks ϵ_{it} are concentrated around a narrow range of values, but large when the probability density function of ϵ_{it} is more spread out.

We note that the mechanism in Proposition 2 allows each division to acquire capacity rights for its use at the outset. Equations (18) and (19) show that the firm would generally face an under-investment problem if only one division were allowed to secure capacity. For instance, if only the upstream division were to acquire capacity, its marginal contribution margin at the efficient capacity level k^* would be:

$$E_\epsilon \left[(1 - \delta) \cdot CM_1'(k^*|w_1, \tilde{\epsilon}_{1t}) + \delta \cdot S(k^*|w, \tilde{\epsilon}_t) \right].$$

This marginal revenue is, however, less than $E_\epsilon [S(k^*|w, \tilde{\epsilon}_t)] = c$ because

$$
\begin{aligned}
E_\epsilon [CM_1'(k^*|w_1, \tilde{\epsilon}_{1t})] &= E_\epsilon \left[R_1'(q^o(k^*, \cdot), \tilde{\epsilon}_{1t}) - w_1 \right] \\
&\leq E_\epsilon \left[R_1'(q_1^*(k^*, \cdot), \tilde{\epsilon}_{1t}) - w_1 \right] \\
&\leq E_\epsilon [S(k^*|w, \tilde{\epsilon}_t)],
\end{aligned}
$$

where the first inequality above is a consequence of the fact that $q_1^o(k^*, \cdot) \geq q_1^*(k^*, \cdot)$. Thus the upstream division would have insufficient incentives to secure the firm-wide optimal capacity level on its own, since it would anticipate a classic hold-up on its investment in the subsequent negotiations.

We now proceed to investigate the possibility of goal congruent transfer pricing rules if two major assumptions in Proposition 2, i.e., limited volatility and quadratic revenue functions, are not met. To solve the divisional coordination problem for a broader class of environments, we consider sequential mechanisms that give the upstream division additional supervisory authority. In effect, the upstream division can now be viewed as a "gatekeeper" whose approval is required for any capacity the downstream division wants to reserve for itself. Specifically, in order to acquire unilateral capacity rights, the downstream division needs to receive approval from the upstream division.[21] If the two divisions reach such an upfront agreement, it

---

[21]We focus on the upstream division as a gatekeeper because this division was assumed to have unique technological expertise in installing and maintaining production capacity. Yet, the following analysis makes clear that the role of the two divisions could be switched. Bockem and Schiller (2008) discuss specific gatekeeper arrangements in the context of a long-term supply contract between Samsung and Fairchild Semiconductor.

specifies the downstream division's unilateral capacity rights k_2 and a corresponding transfer payment $p(k_2)$ that must be made to the upstream division for granting these rights in each subsequent period. The parties report the outcome of this agreement $(k_2, p(k_2))$ to the central office, which commits to enforce this outcome unless the parties renegotiate it.

The upstream division is free to install additional capacity for its own needs in addition to what has been secured by the downstream division. As before, capacity assets are depreciated according to the annuity depreciation rule, and thus the upstream division is charged c for each unit of capacity that it acquires. If the parties fail to reach a mutually acceptable agreement, the downstream division would have no ex-ante claim on capacity, though it may, of course, obtain capacity ex-post through negotiation. Thus, $k_2 = p(k_2) = 0$ is the default scenario of no ex-ante contract. We summarize this negotiated gatekeeper transfer pricing arrangement as follows:

- The two divisions negotiate an ex-ante contract $(k_2, p(k_2))$ which gives Division 2 unilateral rights to k_2 units of capacity in return for a fixed payment of $p(k_2)$ in each period.

- Subsequently, Division 1 installs $k \geq k_2$ units of capacity,

- If Division 2 procures $q_{2t} \leq k_2$ units of output in period t, the corresponding transfer payments is calculated as $TP_t(k_2, q_{2t}) = p(k_2) + w_2 \cdot q_{2t}$.

- After observing the realization of revenue shocks ϵ_t in each period, the divisions can renegotiate the initial capacity rights.

For the result below, we assume that the optimal dedicated capacity level k_i^o is non-zero for each i. A necessary and sufficient condition for k_i^o to be positive is

$$E_{\epsilon_i}\left[R_i'(0, \epsilon_{it})\right] > c_i + w_i. \tag{20}$$

Proposition 3 *Suppose the divisional environments are stationary and the down-stream division's capacity rights are determined through bilateral negotiation. The transfer pricing rule*

$$TP_t(k_2, q_{2t}) = p(k_2) + w_2 \cdot q_{2t}$$

achieves strong goal congruence, provided the divisions are free to renegotiate the initial capacity rights in each period and operating assets are depreciated according to the annuity depreciation rule.

A gatekeeper arrangement will attain strong goal congruence if it induces the two divisions to acquire collectively the efficient capacity level, k^*. The proof of Proposition 3 demonstrates that in order to maximize their joint expected surplus, the divisions will agree on a particular amount of capacity level $k_2^* \in [0, k^*)$ that the downstream can claim for itself in any subsequent renegotiation. Thereafter, the upstream division has an incentive to acquire the optimal amount of capacity k^*, giving this division then an exclusive claim on $k^* - k_2^*$ units of capacity.

Suppose the two divisions have negotiated an ex-ante contract that gives the downstream division rights to k_2 units of capacity in each period. In response to this choice of k_2, the upstream division chooses $r_1(k_2)$ units of capacity for its own use, and thus installs $r_1(k_2) + k_2$ units of aggregate capacity. The upstream division's reaction function, $r_1(k_2)$, will satisfy the first-order condition in (18); i.e.,

$$E_\epsilon \left[(1 - \delta) \cdot CM_1'(r_1(k_2)|w_1, \tilde{\epsilon}_{1t}) + \delta \cdot S(r_1(k_2) + k_2|w, \tilde{\epsilon}_t) \right] = c. \tag{21}$$

As illustrated in Figure 4 below, the reaction function $r_1(k_2)$ is downward-slopping because both $CM_1'(k_1|\cdot)$ and $S(k|\cdot)$ are decreasing functions.
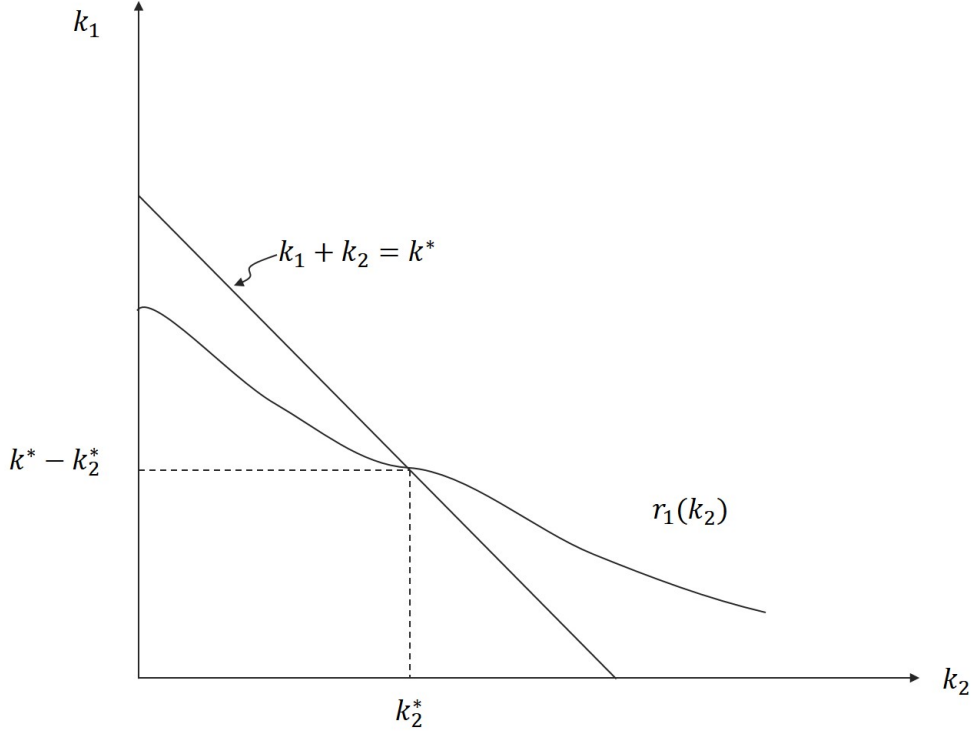
Figure 4: **Division 1's Reaction Function**

Furthermore, the proof of Proposition 3 shows that $r_1(0) \leq k^*$ and $r_1(k^*) > 0$. Therefore, as shown in Figure 4, there exists a unique $k_2^* \in [0, k^*)$ such that the upstream division responds with $r_1(k_2^*) = k^* - k_2^*$, and hence installs the optimal amount of aggregate capacity k_1^* on its own.

The ex-ante agreement $(k_2^*, p(k_2^*))$ must be such that it is preferred by both divisions to the default point of no agreement. If the two divisions fail to reach an ex-ante agreement, the upstream division will choose its capacity level unilaterally, and the downstream division will receive no initial capacity rights. By agreeing to grant k_2^* units of capacity rights to the downstream division, the two divisions can generate additional surplus. The fixed transfer payment $p(k_2^*)$ is chosen such that this additional surplus is split between the two divisions in proportion to their relative bargaining powers.

In the gatekeeper mechanism of Proposition 3, k_2^* denotes the capacity right of the downstream division as approved by the upstream division. Let $k_1^* = r_1(k_2^*)$ denote the upstream division's optimal response with $k_1^* + k_2^* = k^*$. We have the following

comparative statics result:

Corollary to Proposition 3 *The efficient capacity investment for each division, k_i^*, is independent of the bargaining power parameter δ. Furthermore, k_i^* is decreasing in the variable cost of production w_i for each i.*

At first glance, equation (21) for the upstream division 1's reaction function suggests that k_1^* should generally depend on the bargaining power parameter δ. However, the proof of above corollary shows that the downstream division's initial capacity rights are chosen so as to induce the upstream division to acquire k_1^o units of capacity for its own use. That is, $k_1^* = k_1^o$, and hence both k_1^* and $k_2^* = k^* - k_1^o$ are independent of δ. Intuitively, $k_1^* = k_1^o$ ensures that the expected marginal returns from the "stand-alone" use of capacity, $E_{\epsilon_1}[CM_1'(k_1^*, w_1, \tilde{\epsilon}_{1t})]$, and from the "combined" use of capacity, $E_\epsilon[S(k^*|w, \tilde{\epsilon}_t)]$, are both equal to the user cost of capacity c, and hence the upstream division does not face any hold-up problem.

We note that the downstream division is worse-off under the gatekeeper arrangement relative to a setting where the two divisions have symmetric capacity rights. The upstream division will extract some of the expected surplus contributed by the other division. At the same time, the rules in case the parties do not reach an agreement at the initial stage, can be specified differently from the one shown in Proposition 3. The same outcome, albeit with a different transfer payment, would result, for instance, if the mechanism were to specify that in the absence of an agreement the downstream division could claim any share of the available capacity at the transfer price:

$$TP(q_{2t}) = (c + w_2) \cdot q_{2t}.$$

We also note that the divisional efficient capacity levels (k_1^*, k_2^*) remain unchanged with this alternative default specification.

Our goal congruence results in Propositions 2 and 3 do not depend on the direction or extent of any correlation between the periodic shocks $\tilde{\epsilon}_{1t}$ and $\tilde{\epsilon}_{2t}$. Correlation does matter for goal congruence provided the divisional managers have the same expectation regarding future revenues. However, the optimal capacity level, k^* will

generally vary with the extent of correlation in the periodic shocks to the divisional revenues.

The allocation mechanism in Proposition 3 can be interpreted as a hybrid between full cost and negotiated transfer pricing such that the upstream division is charged for the full cost of the entire capacity and output produced by the divisions. Those charges are split between the two divisions through a two-stage negotiation. The latter feature is also the key to the efficiency of the fixed quantity contracts in Edlin and Reichelstein (1995). In their model, a properly set default quantity of a good to be traded provides the parties with incentives to make efficient relationship-specific (unverifiable) investments. In the context of our model, an agreement on the unilateral capacity rights of the downstream division induces the investment center to acquire residual capacity rights for itself such that the overall capacity procured is efficient from a firm-wide perspective.

5   Non-Stationary Environments

Our analysis has thus far focused on characterizing transfer pricing rules and depreciation schedules for environments in which the firm's economic fundamentals are, in expectation, identical across periods. Absent stationarity, divisional managers will generally anticipate that alternative capacity levels will induce tradeoffs in terms of the divisional profits attained in different time periods and thus violate the requirement of strong goal congruence, unless the initial capacity expenditure can be charged in a suitable manner via depreciation charges and transfer prices. In this section, we relax the assumption of stationarity by allowing the divisional revenue parameters x_{it}, variable costs w_{it}, and the distributions of revenue shocks ϵ_{it} to differ across periods.

Returning first to the dedicated capacity setting, the following result extends Lemma 1 to settings with non-stationary environments:

Lemma 3 *If capacity is dedicated, k_i^o, in (5) is given by the unique solution to:*[22]

---

[22]Again, Lemma 3 only holds as stated if $k_i^o > 0$. Furthermore, $k_i^o = 0$ if and only if the left-hand side of (5) is less than $c_i + w_i$ at $k_i = 0$.

$$\sum_{t=1}^{T} E_{\epsilon_{it}} \left[x_{it} \cdot R_i'(q_{it}^o(k_i^o, \tilde{\epsilon}_{it}, x_{it}, w_{it}), \tilde{\epsilon}_{it}) \right] \cdot \gamma^t = v_i + \bar{w}_i \tag{22}$$

where

$$\bar{w}_i = \sum_{t=1}^{T} w_{it} \cdot \gamma^t.$$

It is readily seen that the claim in Lemma 3 reduces to that in Lemma 1 whenever $x_{it} = 1$, $w_{it} = w_i$ and the $\{\epsilon_{it}\}$ are *iid*. Beginning with the work of Rogerson (1997), earlier work on goal congruent performance measures has shown that if the expected cash flows vary across time periods, proper intertemporal cost allocation of the initial investment expenditure can be obtained by *relative benefit cost allocation*. Formally, the weights under the relative benefit cost allocation rule are given by:[23]

$$\hat{z}_{it} \equiv \frac{x_{it}}{\sum_{\tau=1}^{T} x_{i\tau} \cdot \gamma^\tau},$$

with the corresponding intertemporal cost charges per unit of capacity given by $\hat{c}_{it} \equiv \hat{z}_{it} \cdot v_i$.

With dedicated capacity, strong goal congruence for the downstream division hinges entirely on the transfer pricing rule, while the incentives for the upstream division are shaped entirely by the depreciation schedule. Since these two design instruments are separate in the dedicated capacity setting, the following results focus on transfer pricing rules that achieve strong goal congruence for Division 2.

Proposition 4 *With dedicated capacity, the following full cost transfer pricing rules achieve strong goal congruence for the downstream division:*

    i. $TP_t(k_2, q_{2t}) = \hat{c}_{2t} \cdot k_2 + w_{2t} \cdot q_{2t}$, provided $w_{2t} = x_{2t} \cdot w_2$.

    ii. $TP_t(k_2) = \hat{c}_{2t} \cdot k_2 + \hat{z}_{2t} \cdot \bar{w}_2 \cdot k_2$, provided the limited volatility condition holds and the $\{\epsilon_{2t}\}$ are i.i.d.

---

[23]As pointed out by earlier studies, the corresponding relative benefit depreciation charges will coincide with straight-line depreciation if the x_{it} decline linearly over time at a particular rate (Nezlobin et al. 2012).

Part (i) of the above result shows that the two-part full cost transfer pricing rule, as described in the stationary setting of Proposition 1, can also be used to generate strong goal congruence for non-stationary environments provided the intertemporal variations in periodic revenues and variable costs exhibit identical patterns; i.e., $w_{2t} = x_{2t} \cdot w_2$. Absent such coordinated variations in periodic revenues and variable costs, the two-part full cost transfer pricing rule $TP_t(k_2, q_{2t}) = \hat{c}_{2t} \cdot k_2 + w_{2t} \cdot q_{2t}$ will generally fail to meet the requirement of strong goal congruence because the optimal capacity choice k_2^* would no longer be the preferred choice independently of how the manager trades-off performance attained in different time periods.

If the limited volatility assumption introduced in connection with Proposition 2 above is met, part (ii) of Proposition 4 shows that goal congruence is still attainable through a two-part full cost transfer pricing mechanism. The downstream division is then not charged for actually incurred variable costs, but for the present value of anticipated future variable costs, $\bar{w}_2 \cdot k_2$. This may be interpreted as "conservative" since the anticipated variable costs are viewed as intrinsic follow-on costs associated with the initial capacity investment. The overall charge must be apportioned in accordance with the relative benefit rule so as to ensure that the downstream division does not face any intertemporal trade-offs at the initial decision stage.

Suppose the downstream division is entirely focused on its performance measure in period t. Given the transfer pricing rule in part (ii) of Proposition 4, the downstream division will always produce at the capacity limit because it is not charged for actually incurred variable costs. Consequently, Division 2 will choose its capacity k_2 so as to maximize:

$$x_{2t} \cdot E_{\epsilon_2}[R_2(k_2, \epsilon_{2t})] - [\hat{c}_{2t} \cdot k_2 + \hat{z}_{2t} \cdot \bar{w}_2 \cdot k_2]. \tag{23}$$

The above objective function has a unique maximizer since $R_2(q, \cdot)$ is concave in q. We also note that the objective function in (23) is proportional to the objective function of the firm for any $k_2 \leq k_2^o$ because $q_2^o(k_2, w_2, \epsilon_{2t}) \equiv k_2$ for any $k_2 \leq k_2^o$. It thus follows that k_2^o is the unique maximizer of the downstream division's objective function in (23). We finally note that $TP_t(k_2) = \hat{c}_{2t} \cdot k_2 + \hat{z}_2 \cdot \bar{w}_2 \cdot k_2$ is a full-cost transfer pricing rule because, given limited volatility, Division 2 will in equilibrium

31

procure k_2^o and subsequently exhaust the available capacity.

If the limited volatility condition in part (ii) of Proposition 4 fails to hold, the downstream division will generally tend to underinvest. Appendix B shows in examples that as the limited volatility condition fails to hold by an "increasing margin", the corresponding loss in capacity investments and firm-wide value (NPV) increases continuously and monotonically.

The results of Baldenius, Nezlobin and Vaysman (2016, Proposition 1) have shown that goal congruence can be achieved in non-stationary environments where the capacity choices are binary rather than continuous. Specifically, suppose that Division 2 can install a specific amount of capacity $\bar{k}_2$ or not; i.e., $k_2 \in \{0, \bar{k}_2\}$. Suppose further that the divisional manager's private information is a one-dimensional parameter θ_2 that shifts the probability distributions of $\tilde{\epsilon}_{2t}$, but otherwise the revenue function $R_2(\cdot, \tilde{\epsilon}_{2t})$ is publicly known. The essential simplification with binary investment choices is that the accrual accounting rules, i.e., depreciation schedule and transfer pricing rule, only need to separate those θ_2 types for whom capacity investment is in the firm's interest from the ones for whom it is not. Accordingly, we denote the threshold type where the firm is just indifferent between investing and not investing by θ_2^*. Thus,

$$NPV_2(\bar{k}_2|\theta_2^*) = \sum_{t=1}^{T} E_{\epsilon_2}\left[CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t})|\theta_2^*)\right] \cdot \gamma^t - v_2 \cdot \bar{k}_2 = 0. \qquad (24)$$

As before, $CM_{2t}(\cdot)$ denotes the maximized value of the expected future contribution margin in period t:

$$CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \epsilon_{2t}) \equiv x_{2t} \cdot R_2(q_{2t}^o(\bar{k}_2, \cdot), \epsilon_{2t}) - w_{2t} \cdot q_{2t}^o(\bar{k}_2, \cdot).$$

Following the terminology in Baldenius, Nezlobin and Vaysman (2016), we refer to the *Relative Expected Optimized Benefit* (REOB) cost allocation rule as:

$$\bar{z}_{2t} = \frac{E_{\epsilon_2}\left[CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t})|\theta_2^*\right]}{\sum_{\tau=1}^{T} E_{\epsilon_2}\left[CM_{2\tau}(\bar{k}_2|x_{2\tau}, w_{2\tau}, \tilde{\epsilon}_{2\tau})|\theta_2^*\right] \cdot \gamma^\tau}.$$

The REOB rule is effectively the relative benefit rule corresponding to the threshold type θ_2^*. Finally, we denote $\bar{c}_{2t} \equiv \bar{z}_{2t} \cdot v_2$.

Proposition 5 *With dedicated capacity, suppose Division 2' capacity choice is binary and the future realizations of $\tilde{\epsilon}_{2t}$ are drawn according to conditional densities $f(\epsilon_{2t}|\theta_2)$ such that θ_2 shifts $f(\epsilon_{2t}|\theta_2)$ in the sense of first-order stochastic dominance. The full cost transfer pricing rule:*

$$TP_t(k_2, q_{2t}) = \bar{c}_{2t} \cdot k_2 + w_{2t} \cdot q_{2t}$$

then achieves strong goal congruence for the downstream division.

As shown in Appendix A, the net present value of the capacity investment, $NPV_2(\bar{k}_2|\theta_2)$, is increasing in the distributional parameter θ_2 because of the assumed first-order stochastic dominance effect of θ_2 on $\tilde{\epsilon}_{2t}$. If Division 2 were to focus exclusively on its profit measure in period t, $1 \leq t \leq T$, it would seek to maximize:

$$E_{\epsilon_2}\left[\Pi_2(k|\theta_2, \tilde{\epsilon}_{2t})\right] \equiv E_{\epsilon_2}\left[CM_{2t}(k_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t}, \theta_2)\right] - \bar{z}_{2t} \cdot v_2 \cdot k_2.$$

By construction of the REOB rule, $E_{\epsilon_2}\left[\Pi_2(k|\theta_2, \tilde{\epsilon}_{2t})\right] > 0$ if and only if

$$E_{\epsilon_2}\left[CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t})|\theta_2\right] > E_{\epsilon_2}\left[CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t})|\theta_2^*\right] \quad \text{or equivalently} \quad \theta_2 > \theta_2^*.$$

We finally turn to a setting of fungible capacity for non-stationary environments. Generalizing the result in Lemma 2, it can be shown that the optimal capacity k^* is given by:

$$E_\epsilon\left[\sum_{t=1}^T \gamma^t \cdot S_t(k^*|x_t, w_t, \tilde{\epsilon}_t)\right] = v$$

where

$$S_t(k|x_t, w_t, \epsilon_t) \equiv \max\{x_{1t} \cdot R_1'(q_{1t}^*(k, \cdot), \epsilon_{1t}) - w_{1t}, x_{2t} \cdot R_2'(q_{2t}^*(k, \cdot), \epsilon_{2t}) - w_{2t}\}$$

is the shadow price of capacity in period t.

The result below generalizes our finding in Proposition 3 to non-stationary environments. As in the setting of Proposition 3, suppose the divisional revenue functions, $x_{it} \cdot R_i(q, \epsilon_{it}|\theta_i)$, are parameterized by $\theta_i > 0$ and $R_i(q, \epsilon_{it}|\theta_i)$ takes the quadratic form in (17). We assume that while the functional form of divisional revenues is common

knowledge, the central office does not have complete information about the divisional revenue functions because the parameters (θ_1, θ_2) are known only to both divisional managers.

Proposition 6 *Suppose the divisional revenue functions take the quadratic form in (17), the limited volatility condition is satisfied in the dedicated capacity setting, and the revenue shocks ϵ_{it} are i.i.d. A system of decentralized initial capacity choices combined with the full cost transfer pricing rule*

$$TP_t(k_2) = \hat{c}_{2t} \cdot k_2 + \hat{z}_{2t} \cdot \bar{w}_2 \cdot k_2$$

achieves strong goal congruence provided (i) the divisions are free to renegotiate the initial capacity rights, (ii) the anticipated variable production costs of each division, $\bar{w}_i \cdot k_i$, are capitalized, and (iii) the capitalized costs, $(v + \bar{w}_i) \cdot k_i$, are depreciated according to the respective relative benefit rule.

Unlike the goal congruent mechanism identified in the stationary setting of Proposition 3, the above mechanism does not charge the divisions for the variable costs at the actual rates of w_{it} per unit. Instead, the present value of anticipated future variable costs, $\bar{w}_i \cdot k_i$, for each division is capitalized along with the fixed capacity costs at the outset. The capitalized costs are then allocated across subsequent periods according to the relative benefit rule to ensure that the divisions do not face any intertemporal trade-offs in choosing their initial capacity levels.

The transfer pricing rule $TP_t(k_2) = \hat{c}_{2t} \cdot k_2 + \hat{z}_{2t} \cdot \bar{w}_2 \cdot k_2$ ensures that the stand-alone capacity levels (k_1^o, k_2^o) constitute a Nash equilibrium at the initial stage, provided the divisions can renegotiate their initial capacity rights. Thus, the two divisions will install the efficient capacity level k^* at the outset, since the efficient capacity in the fungible setting is the same as in the dedicated setting; i.e., $k^* = k_1^o + k_2^o$, owing to the linearity of the shadow price of capacity in $\tilde{\epsilon}_{it}$.

6   Conclusion

This paper has re-examined the incentive properties of full cost transfer pricing rule in multi-divisional firms. Our analysis is motivated by the fact that this form of internal pricing remains ubiquitous in practice despite the many concerns that have been expressed about it in textbooks and the academic literature. The main ingredients in our model are that divisional managers are responsible for the initial acquisition of capacity as well as its subsequent utilization in future periods. An upstream division installs capacity and provides production services for both divisions, since it has the necessary technical expertise. In each period, the upstream division receives a transfer payment for providing capacity- and production services to the downstream division.

We identify circumstances in which a suitable variant of full cost transfer pricing induces efficient capacity acquisition and subsequent production decisions. From an ex-ante capacity planning perspective, variable cost pricing is clearly inadequate because the buying division will not internalize the relevant capacity costs, and hence this pricing rule generates incentives for the buying division to initially request an excessive amount of capacity. At the same time, a simplistic form of full cost transfer pricing that charges the buying division only for the cost of actually utilized capacity will also not achieve efficient outcomes as this rule again motivates the buying division to request an inefficiently large amount of capacity. Our results demonstrate that, depending on the characteristics of the underlying production and market environment, particular variants of two-part full cost transfer pricing can indeed lead to efficient decentralization.

When the divisions can share the same productive assets for their production needs, efficient capacity investments and ex-post allocations can be achieved through bilateral negotiation. Potential hold-up problems on investments resulting from ex-post negotiation can be alleviated through an appropriate assignment of initial capacity rights that determine the divisions' default payoffs at the negotiation stage.

35

Appendix A

Proof of Lemma 1:

With dedicated capacity, the firm's objective function is additively separable across the two divisions. The firm seeks a capacity level k_i^o for division i to maximize the net present value of expected cash flows

$$NPV_i(k_i) = \sum_{t=1}^{T} E_{\epsilon_i}[R_i(q_i^o(k_i, \cdot), \tilde{\epsilon}_{it}) - w_i \cdot q_i^o(\cdot)]\gamma^t - v_i \cdot k_i. \tag{25}$$

Dividing by the annuity factor $\sum_{t=1}^{T} \gamma^t$, the objective function in (25) can be equivalently expressed as follows:

$$E_{\epsilon_i}[CM_i(k_i|w_i, \tilde{\epsilon}_{it})] - c \cdot k_i,$$

where

$$CM_i(k_i|w_i, \epsilon_{it}) \equiv R_i(q_i^o(k_i, \cdot), \epsilon_{it}) - w_i \cdot q_i^o(k_i, \cdot)$$

is the maximized value of the contribution margin. Once the differentiability of the value function, $CM_i(\cdot)$ has been established, the claim in equation (5) of Lemma 1 essentially amounts to the well-known result that the derivative of the value function with respect to k_i is given by the shadow price of the constrained resource (i.e., c_i); see, for instance, Zangwill (1969, p.69).

Claim: $CM_i(k_i|w_i, \epsilon_{it})$ is differentiable in k_i for all ϵ_{it} and w_i. Furthermore,

$$\frac{\partial}{\partial k_i} CM_i(k_i|w_i, \epsilon_{it}) = R_i'(q_i^o(k_i, \cdot), \epsilon_{it}) - w_i.$$

Proof of Claim:[24] We first note that

---

[24] The validity of this claim can be verified by invoking proposition 4.3 in Oyama and Takenawa (2018). In particular, they postulate as a sufficient condition that the value function, $CM_i(k_i|w_i, \epsilon_{it})$, be concave. This holds in our context due to the assumed concavity of the divisional revenue functions. We provide here an elementary argument that $CM_i(k_i|w_i, \epsilon_{it})$ is in fact everywhere differentiable in k_i.

$$\frac{CM_i(k_i + \Delta | w_i, \epsilon_{it}) - CM_i(k_i | w_i, \epsilon_{it})}{\Delta}$$

$$\geq \frac{R_i(q_i^o(k_i, \cdot) + \Delta, \epsilon_i) - R_i(q_i^o(k_i, \cdot), \epsilon_{it})}{\Delta} - w_i. \tag{26}$$

This inequality follows directly by observing that

$$CM_i(k_i + \Delta | w_i, \epsilon_{it}) \geq R_i(q_i^o(k_i, \cdot) + \Delta, \epsilon_{it}) - w_i \cdot [q_i^o(k_i, \cdot) + \Delta].$$

At the same time, we find that

$$\frac{CM_i(k_i + \Delta | w_i, \epsilon_{it}) - CM_i(k_i | w_i \epsilon_{it})}{\Delta}$$

$$\leq \frac{R_i(q_i^o(k_i + \Delta, \cdot), \epsilon_{it}) - R_i(q_i^o(k_i + \Delta, \cdot) - \Delta, \epsilon_{it})}{\Delta} - w_i. \tag{27}$$

To see this, we note that

$$CM_i(k_i + \Delta | w_i, \epsilon_{it}) - CM_i(k_i | w_i, \epsilon_{it})$$

$$\leq R_i(q_i^o(k_i + \Delta, w_i, \epsilon_{it}), \epsilon_{it}) - w_i \cdot q_i^o(k_i + \Delta, w_i, \epsilon_{it})$$

$$- [R_i(q_i^o(k_i + \Delta, w_i, \epsilon_{it}) - \Delta, \epsilon_{it}) - w_i \cdot (q_i^o(k_i + \Delta, w, \epsilon) - \Delta)].$$

because $q_i^o(k_i + \Delta, w_i, \epsilon_{it}) - \Delta \leq k_i$ if the division invested $k_i + \Delta$ units of capacity. We also note that for Δ sufficiently small, $q_i^o(k_i + \Delta, w_i, \epsilon_{it}) - \Delta \geq 0$ because $q_i^o(k_i, w_i, \epsilon_{it}) > 0$ by the assumption that $R_i'(0, \epsilon_{it}) - w_i > 0$.

By the Intermediate Value Theorem, the right-hand side of (27) is equal to

$$\frac{R_i'(\hat{q}_i(\Delta), w_i, \epsilon_i) \cdot \Delta}{\Delta} - w_i,$$

for some intermediate value $\hat{q}_i(\Delta)$ such that $q_i^o(k_i + \Delta, \cdot) - \Delta \leq \hat{q}_i(\Delta) \leq q_i^o(k_i + \Delta, \cdot)$. As $\Delta \to 0$, the right-hand side in both (26) and (27) converge to the following:

$$R_i'(q_i^o(k, w_i, \epsilon_{it}), \epsilon_{it}) - w_i,$$

proving the claim.

If $k_i^o > 0$ is the optimal capacity level, then

$$\frac{\partial}{\partial k_i} \left[E_\epsilon[CM_i(k_i^o|w_i, \tilde{\epsilon}_i)] - c_i \cdot k_i^o \right] = E_{\epsilon_i} \left[\frac{\partial}{\partial k_i} CM_i(k_i^o|w_i, \tilde{\epsilon}_i) \right] - c_i$$

$$= E_{\epsilon_i}[R_i'(q_i^o(k_i, \cdot), \tilde{\epsilon}_i)] - (c_i + w_i)$$

$$= 0.$$

Thus k_i^o satisfies equation (5) in the statement of Lemma 1.

To verify uniqueness, suppose that both k_i^o and $k_i^o + \Delta$ satisfy equation (5). Since by definition $q_i^o(k_i^o + \Delta, \cdot) \geq q_i^o(k^o, \cdot)$ for all ϵ_{it}, it would follow that in fact

$$q_i^o(k_i^o + \Delta, \cdot) = q_i^o(k_i^o, \cdot)$$

for all ϵ_{it}. That in turn would imply that the optimal production quantity in the absence of a capacity constraint, i.e., $\hat{q}_i(\epsilon_{it}, \cdot)$, is less than k_i^o, and therefore

$$E_{\epsilon_i}[R_i'(\hat{q}_i(\tilde{\epsilon}_{it}, \cdot), \tilde{\epsilon}_{it})] = E_{\epsilon_i}[R_i'(q_i^o(k_i^o, \tilde{\epsilon}_{it}, w_i), \tilde{\epsilon}_{it})] = w_i,$$

which would contradict that k_i^o satisfies equation (5) in the first place. □

Proof of Proposition 1:

Contingent on (k_1, k_2) and $(q_{1t}, q_{2t}) \leq (k_1, k_2)$, Division 1's residual income performance measure in period t is given by

$$\Pi_{1t} = R_1(q_{1t}, \epsilon_{1t}) - w_1 \cdot q_{1t} - w_2 \cdot q_{2t} + TP(q_{2t}, k_2) - z_{1t} \cdot v_1 \cdot k_1 - z_{2t} \cdot v_2 \cdot k_2. \quad (28)$$

Regardless of the decisions made by Division 2, Division 1 will therefore choose the production quantity $q_1^o(k_1, \cdot)$ that maximizes its contribution margin in period t. If Division 1 were to make the initial capacity decision entirely with a focus on its expected performance measure in period t, i.e., $E_{\epsilon_1}[\tilde{\Pi}_{1t}]$, it will choose k_1^o if and only if:

$$z_{1t} \cdot v_1 = \frac{1}{\sum_t \gamma^t} \cdot v_1 \equiv c_1.$$

This follows immediately from the observation that the objective function of the central office, $NPV_1(k_1)$, is proportional to (28) once the terms pertaining to Division

38

2 are disregarded in (28). As observed in Rogerson (1997), there is a 1-1 mapping between depreciation schedules and intertemporal cost charges:

$$z_{1t} = d_{1t} + r(1 - \sum_{\tau=1}^{t-1} d_{1\tau}).$$

Furthermore, if $z_{1t} \cdot v_1 = \frac{1}{\sum_t \gamma^t}$, then the unique corresponding depreciation schedule is the annuity rule (Rogerson, 1997).

To show that only the transfer pricing rule $TP_t(k_2, q_{2t}) = c_2 \cdot k_2 + w_2 \cdot q_{2t}$ achieves strong goal congruence for Division 2, we note that the function $TP_t(k_2, q_{2t})$ is naturally indeterminate for any $q_{2t} > k_2$, as this contingency can never arise. In addition, $TP_t(k_2, q_{2t})$ can only be unique up to an additive constant. In the dedicated capacity scenario with a stationary environment we can drop both the index 2 and the index t, without loss of generality. Suppose now that Division 2's revenue is of the quadratic form:

$$R(q, \epsilon|\theta) = \theta \cdot \epsilon \cdot q - \frac{1}{2} \cdot h \cdot q^2$$

For a given value of θ, we denote by $k^o(\theta)$ the efficient capacity level holding fixed the parameter h and the probability distribution of the shocks $\tilde{\epsilon}$ on the interval $[0, \bar{\epsilon}]$. The first-order condition for the efficient production quantities are:

$$w \equiv \theta \cdot \epsilon - h \cdot q^o(\epsilon, k^o(\theta)) \equiv \frac{\partial}{\partial q} TP(q^o(\epsilon, k^o(\theta)), k^o(\theta)). \tag{29}$$

The first equality reflects that $q^o(\epsilon, k^o(\theta))$ is the ex-post efficient production quantity. The second equation represents the first-order condition for Division 2' quantity choice in any given period, following the realization of ϵ. These equations have to hold as identities for all $\epsilon \geq \epsilon(k^o(\theta))$, where $q^o(\epsilon(k^o(\theta)), k^o(\theta))$ is defined by the requirement that $q^o(\epsilon(k^o(\theta)), k^o(\theta)) = k^o(\theta)$. By the Fundamental Theorem of Calculus:

$$TP(q, k^o(\theta)) = \int_0^q \frac{\partial}{\partial q} TP(y, k^o(\theta)) \, dy + TP(0, k^o(\theta)).$$

As ϵ varies on $[0, \epsilon(k^o(\theta))]$, the quantity $q^o(\epsilon, k^o(\theta))$ varies linearly between zero and $k^o(\theta)$, (and $q^o(\epsilon, k^o(\theta)) = k^o(\theta)$ for $\epsilon \geq \epsilon(k^o(\theta))$). Thus,

$$TP(q, k^o(\theta)) = w \cdot q + C(k^o(\theta)),$$

for $0 \leq q \leq k^o(\theta)$, with $C(k^o(\theta)) \equiv TP(0, k^o(\theta))$. Since $k^o(\theta)$ is an increasing and unbounded function of θ, it follows that $TP(q, k) = w \cdot q + C(k)$, modulo some additive constant, for all $q \leq k$.

For the final step in the proof, suppose the environment is such that $\tilde{\epsilon}$ concentrates its probability mass entirely on the value $\hat{\epsilon}$. The incentive compatibility condition for the initial capacity choice then becomes:

$$\theta \cdot \hat{\epsilon} - h \cdot q^o(\hat{\epsilon}, k^o(\theta)) - w = C'(k^o(\theta)). \tag{30}$$

Since $k^o(\theta)$ is the efficient capacity level, Lemma 1 shows that the left-hand side of (30) is equal to c and thus $C'(k^o(\theta)) = c$. Varying the value of θ then shows that $C'(k) \equiv c$. $\qquad\square$

Proof of Lemma 2:

The maximized contribution margin is given by

$$CM(k|w, \epsilon_t) = \sum_{i=1}^{2}[R_i(q_i^*(k, \cdot), \epsilon_{it}) - w_i \cdot q_i^*(k_i, \cdot)]$$

where $q_i^*(k, \cdot) \equiv q_i^*(k, w, \epsilon_t)$

<u>Claim:</u> $CM(k|w, \epsilon_t)$ is differentiable in k for any ϵ_t and w, such that

$$\frac{\partial}{\partial k} CM(k|w, \epsilon_t) = S(k|w, \epsilon_t),$$

where

$$S(k|w, \epsilon_t) = \max\{R_1'(q_1^*(k, \cdot), \epsilon_{1t}) - w_1, R_2'(q_2^*(k, \cdot), \epsilon_{2t}) - w_2\}.$$

The proof is analogous to the one in Lemma 1 showing the differentiability of $CM_i(\cdot)$.

The expected value of the maximized contribution margin, $E_\epsilon[CM(k|w, \tilde{\epsilon}_t)]$, is identical across periods in the stationary setting. Hence, the firm will choose the

optimal capacity level k^* to maximize the following objective function:

$$E_\epsilon \left[CM(k|w, \tilde{\epsilon}_t) \right] - c \cdot k.$$

Equation (13) in the statement of Lemma 3 then follows from the first-order condition of the above optimization problem. The uniqueness of k^* follows from an argument similar to that in the proof of Lemma 1. □

Proof of Proposition 2:

We first show that with quadratic revenue functions of the form $R_i(q, \epsilon_{it}|\theta_i) = \theta_i \cdot \epsilon_{it} \cdot q - h_i \cdot q^2$ and limited volatility, the efficient capacity level in the fungible scenario is equal to the sum of the efficient capacity levels in the dedicated capacity scenario; that is

$$k^* = k_1^o + k_2^o.$$

From Lemma 1, we know that in the dedicated capacity setting the efficient capacity levels satisfy:

$$E_{\epsilon_i}[R_i'(q_1^o(k_i^o, \tilde{\epsilon}_{it}), \tilde{\epsilon}_{it}|\theta_i)] - w_i = c.$$

The limited volatility condition implies that for all $\tilde{\epsilon}_{it}$

$$q_i^o(k_i^o, \tilde{\epsilon}_{it}) = k_i^o.$$

Furthermore, with quadratic revenue functions, we find that

$$
\begin{aligned}
E_{\epsilon_i}[R_i'(q_i^o(k_i^o, \tilde{\epsilon}_{it}), \tilde{\epsilon}_{it}|\theta_i)] - w_i &= E_{\epsilon_i}[R_i'(k_i^o, \tilde{\epsilon}_{it}|\theta_i)] - w_i \\
&= R_i'(k_i^o, \hat{\epsilon}_{it}|\theta_i)] - w_i \\
&= c,
\end{aligned}
\tag{31}
$$

where $\hat{\epsilon}_{it} \equiv E(\tilde{\epsilon}_{it})$. In the fungible capacity scenarios, Lemma 2 has shown that at the efficient k^*

$$E_\epsilon[R_i'(q_i^*(k^*, \cdot), \tilde{\epsilon}_{it}|\theta_i)] - w_i = c.$$

It is readily seen that in the quadratic revenue scenario, $q_i^*(k^*, \theta, w, \epsilon_t)$ is linear in ϵ_t

provided that $q_i^*(\cdot) > 0$ for all ϵ_t. Thus,

$$
\begin{aligned}
E_\epsilon[S(k^*|\theta, w, \tilde{\epsilon}_t)] &= E_\epsilon[R_i'(q_i^*(k^*, \cdot), \tilde{\epsilon}_{it}|\theta_i)] - w_i \\
&= E_\epsilon[\theta_i \cdot \tilde{\epsilon}_{it} - 2h_i \cdot q_i^*(k^*, \theta, w, \tilde{\epsilon}_t)] - w_i \\
&= \theta_i \cdot \hat{\epsilon}_{it} - 2h_i \cdot q_i^*(k^*, \theta, w, \hat{\epsilon}_t) - w_i \\
&= R_i'(q_i^*(k^*, \theta, w, \hat{\epsilon}_t), \hat{\epsilon}_{it}|\theta_i) - w_i \\
&= c, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (32)
\end{aligned}
$$

where $\hat{\epsilon}_t \equiv E(\epsilon_t)$. It follows from (31) and (32) that

$$
q_i^*(k^*, \theta, w_i, \hat{\epsilon}_t) = k_i^o,
$$

and thus $k^* = k_1^o + k_2^o$.

It remains to show that (k_1^o, k_2^o) is a Nash equilibrium at the initial date. Given the full cost transfer pricing rule $TP_t(k_2, q_{2t}) = c \cdot k_2 + w_2 \cdot q_{2t}$, the divisional profit of Division 2 in period t, contingent on ϵ_t and k_1^o is:

$$
\pi_{2t}(k_2, \epsilon_t|k_1^o) = \delta \cdot CM_2(k_2|\theta_2, w_2, \epsilon_{2t}) + (1-\delta)[CM(k_1^o + k_2|\theta, w, \epsilon_t) - CM_1(k_1^o|\theta_1, w_1, \epsilon_t)] - c \cdot k_2,
$$

where, as before,

$$
CM_i(k_i|\theta_i, w_i, \epsilon_{it}) = \max_{q_i \leq k_i}\{R_i(q_i, \epsilon_{it}|\theta_i) - w_i \cdot q_i\}.
$$

We note that in a stationary environment, the expected value of Division 2's profit, $E_\epsilon[\pi_{2t}(k_2, \tilde{\epsilon}_t|k_1^o)]$, is the same in each period. By definition, k_2^o is the unique maximizer of $E_{\epsilon_2}[\delta \cdot CM(k_2|\theta_2, w_2, \tilde{\epsilon}_{2t})] - \delta \cdot c \cdot k_2$. By Lemma 2, k_2^o maximizes $E_\epsilon[(1-\delta) \cdot CM(k_1^o + k_2|\theta, w, \tilde{\epsilon}_t)] - (1-\delta) \cdot c \cdot k_2$. It thus follows that k_2^o is also a maximizer of Division 2's expected profit in each period, $E_\epsilon[\pi_{2t}(k_2, \tilde{\epsilon}_t|k_1^o)]$.

A symmetric argument can be used to show that in order to maximize its expected residual income in any period, Division 1 will choose k_1^o if it conjectures that the downstream division chooses k_2^o. $\quad\square$

Proof of Proposition 3:

Suppose that the two divisions have agreed to an ex-ante contract under which the downstream division has initial rights for k_2 units of capacity for a transfer payment of $p(k_2) + w_2 \cdot q_{2t}$ in each period. Further, suppose that Division 1 has installed a capacity of k_1 units over which it has unilateral rights.

After observing ϵ_t, the two divisions will renegotiate the initial capacity rights to maximize the joint surplus in each period. Following the same arguments as used in deriving (16), it can be checked that Division 1's effective contribution margin after reallocation of capacity rights is given by

$$CM_1^*(k_1|k_2, w, \epsilon_t) = (1-\delta) \cdot CM_1(k_1|w_1, \epsilon_{1t}) + \delta \cdot [CM(k_1 + k_2|w, \epsilon_t) - CM_2(k_2|w_2, \epsilon_{2t})].$$

For stationary environments, the expected value of effective contribution margin, $E_\epsilon[CM_1^*(k_1|k_2, w, \epsilon_t)]$, is the same in each period. Since capacity assets are depreciated according to the annuity rule, this implies that taking k_2 as given, Division 1 will choose k_1 to maximize

$$E_\epsilon[CM_1^*(k_1|k_2, w, \tilde{\epsilon}_t)] - c \cdot k.$$

As a function of k_2, let $r_1(k_2)$ denote Division 1's optimal response; i.e., $k_1 = r(k_2)$ maximizes the above objective function. Let $r(k_2) \equiv r_1(k_2) + k_2$ denote the corresponding aggregate amount of capacity. Division 1's reaction function, $r_1(k_2)$, will satisfy the following first-order condition:

$$E_\epsilon\left[(1-\delta) \cdot CM_1'(r_1(k_2)|w_1, \tilde{\epsilon}_{1t}) + \delta \cdot S(r(k_2)|w, \tilde{\epsilon}_t)\right] \leq c, \tag{33}$$

which must hold as an equality whenever $r_1(k_2) > 0$. We note from (33) that $r_1(k_2)$ is downward slopping because $CM_1'(k_1|\cdot)$ and $S(k|\cdot)$ are decreasing functions of k_1 and k, respectively.

We now investigate the values of $r_1(k_2)$ at $k_2 = 0$ and $k_2 = k^*$. We first claim that $r_1(0) \leq k^*$. Suppose to the contrary, $r_1(0) > k^*$. This implies that $r(0) > k^*$, and hence

$$E_\epsilon[S(r(0)|w, \tilde{\epsilon}_t)] < E_\epsilon[S(k^*|w, \tilde{\epsilon}_t)] = c. \tag{34}$$

43

Furthermore,

$$
\begin{aligned}
E_{\epsilon_1}\left[CM_1'(r_1(0)|w_1,\tilde{\epsilon}_{1t})\right] &< E_{\epsilon_1}\left[CM_1'(k^*|w_1,\tilde{\epsilon}_{1t})\right] \\
&= E_{\epsilon_1}\left[R_1'(q_1^o(k^*,\tilde{\epsilon}_{1t},\cdot),\tilde{\epsilon}_{1t}) - w_1\right] \\
&\leq E_{\epsilon}\left[R_1'(q_1^*(k^*,\tilde{\epsilon}_t,\cdot),\tilde{\epsilon}_{1t}) - w_1\right] \\
&\leq E_{\epsilon}[S(k^*|w,\tilde{\epsilon}_t)] \\
&= c, \tag{35}
\end{aligned}
$$

where we have used the result that $q_1^o(k^*,\epsilon_{1t},\cdot) \geq q_1^*(k^*,\epsilon_t,\cdot)$ for all ϵ_t to derive the second inequality above. Inequalities in (34) and (35) imply that the first-order condition in (33) cannot hold as an equality, which contradicts the assumption that $r_1(0) > k^*$ is optimal.

We next claim that $r_1(k^*) > 0$, and hence $r(k^*) > k^*$. Suppose to the contrary $r_1(k^*) = 0$. This implies that $r(k^*) = k^*$, and hence

$$
E_{\epsilon}[S(r(k^*)|w,\tilde{\epsilon}_t)] = c.
$$

Furthermore,

$$
E_{\epsilon_1}[CM_1'(r_1(k^*)|w_1,\tilde{\epsilon}_{1t}] = E_{\epsilon_1}[CM_1'(0|w_1,\tilde{\epsilon}_{1t}] > c,
$$

because of the assumption in (20). It thus follows that the left hand side of (33) is strictly greater than c, which contradicts the assumption that $r(k_2) = 0$ is the optimal response to $k_2 = k^*$.

We have thus proven that $r(0) \leq k^*$ and $r(k^*) > k^*$. The Intermediate Value Theorem then implies that there exists a $k_2^* \in [0, k^*)$ such that $r(k_2^*) = k^*$. We have thus shown that if the two divisions sign an ex-ante contract that provides the downstream division with initial capacity rights of k_2^* units, Division 1 will choose the efficient amount of aggregate capacity k^*.

To complete the proof, we need to show that there exists a fixed transfer payment $p(k_2^*)$ such that the ex-ante contract $(k_2^*, p(k_2^*))$ will be preferred by both divisions to the default point of no agreement. If the two divisions fail to reach an ex-ante agreement, Division 1 will choose its capacity level unilaterally, and Division 2 will receive no capacity rights (i.e., $k_2 = 0$). Let $\hat{k}$ denote Division 1's optimal choice of

capacity under the "default" scenario. Division 1's expected periodic payoff under the default scenario is then given by

$$\hat{\Pi}_1 = E_\epsilon \left[(1 - \delta) \cdot CM_1(\hat{k}|w_1, \tilde{\epsilon}_{1t}) + \delta \cdot CM(\hat{k}|w, \tilde{\epsilon}_t) \right] - c \cdot \hat{k}$$

while Division 2's default payoff is

$$\hat{\Pi}_2 = (1 - \delta) \cdot E_\epsilon \left[CM(\hat{k}|w, \tilde{\epsilon}_t) - CM_1(\hat{k}|w_1, \tilde{\epsilon}_{1t}) \right].$$

By agreeing to transfer k_2^* units of capacity rights to Division 2, the two divisions can increase their periodic joint surplus by

$$\Delta\Pi \equiv E_\epsilon \left[CM(k^*|w, \tilde{\epsilon}_t) - c \cdot k^* \right] - E_\epsilon \left[CM(\hat{k}|w, \tilde{\epsilon}_t) - c \cdot \hat{k} \right].$$

The two divisions can then split this additional surplus between them in proportion to their relative bargaining power. The periodic transfer price $p(k_2^*)$ that implements this allocation is given by

$$E_\epsilon \left[(1 - \delta) \cdot CM_1(k^* - k_2^*|w_1, \tilde{\epsilon}_{1t}) + \delta \cdot CM(k^*|w, \tilde{\epsilon}_t) \right] - c \cdot k^* + p(k_2^*) = \hat{\pi}_1 + \delta \cdot \Delta\Pi.$$

Division 2's expected periodic payoff with this choice of transfer payment will be equal to $\hat{\Pi}_2 + (1 - \delta) \cdot \Delta\Pi$. Therefore, both divisions will prefer the ex-ante contract $(k_2^*, p(k_2^*))$ to the default scenario of no contract $(0, 0)$. □

Proof of Corollary to Proposition 3:

Since $r(k_2^*) = k^*$, it follows that $E_\epsilon[S(r(k_2^*)|w, \tilde{\epsilon}_t] = c$. The first-order condition in (33) for division 1's optimal response then yields

$$E_{\epsilon_1}[CM_1'(k_1^*)|w_1, \tilde{\epsilon}_{1t}] \leq c,$$

which can hold as a strict inequality only if $k_1^* \equiv r_1(k_2^*) = 0$. However, k_1^* cannot be zero because $E_{\epsilon_1}[CM_1'(0|w_1, \tilde{\epsilon}_{1t}] > c$ by assumption. It thus follow that

$$E_{\epsilon_1}[CM_1'(k_1^*)|w_1, \tilde{\epsilon}_{1t}] = c, \tag{36}$$

and hence $k_1^* = k_1^o$. Thus, the efficient capacity choices under the gatekeeper mechanism in Proposition 3 are uniquely given by $k_1^* = k_1^o$ and $k_2^* = k^* - k_1^o$. We claim that k_2^* is non-negative; i.e., $k^* - k_1^o \geq 0$. To prove this, suppose to the contrary $k^* < k_1^o$. This implies that

$$c \geq E_\epsilon[R_1'(q_1^*(k^*, \cdot), \epsilon_{1t}) - w_1] \geq E_{\epsilon_1}[R_1'(q_1^o(k^*, \cdot), \epsilon_{1t}) - w_1] > E_{\epsilon_1}[R_1'(q_1^o(k_1^o, \cdot), \epsilon_{1t}) - w_1],$$

which contradicts the optimality condition $c = E_{\epsilon_1}[R_1'(q_1^o(k_1^o, \cdot), \epsilon_{1t}) - w_1]$.

Clearly, k_1^* and k_2^* are independent of the bargaining power parameter δ. It follows from equation (36) that k_1^* is decreasing in w_1 because $CM_1'(k_1^*|w_1, \epsilon_{1t}) = R_1'(q_1^o(k_1^*, \cdot), \epsilon_{1t}) - w_1$ is decreasing in w_1. Since the shadow price $S(k^*|w, \epsilon_t)$ is decreasing in w_2, equation (13) implies that k^* is also decreasing in w_2. It thus follows that $k_2^* = k^* - k_1^o$ is decreasing in w_2. $\square$

Proof of Lemma 3:

The proof proceeds along the lines of the proof of Lemma 1. In particular

$$CM_{it}(k_i|w_{it}, \epsilon_{it}) = \max_{q_{it} \leq k_i}\{x_{it} \cdot R_i(q_{it}, \epsilon_{it}) - w_{it} \cdot q_{it}\}$$

is differentiable in k_i and

$$CM_{it}'(k_i|w_{it}, \epsilon_{it}) = x_{it} \cdot R_i'(q_{it}^o(k_i, \cdot), \epsilon_{it}) - w_{it}.$$

We can interchange the order of differentiation and integration to conclude that the firm's objective function $\Gamma_i(k_i)$ is differentiable with derivative:

$$NPV_i'(k_i) = \sum_{t=1}^{T} E_{\epsilon_{it}}[x_{it} \cdot R_i'(q_{it}^o(k_i, \cdot), \tilde{\epsilon}_{it}) - w_{it}] \cdot \gamma^t - v_i.$$

Given the definition of $\bar{w}_i \equiv \sum_{t=1}^{T} w_{it} \cdot \gamma^t$ in the statement of Lemma 3, the first-order condition in the statement of Lemma 3 now follows immediately. $\square$

Proof of Proposition 4:

Part (i): When $w_{2t} = w_t \cdot x_{2t}$, the firm's objective function regarding the downstream division is:

$$\sum_{t=1}^{T} E_{\epsilon_2}[CM_2(k_2|\tilde{\epsilon}_{2t}, w_{2t})] \cdot \gamma^t - v_2 \cdot k_2,$$

where

$$CM_2(k_2|\epsilon_{2t}, w_{2t}) = \max_{q_{2t} \leq k_2} \{R_2(q_{2t}, \epsilon_{2t}) - w_{2t} \cdot q_{2t}\} = R_2(q_{2t}^o(k_2, \cdot), \epsilon_{2t}) - w_{2t} \cdot q_{2t}^o(k_2)$$

If $w_{2t} = x_{2t} \cdot w_2$, the above objective function reduces to

$$\sum_{t=1}^{T} E_{\epsilon_2}[x_{2t} \cdot CM_2(k_2|\tilde{\epsilon}_{2t}, w_{2t})] \cdot \gamma^t - v_2 \cdot k_2$$

Given the transfer pricing rule

$$TP_t(k_2, q_{2t}) = \hat{c}_{2t} \cdot k_2 + w_{2t} \cdot q_{2t},$$

the expected profit for Division 2 in period t is given by

$$E_{\epsilon_2}[x_{2t} \cdot R_2(q_{2t}^o(k_2, \cdot), \tilde{\epsilon}_{2t}) - x_{2t} \cdot w_2 \cdot q_{2t}^o(k_2, \cdot)] - \hat{c}_{2t} \cdot k_2.$$

Since $\hat{c}_{2t} = \frac{x_{2t}}{\sum_{\tau=1}^{T} x_{2\tau} \cdot \gamma^\tau} \cdot v_2$, Division 2's objective function in period t is proportional to the firm's overall objective function, $NPV_2(k_2)$, and thus Division 2 will choose the optimal capacity level k_2^o at the initial stage.

Part (ii): When the limited condition volatility holds and ϵ_{it} are i.i.d., the optimal k_i^o is such that

$$
\begin{aligned}
NPV_i(k_i^o) &\equiv E_{\epsilon_i}\left[\sum_{\tau=1}^{T} \gamma^\tau \left[x_{i\tau} \cdot R_i(q_{i\tau}^o(k_i^o, \cdot), \tilde{\epsilon}_{i\tau}) - w_{i\tau} \cdot q_{i\tau}^o(k_i^o, \cdot)\right]\right] - v_i \cdot k_i^o \\
&= E_{\epsilon_i}\left[\sum_{\tau=1}^{T} \gamma^\tau \cdot x_{i\tau} \cdot R_i(k_i^o, \tilde{\epsilon}_{i\tau})\right] - (v_i + \bar{w}_i) \cdot k_i^o \\
&\geq NPV_i(k_i) \\
&\geq E_{\epsilon_i}\left[\sum_{\tau=1}^{T} \gamma^\tau \cdot x_{i\tau} \cdot R_i(k_i, \tilde{\epsilon}_{i\tau})\right] - (v_i + \bar{w}_i) \cdot k_i
\end{aligned}
\tag{37}
$$

for all k_i.

We next show that in case Division 2 is only concerned with its profit in period t, it will choose k_2^o. To see this, we recall that because $TP(k_2) = (\hat{c}_{2t} + \hat{z}_{2t}\bar{w}_2) \cdot k_2$, the expected profit for Division 2 in period t is given by

$$E_{\epsilon_2}[\tilde{\Pi}_{2t}(k_2)] = E_{\epsilon_2}[x_{2t} \cdot R_2(q_{2t}^o(k_2, \cdot), \tilde{\epsilon}_{2t})] - (\hat{c}_{2t} + \hat{z}_{2t}\bar{w}_2) \cdot k_2.$$

Clearly $q_{2t}^o(k_2, \cdot) = k_2$ as Division 2 is not charged any variable costs. We recall that $\hat{c}_{2t} = \hat{z}_{2t} \cdot v_2$ and $\hat{z}_{2t} = \frac{x_{2t}}{\sum_{\tau=1}^{T} x_{2t} \cdot \gamma^\tau}$ under the relative benefit depreciation rule, and therefore

$$E_{\epsilon_2}[\tilde{\pi}_{2t}(k_2)] = \hat{z}_{2t} \cdot \left[\sum_{t=1}^{T} x_{2\tau} \cdot \gamma^t \cdot E_{\epsilon_2}[R_2(k_2, \tilde{\epsilon}_{2t})] - (v_2 + \bar{w}_2) \cdot k_2 \right],$$

which, according to (37), is maximized at k_2^o. □

Proof of Proposition 5:

As defined in the main text, the threshold type θ_2^* is the one achieving a zero NPV for the capacity investment $\bar{k}_2$; that is,

$$\sum_{\tau=1}^{T} E_{\epsilon_2}\left[CM_{2\tau}(\bar{k}_2|x_{2\tau}, w_{2\tau}, \tilde{\epsilon}_{2\tau})|\theta_2^*\right] \cdot \gamma^\tau = v_2 \cdot \bar{k}_2.$$

We note that the net present value, $NPV(\bar{k}_2|\theta_i)$, is increasing in θ_2 because (i) θ_2 shifts the densities $f_2(\cdot|\theta_2)$ in the sense of first-order stochastic dominance, and (ii) $CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \epsilon_{2t})$ is increasing in ϵ_{2t} (Theorem 6D1,MasCollel et al., 1995). If the downstream division were to focus exclusively on its profit measure in period t, $1 \leq t \leq T$, it would seek to maximize:

$$E_{\epsilon_2}\left[\Pi_2(k|\theta_2, \tilde{\epsilon}_{2t})\right] \equiv E_{\epsilon_2}\left[CM_{2t}(k_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t})|\theta_2)\right] - \bar{z}_{2t} \cdot v_2 \cdot k_2.$$

Direct substitution for $\bar{z}_{2t}$ according to the REOB rule shows that

$$E_{\epsilon_2}\left[\Pi_2(k|\theta_2, \tilde{\epsilon}_{2t})\right] = E_{\epsilon_2}\left[CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t})|\theta_2\right] - E_{\epsilon_2}\left[CM_{2t}(\bar{k}_2|x_{2t}, w_{2t}, \tilde{\epsilon}_{2t})|\theta_2^*\right],$$

which will be greater than zero if and only if $\theta_2 > \theta_2^*$. □

Proof of Proposition 6:

Following the similar arguments as used in the proof of Proposition 3, it can be shown that the efficient capacity in the fungible setting is the same as in the dedicated setting; i.e., $k^* = k_1^o + k_2^o$.

To show that (k_1^o, k_2^o) is a Nash equilibrium at the initial date, we first prove that Division 2 will optimally choose k_2^o if it conjectures that the upstream divisions chooses k_1^o. Since Division 2 is not charged any variable cost, $q_2^o(k_2, \cdot) = k_2$. Therefore, Division 2's effective contribution margin after reallocation of capacity rights in period t is given by:

$$CM_{2t}^*(k_2|k_1^o, \theta, w_t, \epsilon_t) = \delta \cdot x_{2t} \cdot R_2(k_2, \epsilon_{2t}|\theta_2) + (1 - \delta) \cdot [CM_t(k_1^o + k_2|w_t, , \theta, \epsilon_t) - \Gamma_{1t}],$$

where

$$\Gamma_{1t} \equiv CM_{1t}(k_1^o|w_{1t}, \theta_1, \epsilon_{1t}) - w_{2t} \cdot k_2$$

and

$$CM_t(k_1^o + k_2|w_t, \theta, \epsilon_t) = \sum_{i=1}^{2} [x_{it} \cdot R_i(q_i^*(k_1^o + k_2, \cdot), \epsilon_{it}|\theta_i) - w_{it} \cdot q_i^*(k_1^o + k_2, \cdot)]$$

Given the full cost transfer pricing rule $TP_t(k_2) = \hat{c}_{2t} \cdot k_2 + \hat{z}_{2t} \cdot \bar{w}_2 \cdot k_2 = \hat{z}_{2t} \cdot (v + \bar{w}_2) \cdot k_2$, Division 2's profit in period t, contingent on k_1^o and ϵ_t, is:

$$\pi_{2t}(k_2|k_1^o, \epsilon_t) = CM_{2t}^*(k_2|k_1^o, \theta, w_t, \epsilon_t) - \hat{z}_{2t} \cdot (v + \bar{w}_2) \cdot k_2.$$

It will be convenient to rewrite the above expression as follows:

$$\pi_{2t}(k_2|k_1^o, \epsilon_t) = \delta \cdot \phi_{2t}(k_2|\epsilon_{2t}) + (1 - \delta) \cdot \psi_{2t}(k_2|k_1^o, \epsilon_t) + \alpha,$$

where α is a constant term independent of k_2,

$$\phi_{2t}(k_2|\epsilon_{2t}) = x_{2t} \cdot R_2(k_2, \epsilon_{2t}|\theta_2) - \hat{z}_{2t} \cdot (v + \bar{w}_2) \cdot k_2,$$

and

$$\psi_{2t}(k_2|k_1^o, \epsilon_t) = CM_t(k_1^o + k_2|w_t, \theta, \epsilon_t) + w_{2t} \cdot k_2 - \hat{z}_{2t} \cdot (v + \bar{w}_2) \cdot k_2.$$

To prove that k_2^o is an optimal response to k_1^o, it suffices to show that both $E_{\epsilon_2}[\phi_{2t}(k_2|\epsilon_{2t})]$ and $E_{\epsilon}[\psi_{2t}(k_2|k_1^o, \epsilon_{2t})]$ are maximized at $k_2 = k_2^o$ for each t. We recall that $\hat{z}_{2t} = \frac{x_{2t}}{\sum_{\tau=1}^{T} \gamma^\tau \cdot x_{2\tau}}$, and therefore

$$E_{\epsilon_2}[\phi_{2t}(k_2, \epsilon_{2t})] = \hat{z}_{2t} \cdot \left[\sum_{t=1}^{T} \gamma^t \cdot E_{\epsilon_2}[x_{2t} \cdot R_2(k_2, \epsilon_{2t}|\theta_2) - w_{2t} \cdot k_2] - v \cdot k_2 \right].$$

By definition of k_2^o, the above expression is indeed maximized at $k_2 = k_2^o$ for each t.

Differentiating $\psi_{2t}(k_2|k_1^o, \epsilon_{2t})$ with respect to k_2 yields

$$\psi_{2t}'(k_2|k_1^o, \epsilon_t) = S_t(k_1^o + k_2|w_t, \theta, \epsilon_t) + w_{2t} - \hat{z}_{2t} \cdot (v + \bar{w}_2).$$

Substituting $S_t(k_1^o + k_2|w_t, \theta, \epsilon_t) = x_{2t} \cdot R_2'(q_{2t}^*(k_1^o + k_2, \cdot), \epsilon_{2t}|\theta_2) - w_{2t}$ and noting that $\hat{z}_{2t} = \frac{x_{2t}}{\sum_{\tau=1}^{T} \gamma^\tau \cdot x_{2\tau}}$ and $\bar{w}_2 = \sum_{t=1}^{T} \gamma^t \cdot w_{2t}$, we get:

$$E_\epsilon[\psi_{2t}'(k_2|k_1^o, \epsilon_t)] = \hat{z}_{2t} \cdot \left[\sum_{t=1}^{t} \gamma^t \cdot E_\epsilon \left[x_{2t} \cdot R_2'(q_{2t}^*(k_1^o + k_2, \cdot), \epsilon_{2t}|\theta_2) - w_{2t} \right] - v \right]$$

$$= \hat{z}_{2t} \cdot \left[\sum_{t=1}^{t} \gamma^t \cdot E_\epsilon[S_t(k_1^o + k_2|w_t, \theta, \epsilon_t)] - v \right].$$

We recall that the efficient capacity level k^* is given by the first-order condition:

$$\sum_{t=1}^{t} \gamma^t \cdot E_\epsilon[S_t(k^*|w_t, \theta, \epsilon_t)] - v = 0.$$

Since $k^* = k_1^o + k_2^o$,

$$E_\epsilon[\psi_{2t}'(k_2|k_1^o, \epsilon_t)] = 0$$

at $k_2 = k_2^o$, and hence $E_\epsilon[\psi_{2t}(k_2|k_1^o, \epsilon_t)]$ is maximized at k_2^o. This proves that k_2^o is an optimal response to k_1^o.

A similar argument can be used to show that the upstream division will optimally choose k_1^o when it conjectures that the downstream division chooses k_2^o. This proves that (k_1^o, k_2^o) is a Nash equilibrium. $\qquad\square$

Appendix B

In this Appendix, we demonstrate through examples that the full cost transfer pricing mechanism in part (ii) of Proposition 4 will generally lead the downstream division to underinvest if the limited volatility condition does not hold. We also show that the amount of underinvestment and corresponding efficiency losses increase with the degree to which the limited volatility conditions fails to hold. Thus, the amount of underinvestment and the efficiency loss both increase as the probability density function of the revenue shocks ϵ_{it} becomes more spread out.

Consider a stationary setting in which the downstream division's revenue functions take the following form:

$$R_2(q_2, \epsilon_{2t}) = \epsilon_{2t} \cdot q_{2t} - \frac{h}{2} \cdot q_{2t}^2$$

with $h > 0$. Suppose the revenue shocks ϵ_{2t} are distributed with support over the interval $[\underline{\epsilon}_2, \bar{\epsilon}_2]$ with $\underline{\epsilon}_2 > w_2$. For a given capacity k_2, the efficient quantity choice is given by

$$q_2^o(k_2, \epsilon_{2t}) = \min\left\{\frac{\epsilon_{2t} - w_2}{h}, k_2\right\}.$$

Lemma 1 then yields the following expression for the efficient capacity level k_2^o:

$$E\left[\tilde{\epsilon}_{2t} - h \cdot \min\left\{\frac{\tilde{\epsilon}_{2t} - w_2}{h}, k_2^o\right\}\right] = w_2 + c_2.$$

We note that the limited volatility condition will hold if and only if $\frac{\tilde{\epsilon}_{2t} - w_2}{h} \geq k_2^o$ for all $\tilde{\epsilon}_{2t} \in [\underline{\epsilon}_2, \bar{\epsilon}_2]$. This is equivalent to the condition that

$$\underline{\epsilon}_2 \geq h \cdot k_2^o + w_2. \tag{38}$$

Now suppose the limited volatility condition does not hold, but the two-part full cost mechanism in part (ii) of Proposition 4 is still used for transfer pricing purposes. The downstream division will then produce at the capacity level for all realizations of the revenue shock ϵ_{2t} because it is not charged for actually incurred variable costs. Consequently, this transfer pricing mechanism will induce Division 2 to choose its

capacity so as to maximize $E(\tilde{\epsilon}_{2t}) \cdot k_2 - \frac{h}{2} \cdot k_2^2 - w_2 \cdot k_2 - c_2 \cdot k_2$. The first order condition yields that the induced capacity level k_2^I is given by

$$k_2^I = \frac{E(\tilde{\epsilon}_{2t}) - (w_2 + c_2)}{h}.$$

We note that the induced capacity level k_2^I depends on the distribution of the revenue shocks only through the expected value $E(\tilde{\epsilon}_{2t})$.

Suppose now the revenue shocks $\tilde{\epsilon}_{2t}$ are uniformly distributed over the interval $[m_2 - \sigma_2, m_2 + \sigma_2]$ with $\sigma_2 \geq 0$. As σ_2 increases, the variance of $\tilde{\epsilon}_{2t}$ increases but its expected value remains unchanged at m_2. The revenue volatility can therefore be conveniently parameterized by σ_2. Figure 5 below plots the efficient and induced capacity levels as functions of the revenue volatility parameter σ_2 when $c_2 = w_2 = 1$, $m_2 = 3$, and $h = 0.01$.
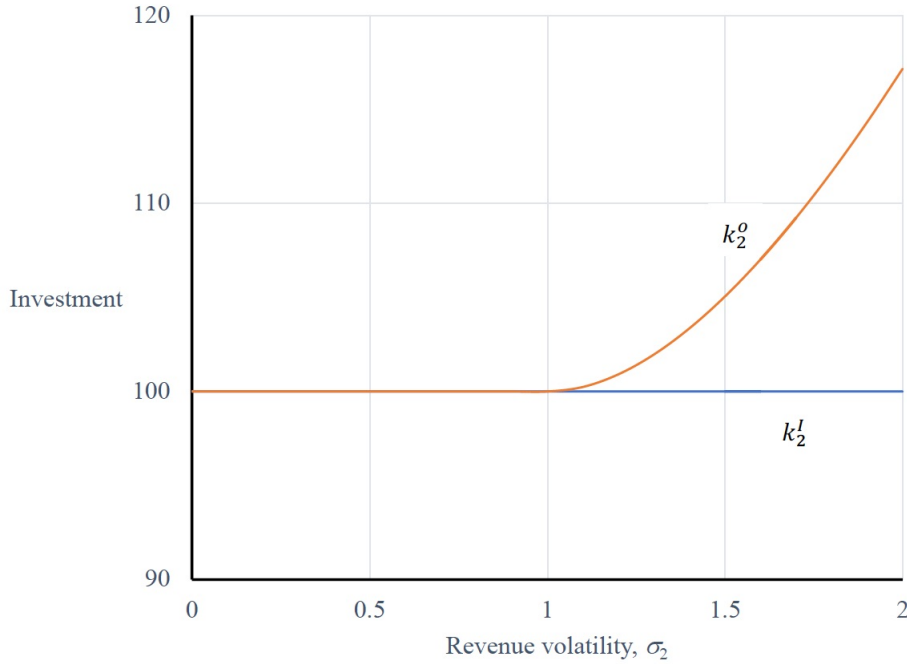

Figure 5: **Investments and revenue volatility**

For these parameter values, it can be verified that the limited volatility condition holds if and only if $\sigma_2 \leq 1$.[25] Figure 5 shows that the full cost transfer pricing

---

[25]Equation (38) yields that the limited volatility condition holds if and only if $m_2 - \sigma_2 \geq h \cdot k_2^o + w_2$. This simplifies to $\sigma_2 \leq c_2 = 1$, since $k_2^o = \frac{m_2 - (w_2 + c_2)}{h}$.

mechanism in part (ii) of Proposition 4 induces the efficient capacity choice (i.e., $k_2^I = k_2^o$) when the limited volatility condition holds (i.e., $\sigma_2 \leq 1$). The downstream division underinvests (i.e., $k_2^I < k_2^o$) when the limited volatility condition fails to hold (i.e., $\sigma_2 > 1$.). The amount of underinvestment increases in the revenue volatility parameter σ_2, and becomes arbitrarily small as σ_2 approaches one from above.

Figure 6 below plots the expected profit per period generated by the full cost transfer pricing mechanism, π_2^I, and the first-best expected profit per period, π_2^o. It confirms that the efficiency loss from using the full cost transfer pricing mechanism, $\pi_2^o - \pi_2^I$, increases in the degree to which the limited volatility condition fails to hold (i.e., the degree to which σ_2 exceeds the threshold value of one).
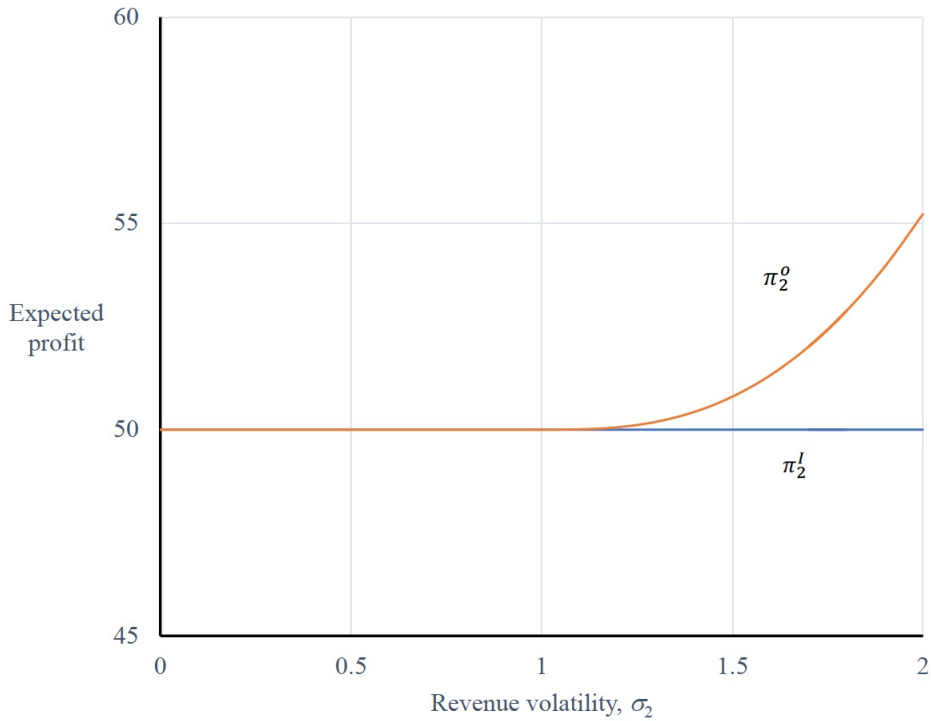

Figure 6: **Expected profits and revenue volatility**

References

Anctil, R. and S. Dutta (1999). Negotiated Transfer Pricing and Divisional vs. Firm-Wide Performance Evaluation. *The Accounting Review* 74(1), 87-104.

Arrow, K. (1964). Optimal Capital Policy, The Cost of Capital and Myopic Decision Rules. *Annals of the Institute of Statistical Mathematics* 1-2, 21-30.

Arya, A., J. Demski, J. Glover, and P. Liang (2009), "Quasi-Robust Multiagent Contracts," *Management Science* 55(5), 752-762.

Arya, A., and B. Mittendorf (2011), "Supply Chains and Segment Profitability: How Input Pricing Creates a Latent Cross-Segment Subsidy," *The Accounting Review*, 86(3) 805-824.

Balachandran, B., Balakrishnan, R. and S. Sivaramakrishnan (1997). On the Efficiency of Cost-Based Decision Rules for Capacity Planning. *The Accounting Research* 72(4), 599-619.

Balakrishnan, R. and S. Sivaramakrishnan (2002). A Critical Overview of Full-Cost Data for Planning and Pricing. *Journal of Management Accounting Research* 3-31.

Baldenius, T., Dutta, S. and S. Reichelstein (2007). Cost Allocations for Capital Budgeting Decisions. *The Accounting Review* 82(4), 837-867.

Baldenius, T., Nezlobin, A., and I. Vaysman (2016). Managerial Performance Evaluation and Real Options. *The Accounting Review* 91(3), 741-766.

Baldenius, T. and S. Reichelstein, and S. Sahay (1999). Negotiated versus Cost-Based Transfer Pricing. *The Review of Accounting Studies* 4(2), 67-91.

Banker, R. and J. Hughes (1994). Product Costing and Pricing. *The Accounting Review* 69(3), 479-494.

Bastian, N. and S. Reichelstein (2004). Transfer Pricing at Timken. Stanford GSB, Case # A-190.

Bockem, S. and U. Schiller (2008). Full-Cost Transfer Pricing and Cost Management. *Journal of Economics and Management Strategy* 28(3), 63-81.

Bouwens, J and B. Steens (2016). Full-Cost Transfer Pricing and Cost Management. *Journal of Management of Accounting Research* 17(1), 219-245.

Carlton, S., & Perloff, J. (2005). *Advanced Industrial Organization.* 4th ed. New

York, NY: Pearson/Addison Wesley.

Datar, S. and M. Rajan. (2014). *Managerial Accounting: Making Decisions and Motivating Performance.* Pearson, NY: Pearson/Addison Wesley.

Dutta, S. (2008). Dynamic Performance Measurement. *Foundations and Trends in Accounting, 2*(3), 175-240.

Dutta, S., and S. Reichelstein (2002). Controlling Investment Decisions: Depreciation and Capital Charges. *Review of Accounting Studies* 7, 253-281.

Dutta, S. and S. Reichelstein (2010). Decentralized Capacity Management and Internal Pricing. *Review of Accounting Studies, 15*(3), 442-478.

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). Negotiated Transfer Pricing: An Efficiency Result. *The Accounting Review* 69(3), 479-494.

Ernst & Young (2003), *Global Transfer Pricing Survey.* www.ey.com.

Feinschreiber, R. and M. Kent. (2012). *Transfer Pricing Handbook.* Wiley Publishers.

Goex, R. (2002), Capacity Planning and Pricing under Uncertainty. *Journal of Management Accounting Research* 59-79.

Goex, R. and U. Schiller (2007). An Economic Perspective on Transfer Pricing. *Handbook of Management Accounting Research.* C.S. Chapman, A. Hopwood and M. Shields

Gramlich, J. and K. Ray (2016). Reconciling full-Cost and Marginal-Cost Transfer Pricing. *Journal of Management Accounting Research* 28(1), 27-37.

Hotelling, H. (1925). A General Mathematical Theory of Depreciation. *Journal of the American Statistical Association* 20, 340-353.

Kaplan, R. (2006). Activity-based Costing and Capacity. *Harvard Business School*, Case # 9-105-059

Marschak, T. and S. Reichelstein (1998). Network Mechanisms and Informationally Efficient Hierarchies. *Journal of Economic Theory* 79, 106-141.

Martinez-Jerez, A. (2007). Understanding Customer Profitability at Charles Schwab. *Harvard Business School*, Case # 106-002.

Mas-Colell, A., Whinston, M. and J. Green (1995), *Microeconomic Theory, Oxford*

University Press, Oxford, New York.

Mount, K. and S. Reiter (1974). The Informational size of Message Spaces. *Journal of Economic Theory* 8, 161-192.

Nezlobin, A., Rajan, M and S. Reichelstein (2012). Dynamics of Rate of Return Regulation. *Management Science* 58(5), 980-995.

Oyama, D. and T. Takenawa (2018). On the (non-)differentiability of the optimal value function when the optimal solution is unique. *Journal of Mathematical Economics* 76, 21-32.

Pfeiffer, T., Lengsfeld, S., Schiller, U. and J. Wagner (2009). Cost-based Transfer Pricing. *Review of Accounting Studies* 7, 253-281.

Pfeiffer, T. and G. Schneider (2007), "Residual Income-Based Compensation Plans for Controlling Investment Decisions Under Sequential private Information," *Management Science* 53(3), 495-507.

Rajan, M. and S. Reichelstein (2009). Depreciation Rules and The Relation Between Marginal and Historical Cost. *Journal of Accounting Research* 47(3),1-43.

Reichelstein, S. and A. Rohlfing-Bastian (2015). Levelized Product Cost: Concept and Decision Relevance. *The Accounting Review* 90 (4), 1653-1682.

Reichelstein, S. and A. Sahoo (2018). Relating Long-Run Marginal Cost to Product Prices: Evidence from Solar Photovoltaic Modules. *Contemporary Accounting Research* 35(3), 1464 - 1498.

Rogerson, W. (1997) Inter-Temporal Cost Allocation and Managerial Investment Incentives: A Theory Explaining the Use of Economic Value Added as a Performance Measure. *Journal of Political Economy* 105, 770-795.

Rogerson, W. (2008). Inter-Temporal Cost Allocation and Investment Decisions. *Journal of Political Economy* 116, 931-950.

Rogerson, W. (2011). On the relationship between historic cost, forward looking cost and long run marginal cost. *Review of Network Economics* 10(2), 1-31.

Sahay, S. (2002). Transfer Pricing Based on Actual Cost. *Journal of Management Accounting Research* 15, 177-192.

Simons, R. (2000). Polysar Limited. *Harvard Business School* Case # 9-187-098.

Solomons, D. (1964), *Divisional Performance Measurement and Control. Irvin.* Home-

wood, Illinois.

Spear, S. (2003). The Electricity Market Game. *Journal of Economic Theory* 109(2), 300-323.

Tang, R. (2002). *Current Trends and Corporate Cases in Transfer Pricing.* Westport, CT: Quorum Books.

Vaysman, I. (1996), "Cost-based Transfer Pricing," *Review of Accounting Studies* Vol.1, 73-108.

Wielenberg, S. (2000). Negotiated Transfer Pricing, Specific Investment and Optimal Capacity Choice. *Review of Accounting Studies* 5, 197-216.

Young, D and O'Byrne, J. (2000), *EVA and Value-Based Management, McGraw Hill*, New York.

Zangwill, W. (1969), *Nonlinear Programming. McGraw Hill*, New York.

Zimmerman, J. (2016), *Accounting for Decision Making and Control.*, 8th ed., *Prentice Hall*, New York.